Title: Reading materials
RDA Data Models Meeting - 30 April-1 May 2007

-- RDA Data Models Meeting wiki page
2 http://dublincore.org/librarieswiki/DataModelMeeting?action=print

---

RDA

-- The RDA Strategic Plan
5 http://www.collectionscanada.ca/jsc/stratplan.html

-- RDA Objectives and Principles
8 http://www.collectionscanada.ca/jsc/docs/5rda-objectives.pdf

-- RDA Scope and Structure
13 http://www.collectionscanada.ca/jsc/docs/5rda-scope.pdf

-- RDA Implementation Scenarios
24 http://www.collectionscanada.ca/jsc/docs/5editor2.pdf

-- Encoding RDA data
29 e:/u/rdameeting/attachments/Encoding-RDA-Data.pdf

-- RDA-Scope analysis note
34 e:/u/rdameeting/attachments/RDA-Scope-analysis-note.pdf

-- RDA-Scope analysis table
35 e:/u/rdameeting/attachments/RDA-Scope-analysis-table.pdf

DCMI, Semantic Web, SKOS

-- Framework for a Bibliographic future (Diane Hillmann et al)
56 e:/u/rdameeting/Framework.pdf
(captured from futurelib.pbwiki.com/Framework)

-- Towards an interoperability framework for metadata standards (Nilsson et al)
62 http://kmr.nada.kth.se/~mini/papers/TowardsAFramework.pdf

-- Eprints Application Profile (Andy Powell et al)
76 http://www.ariadne.ac.uk/issue50/allinson-et-al/

-- Quick Guide to Publishing a Thesaurus on the Semantic Web
86 http://www.w3.org/TR/2005/WD-swbp-thesaurus-pubguide-20050517/

RDA Data Models Meeting - 30 April-1 May 2007

Participants

Tom Baker	Chair of Usage Board, Director Specifications and Documentation, Co-chair DC-Architecture, Dublin Core Metadata Initiative; Co-chair, W3C Semantic Web Deployment Working Group; Co-chair, KIM Working Group (Germany)
Tom Delsey	RDA Editor. Consultant to: IFLA Working Group on Functional Requirements for Bibiliographic Records -- FRBR (1992-1997); AGICOA, CISAC, FIAPF, IFPI, International ISBN Agency, ISO TC46/SC9/WG3, and the ISSN International Centre -- Content Development and Rights Management (ISO Technical Report 21449) (2001); Library of Congress, Networks and MARC Standards Office -- Functional Analysis of the MARC 21 Bibliographic and Holdings Formats (2001); IFLA UBCIM Working Group on Functional Requirements and Numbering for Authority Records -- FRAD (2002-2003); IFLA ISBD Review Group -- Mapping ISBD Elements to FRBR Attributes and Relationships (2004); Library of Congress -- Defining an Access Level MARC/AACR Catalog Record (2004).
Gordon Dunsire	Centre for Digital Library Research, Strathclyde University. Member, CILIP-BL Committee on AACR; Member, RDA Outreach Group; Chair, Cataloguing and Indexing Group in Scotland; Member, Finish Group for DC AP for Collections; Contributor, DC AP for Eprints
Diane Hillmann	Dublin Core Metadata Initiative Liaison to ALA CC-DA; Member, Dublin Core Usage Board and Advisory Board; Co-Coordinator, DC-Education Community; Editor, "Using Dublin Core"; Administrator, AskDCMI; Research Librarian, Cornell University
Alistair Miles	Rutherford Appleton Laboratory, UK; Editor, Simple Knowledge Organisation System (SKOS); Editor, "Best Practice Recipes for Publishing RDF Vocabularies"; Member, W3C Semantic Web Deployment Working Group
Mikael Nilsson	KTH Royal Institute of Technology, Sweden, Co-author of DCMI Abstract Model and DC-in-RDF specifications, co-chair of DC-Architecture and of Joint DCMI-IEEE LTSC Task Force (for expressing IEEE LOM using the DCAM)
Andy Powell	Eduserv Foundation; Co-editor, Eprints Application Profile, Co-author, DCMI Abstract Model
Barbara Tillett	Joint Steering Committee for Revision of AACR
Robina Clayphan	Co-ordinator of Bibliographic and Metadata Standards at The British Library; Moderator of DCMI Libraries Community; member of DCMI Advisory Board; member of CILIP-BL Committee on AACR; member of RDA Outreach Working Group.

Documents

Agenda

Background

IFLA Study Group on the Functional Requirements of Bibliographic Records
FRBR

Functional Requirements for Authority Data
FRANAR

See also http://www.loc.gov/cds/downloads/FRBR.PDF

IME ICC draft Statement of International Cataloguing Principles

RDA

The RDA Strategic Plan

RDA Objectives and Principles

Tom Delsey
RDA - Resource Description and Access: Scope and Structure
RDA Implementation Scenarios

Encoding RDA data

RDA-Scope analysis note

RDA-Scope analysis table

RDA-DCLib mapping

Discussion material

Karen Coyle, Diane Hillmann
RDA, Cataloguing Rules for the 20th Century

Karen Coyle, Diane Hillmann, Jonathan Rochkind, Paul Weiss
Framework for a Bibliographic future
Demonstration of Vocabulary registration using RDA Carriers -- see explanation in Karen Coyle blog post

Dublin Core documents

Andy Powell, Mikael Nilsson et al
DCMI Abstract Model, 2 April revision

Julie Allinson, Pete Johnston and Andy Powell
A Dublin Core Application Profile for Scholarly Works

Nilsson et al. DC2007, Towards an interoperability framework for metadata standards (Presentation)

SKOS documents

Alistair Miles

SKOS: Requirements for Standardisation, DCMI Conference 2006 [presentation] [paper]

Alistair Miles and Dan Brickley eds.
SKOS Core Guide, W3C Working Draft 2 November 2005

Alistair Miles ed.
Quick Guide to Publishing a Thesaurus on the Semantic Web, W3C Working Draft 17 May 2005

For information

Working Group on the Future of Bibliographic Control
Link to papers from the meetings of Bibliographic Control Working Group

Gordon Dunsire
Distinguishing Content from Carrier: The RDA/ONIX Framework for Resource Categorization
http://www.dlib.org/dlib/january07/dunsire/01dunsire.html

Overview
News and Announcements
RDA
Submitting Proposals to Revise AACR
Strategic Plan
JSC Members
Annual Reports
Working Documents
Historic Documents
Links to Related Resources

Joint Steering Committee for Revision of Anglo-American Cataloguing Rules

Strategic Plan for RDA 2005-2008

To print, use the PDF version.

29 May 2006

A Strategic Plan was first developed for AACR in 2002, and revised in 2003 and 2004.

The current Strategic Plan has been comprehensively revised to reflect goals and strategies for the development of *RDA - Resource Description and Access* and is reviewed annually.

The Joint Steering Committee and the Committee of Principals share overall responsibility for the plan. The JSC, the CoP, the Co-publishers, the RDA Project Manager, the RDA Editor and the JSC Chair each have responsibilities in relation to specific aspects of this plan.

STATEMENT OF PURPOSE FOR RDA

RDA - Resource Description and Access will be a new standard for resource description and access, designed for the digital world.

Built on foundations established by the *Anglo-American Cataloguing Rules* (AACR), RDA will provide a comprehensive set of guidelines and instructions on resource description and access covering all types of content and media.

RDA will enable users of library catalogues and other systems of information organization to find, identify, select, and obtain resources appropriate to their information needs.

Return to Top of Page

GOALS FOR RDA

1. The guidelines and instructions in RDA will be designed to:

- Provide a consistent, flexible and extensible framework for both the technical and content description of all types of resources and all types of content.
- Be compatible with internationally established principles, models, and standards.
- Be usable outside the library community, and be capable of adaptation by various communities to their specific needs.

2. Descriptions and access points produced through the application of RDA guidelines and instructions will:

- Enable users to find, identify, select, and obtain resources appropriate to their information needs.
- Be compatible with those descriptions and access points devised using AACR2, and present in

existing catalogues and databases.
- Be independent of the format, medium or system used to store or communicate the data.
- Be readily adaptable to newly-emerging database structures.

3. RDA will be developed as a resource description standard that is:

- Optimised for use as an online tool (although a print edition will also be published).
- Derived from English language conventions and customs, written in plain English, and able to be used in other language communities.
- Easy and efficient to use, both as a working tool and for training purposes.

Return to Top of Page

STRATEGIES FOR ACHIEVING THE GOALS

1. To ensure that the guidelines and instructions in RDA meet the goals for consistency, flexibility, etc., outlined under Goal 1 above, we will:

- Resolve problems associated with the class of materials concept and the related issue of GMDs.
- Resolve residual problems associated with cataloguing resources that change over time, including multipart resources.
- Revise and modernise terminology to eliminate unnecessary library jargon, and eliminate ambiguous and inconsistent terminology.
- Extend consistency within the technical and content description, and eliminate unnecessary redundancy.
- Review and update the examples throughout RDA.
- Include a statement of the principles of RDA, and a description of the functions of the catalogue, and relate these to the *Statement of International Cataloguing Principles*.
- Align the structure, concepts and terminology of the instructions more directly with the *Functional Requirements for Bibliographic Records* (FRBR) and *Functional Requirements for Authority Records* (FRAR) models.
- Improve the coverage of materials such as digital resources, three-dimensional objects, visual materials, manuscripts and archives.
- Identify and actively involve other resource description communities and other stakeholders in the development of RDA in accordance with the RDA communication plan.

2. To ensure that descriptions and access points produced through the application of RDA meet the goals for functionality, compatibility, etc., outlined under Goal 2 above, we will:

- Directly relate the elements of the description and access points to the user tasks that they support.
- Revise the instructions to facilitate collocation at the FRBR work and expression levels.
- Incorporate the concept of authority control.
- Revise the instructions relating to the choice of access points to address issues associated with the concept of “authorship” as it is currently reflected in AACR2 and restrictions imposed by the “rule of three”.
- Separate instructions on the recording of data from those on the presentation of data.

3. To ensure that RDA is developed as a resource description standard that meets the goals for ease and efficiency of use outlined under Goal 3 above, we will:

- Work with an online product developer to develop an online tool according to functional specifications that reflect how RDA will be used.
- Provide conceptual background within RDA to assist the user of the instructions to understand the process to be followed.
- Write new text in plain English, and revise existing text wherever practicable.
- Provide a revised Glossary, and incorporate definitions into the text where needed.
- Consult with educators and trainers, and plan for activities that support the implementation of RDA.

4. To ensure that RDA is ready for publication in 2008, we will:

- Develop, implement and maintain a Business Plan.
- Establish an effective project management structure, including the appointment of a project manager.
- Balance community input with the need for RDA to be ready for publication within the timeframe.
- Employ appropriate risk management processes.

Return to Top of Page

`http://www.collectionscanada.ca/jsc/stratplan.html`
Last updated: 19 June 2006
Webmaster: Nathalie Schulz, JSC Secretary

To: Joint Steering Committee for Revision of AACR

From: Deirdre Kiorgaard, Chair, JSC

Subject: Draft statement of objectives and principles for RDA

[Draft]

RDA — Resource Description and Access

Objectives and Principles

The draft statement of objectives and principles for RDA that follows reflects discussions that have been conducted to date within the Joint Steering Committee. Those discussions have taken into account comments received on the earlier version of the statement that was issued for constituency review in December 2004 with the draft of AACR3 Part I.

The statement is still in draft form. While there have been several rounds of discussion on the objectives and principles as they relate to Part I of RDA, those that relate to Part II have only been discussed in a preliminary way, and will require further review by JSC once a first draft of Part II has been completed. The draft does not yet reflect specific objectives and principles relating to Part III; those will be discussed and added as work on preparing the first draft of Part III begins in the second half of 2006.

It should be noted as well that while the statement of objectives and principles serves to provide overall guidance for the development of RDA, trade-offs sometimes have to be made between one principle and another. For example, the principle of uniformity needs to be balanced with the principle of common usage, the principle of accuracy needs to be balanced with the principle of representation, etc.

1. Objectives and Principles for the Design of RDA

The objectives and principles set out in this section are those that govern the overall design of *RDA* as a standard for resource description and access. They address matters of scope, formulation, currency, etc.

Objectives

Comprehensiveness

The guidelines and instructions should cover all types of resources and all types of content represented in catalogues or similar tools.

Consistency

The guidelines and instructions should be consistent in their formulation.

Clarity

The guidelines and instructions should be clear and unambiguous with respect to underlying concepts, terminology, and scope of application. They should be written in plain English.

Rationality

The guidelines and instructions should reflect rational, non-arbitrary decisions.

Currency

The guidelines and instructions should be responsive to new developments affecting the range, nature, and characteristics of the resources and types of content covered, and to the emergence of new types of resources and content.

Compatibility

The guidelines and instructions should be compatible with internationally established principles, models, and standards.

Adaptability

The guidelines and instructions should be amenable to adaptation by various communities to meet their specific needs.

Ease and efficiency of use

The guidelines and instructions should be easy and efficient to use.

Format

The guidelines and instructions should be amenable to presentation in either a conventional print format or in a digital format embodying features such as hypertext links, selective display, etc.

Principles

Generalization

For each element of the description the guidelines and instructions should provide substantive basic instructions that are applicable to all types of resources.

The guidelines and instructions on relationships should provide substantive basic instructions that are applicable to all types of content.

Specificity

Where required, the guidelines and instructions should provide supplementary or special instructions applicable to a specific type of content, medium, mode of issuance, etc.

Non-redundancy

The guidelines and instructions should avoid unnecessary repetition.

Terminology

The terminology used in the guidelines and instructions should be consistent with the concepts and terms defined in *Functional Requirement for Bibliographic Records* and *Functional Requirements for Authority Records.*

Reference structure

The reference structure embodied in the guidelines and instructions should be consistent in its overall design and serve as an efficient means of indicating the location of additional guidelines and instructions that may be relevant to the application of the guideline or instruction from which the reference is made.

2. Functionality of Records Produced Using RDA

The objectives and principles set out in this section are those that relate to the functionality of the descriptions and access points produced through the application of *RDA* guidelines and instructions. They address matters of responsiveness to user needs, cost efficiency, etc.

Objectives

Responsiveness to user needs

The descriptive data provided for in the guidelines and instructions should enable the user to:

- *identify* the resource described (i.e., to confirm that the resource described corresponds to the resource sought, or to distinguish between two or more resources with similar characteristics);
- *select* a resource that is appropriate to the user's requirements with respect to content, format, etc.

The access points provided for in the guidelines and instructions should enable the user to locate:

- all resources described in the catalogue that embody a particular work or a particular expression of that work;
- all resources described in the catalogue that embody works and expressions of works associated with a particular person, family, or corporate body;
- a specific resource described in the catalogue that is searched under a title appearing in that resource;
- works, expressions of works, and manifestations represented in the catalogue that are related to those retrieved in response to the user's search.

Cost efficiency

The descriptive data and access points provided for in the guidelines and instructions should meet functional requirements in a cost-efficient manner.

Flexibility

The descriptive data and access points provided for in the guidelines and instructions should function independently of the format, medium, or system used to store or communicate the data. They should be amenable to use in a variety of environments.

Continuity

The descriptive data and access points provided for in the guidelines and instructions should be amenable to integration into existing files (particularly those developed using *AACR* and related standards) with a minimum of retrospective adjustment to those files.

Principles

Differentiation

The descriptive data provided for in the guidelines and instructions should serve to differentiate the resource described from other resources represented in the file.

The access point data provided for in the guidelines and instructions should serve to differentiate the person, family, corporate body, work, expression, or manifestation represented from other entities represented in the file.

Sufficiency

The descriptive data provided for in the guidelines and instructions should be sufficient to meet the needs of the user with respect to selection of an appropriate resource.

Relationships

The descriptive data provided for in the guidelines and instructions should indicate significant bibliographic relationships between the resource described and other resources.

The access point data provided for in the guidelines and instructions should reflect all significant bibliographic relationships between works, expressions, manifestations, persons, families, and corporate bodies.

Representation

The descriptive data provided for in the guidelines and instructions should reflect the resource's representation of itself.

Accuracy

The descriptive data provided for in the guidelines and instructions should furnish supplementary information to correct or clarify ambiguous, unintelligible, or misleading representations made by the resource itself.

Attribution

The guidelines and instructions for providing access points should reflect attributions of responsibility made either in manifestations embodying the work or expression or in reference sources, irrespective of whether the attribution of responsibility is accurate.

Common usage

The guidelines and instructions for recording data elements other than those transcribed from the resource itself should reflect common usage.

Common practice

The guidelines and instructions for citing works, expressions, and manifestations should follow common practice.

Uniformity

The guidelines and instructions provided in the appendices on capitalization, numerals, abbreviations, order of elements, punctuation, etc., should serve to promote uniformity in the presentation of descriptive and access point data.

7 December 2006

1

To: Joint Steering Committee for Revision of AACR

From: Deirdre Kiorgaard, Chair, JSC

Subject: RDA Scope and Structure

RDA — Resource Description and Access

Scope and Structure

This document is one of three that define the framework for the development of *RDA*. The *RDA Strategic Plan* establishes long-term goals for *RDA* and the strategies for achieving those goals in the period 2005-2008. The *RDA Objectives and Principles* document sets out the objectives and principles that govern the overall design of *RDA* as well as objectives and principles relating to the functionality of the data produced through the application of *RDA*. This document defines the scope and structure of *RDA* in relation to its underlying conceptual models (*FRBR*[1] and *FRAD*[2]) and to two related metadata models (the *DCMI Abstract Model*[3] and *The <indecs> Metadata Framework*[4]).

1. Scope

RDA provides a set of guidelines and instructions on formulating descriptive data and access point control data to support resource discovery.

1.1 Definitions

For purposes of defining the scope of *RDA*, the terms *resource*, *resource discovery*, *descriptive data*, and *access point control data* are defined as follows:

Resource

A *resource* is an identifiable information object, either tangible or intangible in nature.

---

[1] IFLA Study Group on the Functional Requirements for Bibiographic Records, *Functional Requirements for Bibliographic Records: Final Report* (München: K.G. Saur, 1998). Cited hereafter as *FRBR*.

[2] IFLA Working Group on Functional Requirements and Numbering of Authority Records, *Functional Requirements for Authority Data: A Conceptual Model* (Draft 2006-08-15). Cited hereafter as *FRAD*.

[3] Andy Powell, Mikael Nilsson, Ambjorn Naeve, and Peter Johnston, DCMI Abstract Model (2005-03-07). Cited hereafter as *DCMI Abstract Model*.

[4] Godfrey Rust and Mark Bide, The <indecs> Metadata Framework: Principles, Model and Data Dictionary (June 2000). Cited hereafter as *Indecs*.

Resource discovery

Resource discovery encompasses the following generic user tasks:[5]

- *FIND* — i.e., to find resources that correspond to the user's stated search criteria
- *IDENTIFY* — i.e., to confirm that the resource described corresponds to the resource sought, or to distinguish between two or more resources with similar characteristics
- *SELECT* — i.e., to select a resource that is appropriate to the user's needs
- *OBTAIN* — i.e., to acquire or access the resource described

Descriptive data

Descriptive data are data (i.e., value representations[6]) that reflect the characteristics (i.e., properties[7]) of a resource.

Access point control data

Access point control data are data (i.e., value representations) that reflect the characteristics (i.e., properties) of an entity represented by a controlled access point.[8]

1.2 Descriptive data

The descriptive data covered in *RDA* reflect the attributes and relationships associated with the entities *work*, *expression*, *manifestation*, and *item*, as defined in *FRBR*.[9]

The scope of descriptive data covered in *RDA* may be extended in future releases to cover additional attributes and relationships associated with the entities *work*, *expression*, *manifestation*, and *item* not currently defined in *FRBR* that support resource discovery.

Attributes and relationships associated with the entities *work*, *expression*, *manifestation*, and *item* whose primary function is to support user tasks related to resource management (e.g., acquisition, preservation) and data management (e.g., data processing, data display) are currently out of scope.

Attributes and relationships associated with the entities *concept*, *object*, *event*, and *place*, as defined in *FRBR*, fall outside the scope of *RDA*. Subject relationships, as defined in *FRBR*, are also out of scope.

1.3 Access point control data

The access point control data covered in *RDA* reflect the attributes and relationships associated with the entities *person*, *family*, *corporate body*, *place*, *work*, *expression*, *manifestation*, and *item*, as defined in *FRAD*.[10]

---

[5] Based on the user tasks defined in *FRBR*, p.82.
[6] The term *value representation* is used as defined in the *DCMI Abstract Model*, p. 8
[7] The term *property* is used as defined in the *DCMI Abstract Model*, p. 7
[8] In this context, a controlled access point is a name formulated according to the guidelines and instructions in part B of *RDA*.
[9] See the attributes defined in sections 4.2-4.5 and the relationships defined in sections 5.2-5.3 of *FRBR*.

Attributes associated with the entities *name*, *identifier*, *controlled access point*, and *rules*, as defined in *FRAD*, are covered on a selective basis.

The scope of access point control data covered in *RDA* may be extended in future releases to cover additional attributes and relationships associated with the entities *person*, *family*, *corporate body*, *place*, *work*, *expression*, *manifestation*, *item*, *name*, *identifier*, *controlled access point*, and *rules* not currently defined in *FRAD* that support resource discovery.

Attributes and relationships associated with the entities *person*, *family*, *corporate body*, *place*, *work*, *expression*, *manifestation*, *item*, *name*, *identifier*, *controlled access point*, and *rules* whose primary function is to support user tasks associated with data management (e.g., data processing, data display) are currently out of scope.

Attributes and relationships associated with the entities *concept*, *object*, and *event*, as defined in *FRAD*, fall outside the scope of *RDA*. Relationships between *controlled access points*, as defined in *FRAD*, are also out of scope.

1.4 Elements

Attributes and relationships associated with a resource or other entity are formally represented in *RDA* as elements.

Each RDA element corresponds to an attribute or relationship as defined in *FRBR* or *FRAD* (e.g., the RDA *title* element corresponds to the FRBR attribute *title of manifestation*). The scope of each RDA element is determined by the scope of the corresponding attribute or relationship, as defined in *FRBR* or *FRAD*.

For any RDA element, one or more element sub-types (i.e., element refinements) may be defined. For example, for the RDA *title* element, sub-types are defined for *title proper*, *parallel title*, *other title information*, *variant title*, *earlier/later title*, and *key title*. Each element sub-type is a sub-class of the element under which it is defined (i.e., the defined scope of the element sub-type is co-extensive with the defined scope of the element). RDA element sub-types are generally defined for purposes of mapping more precisely to element refinements defined in related metadata schemes for encoding or presentation (e.g., MARC 21,[11] ISBD[12]).

For any RDA element or element sub-type, one or more sub-elements (i.e., element components) may be defined. For example, for the RDA *edition* element, sub-elements are defined for *edition statement* and *statement relating to a named revision of an edition*. Each sub-element is a discrete component of the element or element sub-type under which it is defined (i.e., the defined scope of the sub-element covers only a part or component of the defined scope of the element or

---

[10] See the attributes defined in sections 4.1-4.7 and the relationships defined in sections 5.3-5.4 of *FRAD*.

[11] *MARC 21 Format for Bibliographic Data* (Washington: Library of Congress; Ottawa: National Library of Canada, 1999-) and *MARC 21 Format for Authority Data* (Washington: Library of Congress; Ottawa: National Library of Canada, 1999-).

[12] *ISBD(G): General International Standard Bibliographic Description*, 2004 Revision (International Federation of Library Associations and Institutions, 2004).

element sub-type). RDA sub-elements are generally defined for purposes of mapping more precisely to sub-elements defined in related metadata schemes for encoding or presentation.

1.5 Value representations

RDA guidelines and instructions cover the following categories of value representations for the attributes of a resource or other entity:[13]

Label

A string whose function is to distinguish one entity from another (e.g., identifiers, names, titles).

Quantity

A number measuring some aspect of an entity (e.g., extent, dimensions, duration).

Quality

A characteristic of the structure or nature of an entity (e.g., colour, language, gender).

Type

A categorization of one or more characteristics of an entity (e.g., media type, carrier type, content type).

Role

A part played or function fulfilled by an entity in relation to another entity or entities (e.g., the function played by a person, family, or corporate body in relation to the content of a resource).

RDA guidelines and instructions cover the following categories of value representations for a related resource or related entity:

Identifier

A *label* in the form of a number or code uniquely associated with the related resource or entity (e.g., a standard number).

Name

A *label* in the form of a *controlled access point* representing the related resource or entity (e.g., a controlled access point representing a person).

Description

A set of one or more value representations reflecting attributes of the related resource or entity (e.g., the title, publisher, and date of publication of the related resource).

[13] Based on the generic attributes defined in *Indecs* p. 17.

1.6 Value strings

All RDA value representations are formulated as value strings[14] (i.e., alpha-numeric strings representing the value of an attribute or relationship).

An RDA value string may take any of the following forms:

Transcribed string

A *label* transcribed from the resource itself or from another source.

Structured string

A *quantity*, *quality*, *type*, or *role*, or a reference to a related resource or entity, recorded in a specified form.

Unstructured string

A *quantity*, *quality*, *type*, or *role*, or a reference to a related resource or entity, recorded as free text.

Value representations in the form of rich representations[15] (i.e., marked-up text, images, videos, audio, etc.) are out of scope for *RDA*.

1.7 Application

For each element of descriptive data, *RDA* provides general guidelines and instructions that can be applied to any resource exhibiting the characteristic represented in that element. Where necessary, *RDA* specifies exceptions to the general guidelines and instructions that apply to specific types of media, content, mode of issuance, etc. Supplementary guidelines and instructions provide additional detail on formulating descriptive data for specific types of media, etc., and for resources that exhibit characteristics not covered by the general guidelines and instructions.

For each type of entity represented by a controlled access point (i.e., *person*, *family*, *corporate body*, etc.), *RDA* provides general instructions on elements of access point control data that can be applied to any entity of that type that exhibits the characteristic reflected in that element. Where necessary, *RDA* specifies exceptions for specific entity sub-types (e.g., government bodies as a sub-type of corporate body). Supplementary guidelines and instructions provide additional detail on formulating access point control data for specific entity sub-types, and for specific element sub-types (e.g., names of persons in specific languages) not covered by the general guidelines and instructions.

1.8 Record structure

RDA does not specify a record structure for the encoding or presentation of descriptive data or access point control data. Value representations formulated according to the guidelines and instructions in *RDA* are treated as discrete elements that can be stored or presented in a variety of record structures.

---

[14] The term *value string* is used as defined in the *DCMI Abstract Model*, p. 8.
[15] The term *rich representation* is used as defined in the *DCMI Abstract Model*, p. 7.

Mappings of RDA elements to a select number of encoding and presentation structures (e.g., MARC 21, ISBD) are provided in RDA appendices.

2. Structure

RDA is divided into two parts: part A covers descriptive data; part B covers access point control data.

2.1 Part A –Description

The initial chapter in part A provides general guidelines relating to various types of description, changes requiring a new description, required elements, language and script of the description, and conventions used in formulating transcribed, structured, and unstructured strings, etc. The remaining six chapters cover descriptive elements reflecting attributes of *work*, *expression*, *manifestation*, and *item* organized as follows:

Resource identification

The elements covered reflect the attributes of *manifestation* and *item* that are most commonly used to identify a resource. For the most part, the value representations for the elements are labels (e.g., title, statement of responsibility, edition) transcribed from the resource itself. Also included are instructions on recording certain elements as in the form of structured strings (e.g., frequency), and on providing additional details relating to the elements covered in the form of unstructured strings (e.g., note on the source of the title proper).

Carrier

The elements covered reflect attributes of *manifestation* and *item* associated with the carrier of a resource and with the formatting and encoding of the information stored on the carrier. The elements reflect both general and media-specific attributes. For the most part, the value representations for the elements are recorded as structured strings representing quantities (e.g., extent), qualities (e.g., layout, colour, digital characteristics), and types (e.g., media type, carrier type). Also included are instructions on providing additional details relating to characteristics of the carrier, or to the formatting or encoding of information stored on the carrier, in the form of unstructured strings (e.g., notes on equipment and system requirements).

Content

The elements covered reflect attributes of *work* and *expression* associated with the intellectual or artistic content of a resource. The elements reflect attributes that may apply to any type of content as well as those associated with specific types of content. The value representations for the elements are recorded as either structured or unstructured strings representing qualities (e.g., nature and scope of the content, intended audience, language) and types (e.g., content type).

Acquisition and access

The elements covered reflect attributes of *manifestation* and *item* associated with acquiring or obtaining access to a resource (e.g., terms of availability, contact information, restrictions on access). The value representations for the elements are recorded as unstructured strings representing quantities (e.g., price) or qualities (e.g., address of a supplier).

Persons, families, and corporate bodies associated with a resource

The elements covered reflect relationships between the resource described and *persons*, *families*, and *corporate bodies* associated with the resource (e.g., creators, contributors, publishers, custodians). The value representations for the elements are recorded as names (i.e., controlled access points) representing persons, families, and corporate bodies associated with the resource, and as structured strings representing the role played by the person, etc., in relation to the resource. Supplementary instructions on relationships pertaining to specific types of content (e.g., legal works) are also included.

Related resources

The elements covered reflect relationships between the resource described and other resources (i.e., related *works*, *expressions*, *manifestations*, and *items*). The value representations are recorded as identifiers, names (i.e., controlled access points), or descriptions for the related works, expressions, manifestations, and items, and as structured strings representing the nature of the relationship. Supplementary instructions on relationships pertaining to specific types of content (e.g., musical works, art works) are also included.

2.2 Part B – Access point control

The initial chapter in part B provides general guidelines relating to preferred access points and references, required elements, language and script of access points, and conventions used in recording names and titles for use in controlled access points. The remaining six chapters cover access point control elements reflecting attributes associated with the entities *person*, *family*, *corporate body*, *place*, *work*, *expression*, *manifestation*, and *item*, organized as follows:

Access points representing persons

The elements covered reflect attributes used to represent a *person* by means of a controlled access point. The value representations for the elements are recorded as structured strings representing labels (e.g., personal name, title) and qualities (e.g., date of birth). Supplementary instructions on names of persons in specific languages are also included.

Access points representing families

The elements covered reflect attributes used to represent a *family* by means of a controlled access point. The value representations for the elements are recorded as structured strings representing labels (e.g., family name) and qualities (e.g., place associated with the family).

Access points representing corporate bodies

The elements covered reflect attributes used to represent a *corporate body* by means of a controlled access point. The value representations for the elements

are recorded as structured strings representing labels (e.g., corporate name) and qualities (e.g., place associated with the body). Supplementary instructions on names of specific types of corporate bodies (e.g., government bodies, religious bodies) are also included.

Access points representing places

The elements covered reflect attributes used to represent a *place* by means of a controlled access point or as part of a controlled access point representing a corporate body. The value representations for the elements are recorded as structured strings representing labels (e.g., place name) and qualities (e.g., type of jurisdiction).

Access points representing works, expressions, manifestations, and items

The elements covered reflect attributes used to represent a *work*, *expression*, *manifestation*, or *item* by means of a controlled access point. The value representations for the elements are recorded as structured strings representing labels (e.g., title of work) and qualities (e.g., language of expression). Supplementary instructions on names of specific types of works (e.g., laws, sacred scriptures) are also included.

Other information used in access point control

The elements covered reflect additional attributes associated with a *person*, *family*, *corporate body*, *place*, *work*, *expression*, *manifestation*, or *item* as well as selected attributes of *name*, *identifier*, *controlled access point*, and *rules* that are used for purposes of access point control. The value representations for the elements are recorded as structured strings representing labels (e.g., identifier) or as unstructured strings representing qualities (e.g., gender, field of activity).

2.3 Format

RDA guidelines and instructions relating to specific elements are presented in the following format:

ELEMENT NAME

The name or label used to identify the element in RDA.

REQUIREMENT

The element requirement set by RDA (required, required if applicable, optional)

BASIC INSTRUCTIONS

Scope

The scope of the element as defined in RDA.

Sources of information

Sources from which information used as the basis for representing the element may be derived.

Instructions

Basic instructions on formulating value representations for the element.

Guidelines and instructions relating to element sub-types are presented following the basic instructions for the element in the following format:

ELEMENT SUB-TYPE NAME

The name or label used to identify the element sub-type in RDA.

REQUIREMENT

The element sub-type requirement set by RDA (required, required if applicable, optional)

Scope

The scope of the element sub-type as defined in RDA.

Sources of information

Sources from which information used as the basis for representing the element sub-type may be derived.

Instructions

Basic instructions on formulating value representations for the element sub-type.

Guidelines and instructions relating to sub-elements are presented following the basic instructions for the element or element sub-type in the following format:

SUB-ELEMENT NAME

The name or label used to identify the sub-element in RDA.

REQUIREMENT

The sub-element requirement set by RDA (required, required if applicable, optional)

Scope

The scope of the sub-element as defined in RDA.

Sources of information

Sources from which information used as the basis for representing the sub-element may be derived.

Instructions

Basic instructions on formulating value representations for the sub-element.

In cases where *RDA* includes instructions on providing additional detail on the attribute or relationship represented in an element in the form of notes (i.e., as unstructured strings), instructions relating to notes are presented following the instructions on recording the element in the form of a transcribed or structured string as follows:

NOTES ON [ELEMENT NAME]

REQUIREMENT

The requirement set by RDA (required, required if applicable, optional)

Scope

The scope of the notes as defined in RDA.

Sources of information

Sources from which information used as the basis for the notes may be derived.

Instructions

Basic instructions on formulating the notes.

2.4 Specificity of instructions

Instructions on recording an element, element sub-type, or sub-element are presented in order of increased specificity. Basic instructions address aspects of the attribute or relationship reflected in the element, element sub-type, or sub-element that are most commonly encountered when formulating descriptive data or access point control data. Detailed instructions addressing less frequently encountered aspects of the attribute or relationship are presented under specific headings following the basic instructions, as required.

2.5 Appendices

The appendices to *RDA* provide information on the following:

Capitalization

Guidelines on capitalization conventions used in English and a selected number of other languages.

Abbreviations

Lists of abbreviations used in English and a selected number of other languages.

Initial articles

Lists of initial articles used in English and a selected number of other languages.

Record structures for descriptive data

Mappings of RDA descriptive elements to a selected number of related metadata schemes for encoding and presentation of descriptive data (e.g., MARC 21, ISBD).

Record structures for access point control data

Mappings of RDA access point control elements to a selected number of related metadata schemes for encoding and presentation of access point control data (e.g., MARC 21, GARR).

5JSC/Editor/2
14 January 2007

To: Joint Steering Committee for Revision of AACR

From: Tom Delsey, RDA Editor

Subject: RDA Database Implementation Scenarios

Attached are updated versions of the RDA database implementation scenarios discussed at the October 2006 meeting. The scenarios depicted are intended simply to illustrate some of the potential implementations of RDA data in various database structures.

No constituency responses to this document are required.

RDA Implementation Scenarios

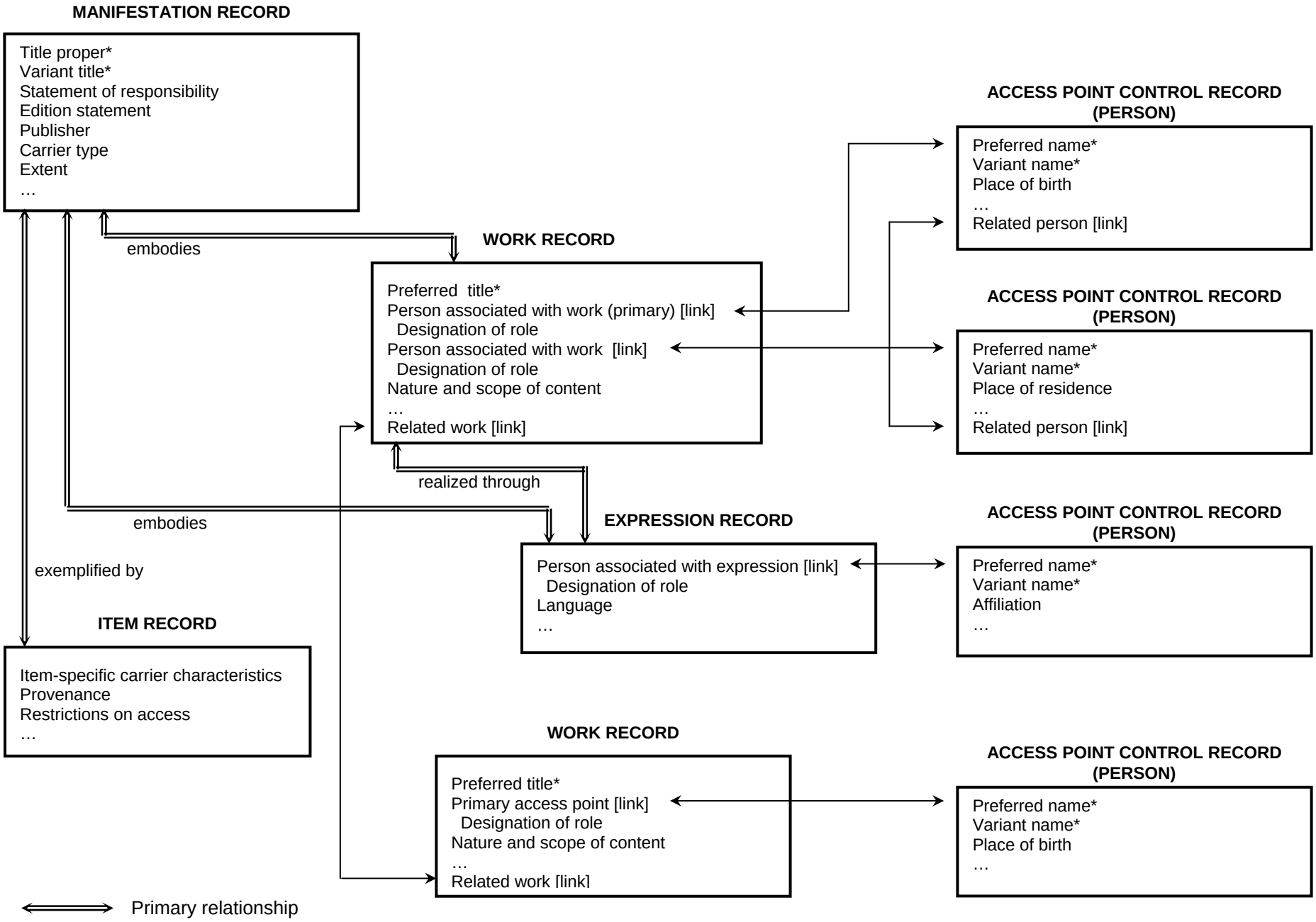
The attached diagrams illustrate three potential implementation scenarios for RDA data.

In the first scenario, RDA data are stored in a relational or object-oriented database structure that mirrors the FRBR and FRAD conceptual models. Descriptive data elements are stored in records that parallel the primary entities in the FRBR model: work records, expression records, manifestation records, and item records. Data elements used for access point control are stored in records that are centred on the primary entities in the FRAD model: persons, families, corporate bodies, etc. Data elements indexed as access points (both controlled and uncontrolled) are marked with an asterisk. Relationships between the primary FRBR entities are reflected through links from one record to another. For example, the link from the manifestation record to the work record reflects the primary relationship between the manifestation and the work that it embodies. Similarly, a relationship between one work and another (e.g., a derivative relationship) is reflected in a link from one work record to another. Relationships between the primary FRBR entities and a person, family, corporate body, etc., are reflected through links from work records, etc., to access point control records for persons, etc. The relationship between one person and another, etc., is reflected in a link from one access point control record to another.

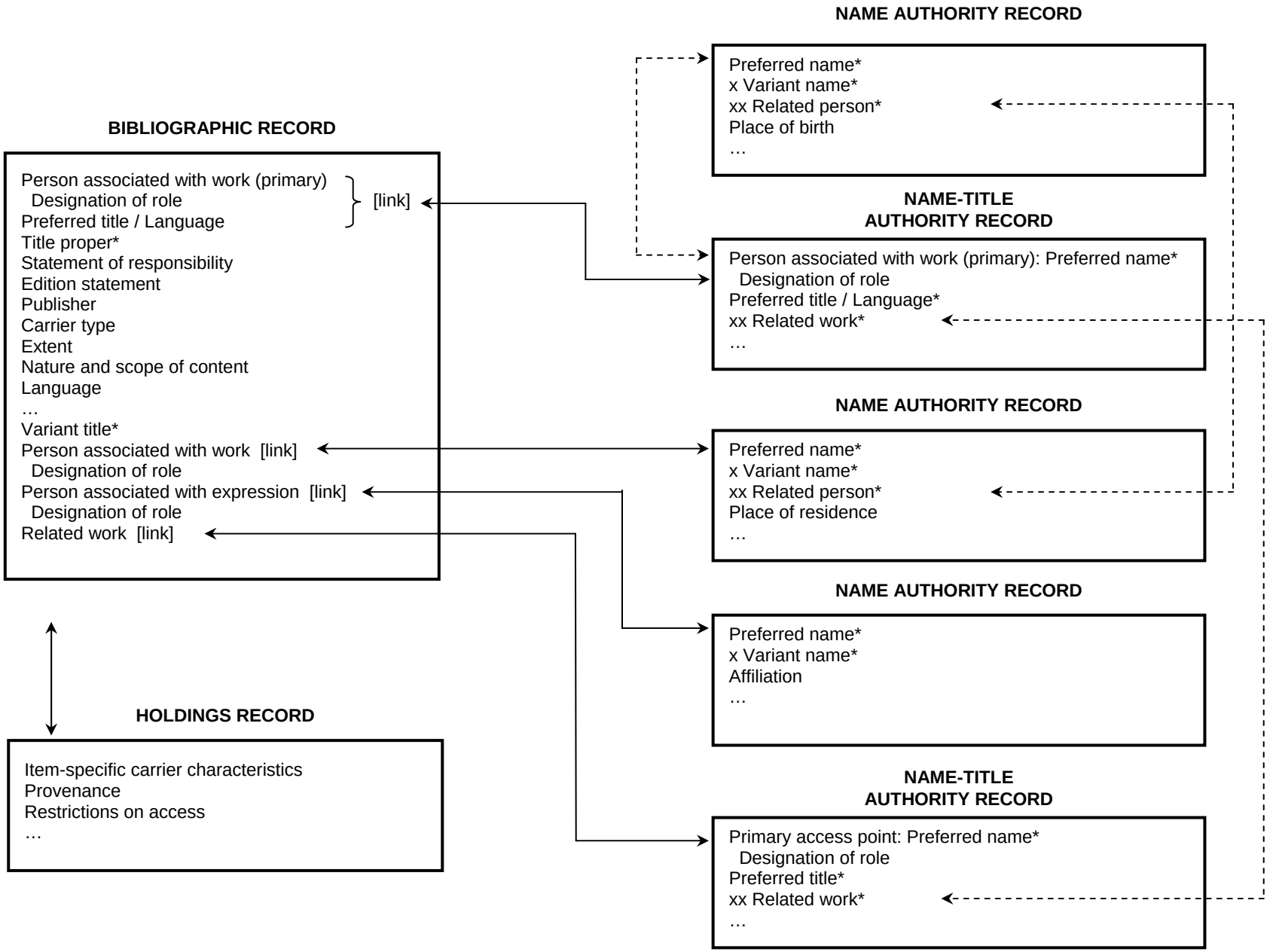
In the second and third scenarios, RDA data is stored in database structures conventionally used in library applications. In those structures, data is stored in bibliographic records and in authority records, and in some implementations in holdings records as well (as shown in scenario 2). Descriptive data elements are stored in bibliographic records. In implementations where bibliographic files and authority files are linked (scenario 2), the bibliographic record also contains links to authority records for persons, families, corporate bodies, etc., associated with the work, etc., embodied in the resource described. In implementations where bibliographic files and authority files are not linked (scenario 3), access points using the preferred name or title for the person, etc., are stored in the bibliographic record along with the descriptive data. In both types of implementation, variant names and other data used for access point control are stored in authority records.

RDA data can be readily mapped to any one of the implementation scenarios (or to variations on the three scenarios illustrated). In all implementations the data will support the functional objectives that RDA is designed to fulfil. The data structures used to store the data and to reflect relationships, however, will have a bearing both on the efficiency of data creation and maintenance, and on the ease and effectiveness with which users are able to access the data and navigate the database. For example, the use of records for works and expressions in the relational and object-oriented database structures ensures access not only to all works and expressions associated with a particular person, etc., but to all related works (adaptations, etc.) as well, regardless of whether the name of that person is used as the primary access point for those works or not.

Scenario 1: Relational / object-oriented database structure



Scenario 2: Linked bibliographic and authority records



Scenario 3: 'Flat file' database structure (no links)

BIBLIOGRAPHIC RECORD

Person associated with work (primary): Preferred name*
  Designation of role
Preferred title / Language*
Title proper*
Statement of responsibility
Edition statement
Publisher
Carrier type
Extent
Nature and scope of content
Language
…
Item-specific carrier characteristics
Provenance
Restrictions on access
…
Variant title*
Person associated with work: Preferred name*
  Designation of role
Person associated with expression: Preferred name*
  Designation of role
Related work: Primary access point/ Preferred title*

NAME AUTHORITY RECORD

Preferred name*
x Variant name*
xx Related person*
Place of birth
…

NAME-TITLE AUTHORITY RECORD

Person associated with work (primary): Preferred name*
  Designation of role
Preferred title / Language*
xx Related work*
…

NAME AUTHORITY RECORD

Preferred name*
x Variant name*
xx Related person*
Place of residence
…

NAME AUTHORITY RECORD

Preferred name*
x Variant name*
Affiliation
…

NAME-TITLE AUTHORITY RECORD

Primary access point: Preferred name*
  Designation of role
Preferred title*
xx Related work*
…

Encoding RDA data

RDA as a metadata element set

RDA can be viewed as a metadata element set (similar to the Dublin Core Metadata Element Set) insofar as it:

a) specifies a set of elements, element sub-types, and sub-elements that reflect the properties of a resource
b) defines each element, element sub-type, and sub-element
c) establishes parameters for the value representations recorded for each element, element sub-type, and sub-element.

As a metadata element set, RDA incorporates the following additional features (similar to those in the Dublin Core Library Application Profile):

a) it establishes requirements (required, required if applicable, and optional) for each element, element sub-type, and sub-element
b) it incorporates (either directly or by reference) encoding schemes for certain elements, element sub-types, and sub-elements (e.g., the ISO encoding schemes for standard identifiers such as ISBNs and ISSNs).

RDA has not, at this stage, been formally registered as a metadata element set, nor have Uniform Resource Identifiers (URIs) been assigned to RDA data elements, element sub-types, and sub-elements.

Encoding syntax

In the absence of a formally registered RDA metadata element set, data created using RDA must be encoded using a "proxy" syntax. That is to say that RDA elements, element sub-types, and sub-elements must be encoded using a syntax that can be mapped to RDA specifications. The candidate proxies include MARC 21, MODS, and XML implementations of Dublin Core.

Using a proxy encoding syntax means that the data is encoded using the element set defined for that syntax rather than the RDA element set *per se*. RDA elements, element sub-types, and sub-elements are mapped to the nearest corresponding element, element sub-type, or sub-element defined in the proxy syntax, and encoded accordingly.

Once the initial phase of RDA development has been completed, it should be possible to formally register an RDA metadata element set. At that stage, RDA data could be encoded directly in XML using the registered RDA terms to identify the elements, element sub-types, and sub-elements. The mappings of RDA to MARC 21, MODS, and Dublin Core could then be used as the basis for conversion to and from the RDA encoding syntax when importing or exporting data in MARC 21, MODS, or Dublin Core in XML.

Encoding schemes for RDA value representations

As noted above, for certain elements, element sub-types, and sub-elements defined in the RDA element set, the RDA instructions reference "external" encoding schemes (e.g., the ISO encoding schemes for standard identifiers such as ISBNs and ISSNs). For certain other elements, element sub-types, and sub-elements, RDA provides instructions on recording value representations in a structured form that function, in effect, as "internal" encoding schemes. For example, the controlled lists of values for elements such as Media type, Carrier type, and Content type function as vocabulary encoding schemes for those elements. At this stage, however, none of the controlled lists of values specified in RDA have been formally registered as vocabulary encoding schemes.

When RDA data is encoded in a proxy encoding syntax, the values for which RDA provides an "internal" vocabulary encoding scheme must be recorded in the corresponding element, element sub-type, or sub-element defined in the proxy encoding syntax. In some instances, the corresponding element, element sub-type, or sub-element in the proxy encoding syntax will have defined its own "internal" vocabulary encoding scheme (e.g., the coded data in the 007 and 008 fields in MARC 21). In those cases, the vocabulary specified for the RDA element, element sub-type, or sub-element must be mapped to the vocabulary specified in the proxy encoding syntax. In other instances, there will be no vocabulary encoding scheme defined in the proxy encoding syntax (e.g., the variable data for "other technical details" recorded in field 300, subfield b in MARC 21), and the RDA value representation must be recorded simply as an unstructured value string.

Once the initial phase of RDA development has been completed, it should be possible to formally register the controlled lists of terms specified for various RDA elements, element sub-types, and sub-elements as vocabulary encoding schemes, with assigned URIs. At that stage, the URI could be used to identify such a term recorded in the corresponding element, element sub-type, or sub-element in a proxy encoding syntax as a term derived from a specific RDA vocabulary encoding scheme (assuming the proxy syntax allows the use of "external" vocabulary encoding schemes and makes provision for identifying the scheme from which the recorded term was derived).

Alternatives for recording RDA values

The Joint Steering Committee (JSC) is currently assessing the implications of incorporating into RDA alternatives for recording values for three specific types of data: (1) elements, element sub-types, and sub-elements for which RDA specifies a controlled list of values; (2) designations used to indicate roles and relationships; and (3) elements, element sub-types, and sub-elements for which there is an international standard encoding scheme.

Controlled lists of values

The current draft of RDA specifies controlled lists of values for a number of elements reflecting characteristics of both the carrier and the content of the resource being described. The values specified in those controlled lists are all represented as terms (i.e., text strings). Consideration is being given to incorporating a general guideline into RDA that would allow the recording of a coded value as a substitute for a term specified in a controlled list.

It is assumed that when data formulated according to RDA specifications is recorded using a proxy encoding syntax such as MARC 21, a term derived from a controlled list for a specific RDA element, element sub-type, or sub-element would normally be recorded as an unstructured value string in the corresponding variable field of the encoding syntax.

If the alternative under consideration is introduced, a coded value defined in the proxy encoding syntax (e.g., in field 007 or 008 in MARC 21) that corresponds to the term specified in RDA could be recorded as a substitute for recording the RDA term in the corresponding variable field. If there is a coded value defined in the proxy encoding syntax that is the direct equivalent of the RDA term, that coded value would be recorded. If, however, there is no coded value defined in the proxy encoding syntax that is the direct equivalent of the RDA term, a coded value for "other" would have to be recorded in the coded data field. In that case, the term specified in RDA (or a more appropriate or specific term, as provided for with all "open" lists in RDA) could be recorded in the corresponding variable field in the proxy encoding syntax.

The key factors relating to the introduction of such an alternative that JSC needs to consider are the following:

a) the degree of correspondence (and/or conflict) between the RDA values and the coded values defined in the encoding syntaxes that are most likely to be used as proxies for encoding RDA data
b) the potential advantages to be gained by recording values in a coded form that could be used to facilitate automated processing of the data (e.g., in record matching, as search limiters, or to generate the display of an equivalent term in a form suited to the community that the database is designed to serve)
c) the implications for user displays of recording certain of the specified values for an element, element sub-type, or sub-element as coded values (e.g., in field 007 or 008 in MARC 21) while coding other specified values (i.e., those that could only be coded as "other" in a coded data field in the proxy encoding syntax) as text strings (e.g., in the corresponding variable field in MARC 21)

The other alternative that could be considered is simply to acknowledge the possibility of recording an RDA value in an equivalent coded form in a proxy encoding syntax not as a substitute for, but in addition to, recording the specified RDA value as a term (i.e., a text string) in the corresponding variable field. That alternative would not necessarily require "redundant" data entry if the data entry software provided a menu of specified terms to be recorded, and the selection of a term triggered both the recording of the term in the corresponding variable field and the generation of the equivalent coded value for the corresponding coded data element.

Ultimately, JSC's assessment of the alternatives will likely hinge on what can realistically be expected by way of support for both data entry and data display from the developers of the software that will manage the data created using RDA.

Designations of roles and relationships

The current draft of RDA provides for the optional use of a designation of role in conjunction with an access point representing a person, family, or corporate body

associated with the resource being described, to indicate the role played by that person, family, or corporate body in relation to the resource. Similar provisions are made for the use of a designation to indicate the nature of the relationship between the resource being described and a related resource. JSC is currently assessing the implications of allowing the use of either a coded value or a term for such designations.

For designations of role, the current draft does not specify the designations, but recommends the use of "standard" lists. The draft does, however, define element sub-types for associated persons, families, and corporate bodies that reflect the role played by that person, family, or corporate body in relation to the resource. Similarly, for designations of relationship, the current draft does not specify the designations (nor, in this case, does it explicitly recommend the use of "standard lists"). It does, however, define element sub-types that reflect the nature of the relationship between the resource being described and the related resource.

To the extent that the element sub-types defined in RDA for persons, families, and corporate bodies associated with the resource and for related resources can be mapped to an equivalent element sub-type in the proxy encoding syntax, the definition of the element sub-type itself will serve as a means of designating the role or relationship. However, where a more specific designation of role or relationship is needed, a term or code derived from a "standard" list would have to be used in conjunction with the appropriate element sub-type.

The key factors relating to allowing the use of designations of role and relationship derived from "standard lists" that JSC needs to consider are the following:

a) the degree of correspondence (and/or conflict) between the RDA element sub-types that reflect roles and relationships and the values defined in the "standard lists" defined for use with the proxy encoding syntaxes that are most likely to be used for encoding RDA data
b) the implications for user displays of recording designations derived from a variety of "standard" lists that are independent of RDA specifications.

The alternative that has been suggested for consideration is to specify RDA values for designations of role and relationship (i.e., at a more specific level than that identified by the element sub-type).

International standard encoding schemes

As previously noted, there is at least one instance in the current draft of RDA where the instructions specify (by reference) the use of an international standard encoding scheme when recording the value for an element (i.e., resource identifiers for which there is an international standard). However, when considering the suggested use of the same approach for the Duration element, JSC chose not to make specific reference to the relevant ISO encoding scheme in the instructions, but rather to word the instruction in neutral terms and to add an example illustrating duration formatted according to the ISO standard.

The key factors relating to referencing international standard encoding schemes (either as the prescribed form for recording an RDA element, element sub-type, or sub-element, or as an alternative form) that JSC needs to consider are the following:

a) the potential advantages to be gained by recording values in accordance with a standard encoding scheme that could be used to facilitate automated processing of the data (e.g., for retrieval or matching)
b) the implications for user displays of recording data in a form that is designed primarily for automated processing and may not be in a form that is readily recognized or understood by a user.

Again, JSC's assessment of the alternatives will likely hinge on what can realistically be expected by way of support for data display from the developers of the software that will manage the data created using RDA.

RDA Scope Analysis

The attached table provides an analysis of each element, element sub-type, and sub-element specified in Part A of RDA *vis à vis* the scope of RDA as stated in the *RDA Scope and Structure* document.

The analysis is designed to assist JSC in assessing comments on the scope of RDA as articulated in the draft *RDA Scope and Structure* document, and in making decisions on any reworking of either the scope statement or the specifications for RDA elements that may be considered necessary to better align the two.

The table layout is as follows:

- The first column in the table lists the RDA "elements" (i.e., elements, element sub-types, and sub-elements) in the order in which they appear in the most recent drafts of RDA.
- The second column identifies the classification of the "element" (i.e., as an element, an element sub-type, or a sub-element).
- The third column identifies the corresponding FRBR attribute or relationship.
- The fourth column identifies the generic attribute type reflected by the element as defined in the <in*d*ecs> Metadata Framework (i.e., label, quantity, quality, type, or role).
- The fifth column identifies the category of value string used to record the element (i.e., transcribed, structured, or unstructured).
- The sixth column identifies the syntax encoding scheme specified for the element (e.g., the ISO encoding scheme for the ISSN of a series).
- The seventh column identifies the vocabulary encoding scheme specified for the element. (Vocabulary encoding schemes identified as "RDA …" refer to the controlled lists of terms given in the RDA instructions on recording the element.)

Elements for which there are new proposals currently under review by JSC, or for which structural changes have been proposed since the last draft of the relevant chapter of RDA was issued are shaded. (The structures reflected for those elements are only tentative; they reflect what the Editor judges to be most consistent with the overall structure of RDA, but are subject to review by JSC.)

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
RESOURCE IDENTIFICATION						
Title	element	title of the manifestation	label	transcribed / unstructured	n/a	n/a
Title proper	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Alternative title	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Parallel alternative title	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Parallel title	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Other title information	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Parallel other title information	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Variant title	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Earlier/later title	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Key title	element sub-type	title of the manifestation	label	transcribed	n/a	n/a
Abbreviated title	element sub-type	title of the manifestation	label	unstructured	n/a	n/a
Devised title	element sub-type	title of the manifestation	label	unstructured	n/a	n/a
Notes on titles	?	?	?	unstructured	n/a	n/a
Statement of responsibility	element	statement of responsibility	label	transcribed / unstructured	n/a	n/a
Statement of responsibility relating to title	element sub-type	statement of responsibility	label	transcribed	n/a	n/a
Parallel statement of responsibility relating to title	element sub-type	statement of responsibility	label	transcribed	n/a	n/a
Statement of responsibility relating to the edition	element sub-type	statement of responsibility	label	transcribed	n/a	n/a
Parallel statement of responsibility relating to the edition	element sub-type	statement of responsibility	label	transcribed	n/a	n/a
Statement of responsibility relating to a named revision of an edition	element sub-type	statement of responsibility	label	transcribed	n/a	n/a
Parallel statement of responsibility relating to a named revision of an edition	element sub-type	statement of responsibility	label	transcribed	n/a	n/a
Statement of responsibility relating to series	element sub-type	statement of responsibility	label	transcribed	n/a	n/a
Parallel statement of responsibility relating to series	element sub-type	statement of responsibility	label	transcribed	n/a	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Notes on statements of responsibility	?	?	?	unstructured	n/a	n/a
Edition	element	edition/issue designation	label	transcribed / unstructured	n/a	n/a
Edition statement	element sub-type	edition/issue designation	label	transcribed	n/a	n/a
Parallel edition statement	element sub-type	edition/issue designation	label	transcribed	n/a	n/a
Statement relating to a named revision of an edition	element sub-type	edition/issue designation	label	transcribed	n/a	n/a
Parallel statement relating to a named revision of an edition	element sub-type	edition/issue designation	label	transcribed	n/a	n/a
Notes on edition	?	?	?	unstructured	n/a	n/a
Numbering of serials	element	numbering (serial)	label	unstructured	n/a	n/a
Numeric and/or alphbetic designation of first issue or part	element sub-type	numbering (serial)	label	unstructured	n/a	n/a
Chronological designation of first issue or part	element sub-type	numbering (serial)	label	unstructured	n/a	n/a
Numeric and/or alphbetic designation of last issue or part	element sub-type	numbering (serial)	label	unstructured	n/a	n/a
Chronological designation of last issue or part	element sub-type	numbering (serial)	label	unstructured	n/a	n/a
Notes on numbering	?	?	?	unstructured	n/a	n/a
Finding	element	n/a	role?	unstructured	n/a	n/a
Finder	sub-element	n/a	role?	unstructured	n/a	n/a
Place of finding	sub-element	n/a	role?	unstructured	n/a	n/a
Date of finding	sub-element	n/a	role?	unstructured	n/a	n/a
Notes on finding	?	?	?	unstructured	n/a	n/a
Production	element	~	label	transcribed / unstructured	n/a	n/a
Producer	sub-element	fabricator/manufacturer	label	transcribed	n/a	n/a
Place of production	sub-element	n/a	label	transcribed	n/a	n/a
Date of production	sub-element	n/a	label?	unstructured	n/a	n/a
Notes on production	?	?	?	unstructured	n/a	n/a
Publication	element	~	label	transcribed / unstructured	n/a	n/a
Publisher	sub-element	publisher/distributor	label	transcribed	n/a	n/a
Place of publication	sub-element	place of publication/distribution	label	transcribed	n/a	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Date of publication	sub-element	date of publication/distribution	label?	unstructured	n/a	n/a
Notes on publication	?	?	?	unstructured	n/a	n/a
Distribution	element	~	label	transcribed / unstructured	n/a	n/a
Distributor	sub-element	publisher/distributor	label	transcribed	n/a	n/a
Place of distribution	sub-element	place of publication/distribution	label	transcribed	n/a	n/a
Date of distribution	sub-element	date of publication/distribution	label?	unstructured	n/a	n/a
Notes on distribution	?	?	?	unstructured	n/a	n/a
Series	element	series statement	label	transcribed / structured	n/a	n/a
Title proper of series	sub-element	series statement	label	transcribed	n/a	n/a
Parallel title of series	sub-element	series statement	label	transcribed	n/a	n/a
Other title information of series	sub-element	series statement	label	transcribed	n/a	n/a
Parallel other title information of series	sub-element	series statement	label	transcribed	n/a	n/a
ISSN of series	sub-element	series statement	label	structured	ISO specs	n/a
Title proper of subseries	sub-element	series statement	label	transcribed	n/a	n/a
Parallel title of subseries	sub-element	series statement	label	transcribed	n/a	n/a
Other title information of subseries	sub-element	series statement	label	transcribed	n/a	n/a
Parallel other title information of subseries	sub-element	series statement	label	transcribed	n/a	n/a
ISSN of subseries	sub-element	series statement	label	structured	ISO specs	n/a
Frequency	element	expected frequency of issue (serial)	quality?	unstructured	n/a	n/a
Resource identifier	element	manifestation identifier / item identifier / fingerprint?	label	structured / unstructured	as specified for identifier (if applicable)	n/a
Preferred citation	element	n/a	label	unstructured	n/a	n/a
Issue, part, or iteration used as the basis for the description	element	n/a	?	unstructured	n/a	n/a
Creation, custodial history and immediate source of acquisition	element	provenance of the item	role?	unstructured	n/a	n/a
Creation of an archival resource or collection	element sub-type	provenance of the item	role?	unstructured	n/a	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Custodial history	element sub-type	provenance of the item	role?	unstructured	n/a	n/a
Immediate source of acquisition	element sub-type	provenance of the item	role?	unstructured	n/a	n/a
CARRIER						
Media type	element	n/a	type	unstructured	n/a	RDA media type
Carrier type	element	form of carrier	type	unstructured	n/a	RDA carrier type
Extent	element	extent of the carrier / extent of the expression	quantity	structured / unstructured	RDA extent	RDA carrier type / RDA subunit
Extent of cartographic resources	element sub-type	extent of the expression / extent of the carrier	quantity	structured / unstructured	RDA extent of cartographic resources	RDA extent of cartographic resources
Extent of notated music	element sub-type	extent of the expression / extent of the carrier	quantity	structured / unstructured	RDA extent of notated music	RDA extent of notated music
Extent of still images	element sub-type	extent of the expression / extent of the carrier	quantity	structured / unstructured	RDA extent of still images	RDA extent of still images
Extent of text	element sub-type	extent of the expression / extent of the carrier	quantity	structured / unstructured	RDA extent of text	RDA extent of text
Extent of three-dimensional forms	element sub-type	extent of the expression / extent of the carrier	quantity	structured	RDA extent	RDA extent of three-dimensional forms
Notes on extent	element sub-type	extent of the expression / extent of the carrier	quantity / quality	unstructured	n/a	n/a
Dimensions	element	dimensions of the carrier	quantity	structured	RDA dimensions	n/a
Dimensions of maps, etc.	element sub-type	n/a	quantity	structured	RDA dimensions of maps, etc.	n/a
Dimensions of still images	element sub-type	n/a	quantity	structured	RDA dimensions of still images	n/a
Notes on dimensions	element sub-type	dimensions of the carrier +	quantity / quality	unstructured	n/a	n/a
Base material	element	physical medium	quality	unstructured	n/a	RDA base material

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Base material for microfilm, microfiche, and motion picture film	element sub-type	physical medium	quality	unstructured	n/a	RDA base material for microfilm, microfiche, and motion picture film
Notes on base material	element sub-type	physical medium	quality	unstructured	n/a	n/a
Applied material	element	physical medium	quality	unstructured	n/a	RDA applied material
Emulsion on microfilm and microfiche	element sub-type	physical medium	quality	unstructured	n/a	RDA emulsion on microfilm and microfiche
Notes on applied material	element sub-type	physical medium	quality	unstructured	n/a	n/a
Mount	element	n/a	quality	unstructured	n/a	RDA base material
Notes on mount	element sub-type	n/a	quality	unstructured	n/a	n/a
Production method	element	technique (graphic or projected image) / recording technique (remote sensing image) / +	quality	unstructured	n/a	RDA production method
Production method for manuscripts	element sub-type	n/a	quality	unstructured	n/a	RDA production method for manuscripts
Production method for tactile resources	element sub-type	n/a	quality	unstructured	n/a	RDA production method for tactile resources
Notes on production method	element sub-type	n/a	quality	unstructured	n/a	n/a
Generation	element	~	quality	unstructured	n/a	see below
Audio recordings	element sub-type	n/a	quality	unstructured	n/a	RDA generation (audio recordings)
Digital resources	element sub-type	n/a	quality	unstructured	n/a	RDA generation (digital resources)
Microforms	element sub-type	generation (microform or visual projection)	quality	unstructured	n/a	RDA generation (microforms)

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Motion picture films	element sub-type	generation (microform or visual projection)	quality	unstructured	n/a	RDA generation (motion picture films)
Videotapes	element sub-type	n/a	quality	unstructured	n/a	RDA generation (videotapes)
Notes on generation	element sub-type	generation (microform or visual projection) +	quality	unstructured	n/a	n/a
Layout	element	n/a	quality	unstructured	n/a	RDA layout
Layout of maps, etc.	element sub-type	n/a	quality	unstructured	n/a	RDA layout of maps, etc.
Layout of tactile music	element sub-type	n/a	quality	unstructured	n/a	RDA layout of tactile music
Layout of tactile text	element sub-type	n/a	quality	unstructured	n/a	RDA layout of tactile text
Notes on layout	element sub-type	n/a	quality	unstructured	n/a	n/a
Colour	element	~	quality	unstructured	n/a	RDA colour
Still images	element sub-type	colour (image)	quality	unstructured	n/a	RDA colour +
Moving images	element sub-type	colour (image)	quality	unstructured	n/a	RDA colour +
Three-dimensional forms	element sub-type	n/a	quality	unstructured	n/a	RDA colour +
Resources designed for persons with visual impairments	element sub-type	n/a	quality	unstructured	n/a	n/a
Notes on colour	element sub-type	n/a	quality	unstructured	n/a	n/a
Foliation	element	foliation (hand-printed books)	quality	unstructured	n/a	RDA foliation
Notes on foliation	element sub-type	foliation (hand-printed books)	quality	unstructured	n/a	n/a
Font size	element	type size (printed book)	quality	unstructured	n/a	RDA font size +
Notes on font size	element sub-type	type size (printed book)	quality	unstructured	n/a	n/a
Polarity	element	polarity (microform or visual projection)	quality	unstructured	n/a	RDA polarity
Notes on polarity	element sub-type	polarity (microform or visual projection)	quality	unstructured	n/a	n/a
Reduction ratio	element	reduction ratio (microform)	quality	unstructured	n/a	RDA reduction ratio +
Notes on reduction ratio	element sub-type	reduction ratio (microform)	quality	unstructured	n/a	n/a
Sound characteristics	element	~	quality	unstructured	n/a	RDA sound characteristics

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Type of recording	element sub-type	capture mode	quality	unstructured	n/a	RDA type of recording
Playing speed	element sub-type	playing speed (sound recording)	quality	unstructured	n/a	RDA playing speed
Groove characteristic	element sub-type	kind of cutting (sound recording)	quality	unstructured	n/a	RDA groove characteristic
Track configuration	element sub-type	n/a	quality	unstructured	n/a	RDA track configuration
Tape configuration	element sub-type	tape configuration (sound recording)	quality	unstructured	n/a	RDA tape configuration
Configuration of playback channels	element sub-type	kind of sound (sound recording)	quality	unstructured	n/a	RDA configuration of playback channels
Special playback characteristics	element sub-type	special reproduction characteristics (sound recording)	quality	unstructured	n/a	RDA special playback characteristics
Notes on sound characteristics	element sub-type	n/a	quality	unstructured	n/a	n/a
Projection characteristics	element	~	quality	unstructured	n/a	see below
Presentation format	element sub-type	presentation format (visual projection)	quality	unstructured	n/a	RDA presentation format
Projection speed	element sub-type	n/a	quality	unstructured	n/a	RDA projection speed
Notes on projection characteristics	element sub-type	n/a	quality	unstructured	n/a	n/a
Video characteristics	element	~	quality	unstructured	n/a	see below
Video format	element sub-type	presentation format (visual projection) [sic]	quality	unstructured	n/a	RDA video format
Broadcast standard	element sub-type	n/a	quality	unstructured	n/a	RDA broadcast standard
Notes on video characteristics	element sub-type	n/a	quality	unstructured	n/a	n/a
Digital file characteristics	element	~	quality	unstructured	n/a	see below
File type	element sub-type	file characteristics (electronic resource)	quality	unstructured	n/a	RDA file type
Encoding format	element sub-type	file characteristics (electronic resource)	quality	unstructured	n/a	RDA encoding format

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
File size	element sub-type	file characteristics (electronic resource)	quality	unstructured	n/a	RDA file size
Transmission speed	element sub-type	n/a	quality	unstructured	n/a	RDA transmission speed
Representation of cartographic images	element sub-type	n/a	quality	unstructured	n/a	RDA representation of cartographic images
Notes on digital file characteristics	element sub-type	n/a	quality	unstructured	n/a	n/a
Changes in carrier characteristics	element	n/a	quality	unstructured	n/a	n/a
Equipment and system requirements	element	system requirements (electronic resource)	role	unstructured	n/a	n/a
Item-specific carrier characteristics	element	condition of the item	quality	unstructured	n/a	n/a
Item-specific carrier characteristics of early printed resources	element sub-type	condition of the item	quality	unstructured	n/a	n/a
CONTENT						
Content type	element	form of expression	type	unstructured	n/a	RDA content type
Nature and scope of the content	element	~	quality	unstructured	n/a	n/a
Intended audience	element	intended audience	quality	unstructured	n/a	n/a
Summarization of the content	element	summarization of content	quality	unstructured	n/a	n/a
System of arrangement	element	n/a	quality	unstructured	n/a	n/a
Dissertations	element	n/a	role?	unstructured	n/a	n/a
Degree	sub-element	n/a	role?	unstructured	n/a	n/a
Institution	sub-element	n/a	role?	unstructured	n/a	n/a
Year degree granted	sub-element	n/a	role?	unstructured	n/a	n/a
Capture	element	n/a	role?	unstructured	n/a	n/a
Capturer	sub-element	n/a	role?	unstructured	n/a	n/a
Place of capture	sub-element	n/a	role?	unstructured	n/a	n/a
Date of capture	sub-element	n/a	role?	unstructured	n/a	n/a
Notes on capture	?	?	?	unstructured	n/a	n/a
Language of the content	element	language of expression	quality	unstructured	n/a	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Notational system	element	n/a	quality	unstructured	n/a	n/a
Script	element sub-type	n/a	quality	unstructured	n/a	n/a
Music notation system	element sub-type	n/a	quality	unstructured	n/a	n/a
System of tactile notation	element sub-type	n/a	quality	unstructured	n/a	RDA tactile system of notation
Illustrative content	element	n/a	quality	unstructured	n/a	RDA illustrative content
Notes on illustrative content	element sub-type	n/a	quality	unstructured	n/a	n/a
Format of notated music	element	type of score	label	transcribed	n/a	n/a
Medium of performance of musical content	element	medium of performance (musical work)	quality	unstructured	n/a	n/a
Duration	element	extent of the expression	quantity	unstructured	n/a	n/a
Scale	element	~	quantity	unstructured	n/a	n/a
Scale of still image or three-dimensional form	element sub-type	n/a	quantity	structured	RDA scale of still image or three-dimensional form	n/a
Scale of cartographic content	element sub-type	scale (cartographic image/object)	quantity	structured	RDA scale of cartographic content	n/a
Additional scale information [cartographic content]	element sub-type	scale (cartographic image/object)	quantity / quality	unstructured	n/a	n/a
Variations in scale [cartographic content]	element sub-type	scale (cartographic image/object)	quantity / quality	structured / unstructured	RDA variations in scale [cartographic content]	RDA variations in scale [cartographic content]
Content not drawn to scale [cartographic content]	element sub-type	scale (cartographic image/object)	quality	unstructured	n/a	RDA cartographic content not drawn to scale
Non-linear scale [cartographic content]	element sub-type	scale (cartographic image/object)	quantity / quality	unstructured	n/a	RDA non-linear scale [cartographic content]

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Vertical scale [cartographic content]	element sub-type	scale (cartographic image/object)	quantity	structured	RDA vertical scale [cartographic content]	n/a
Projection of cartographic content	element	projection (cartographic image/object)	quality	unstructured	n/a	n/a
Coordinates of cartographic content	element	coordinates (cartographic work)	quantity	structured	see below	n/a
Longitude and latitude	element sub-type	coordinates (cartographic work)	quantity	structured	RDA longitude and latitude	n/a
Strings of coordinate pairs	element sub-type	coordinates (cartographic work)	quantity	structured	RDA strings of coordinate pairs	n/a
Ascension and declination	element sub-type	coordinates (cartographic work) / equinox (cartographic work)	quantity	structured	RDA ascension and declination	n/a
Magnitude of cartographic content	element	n/a	quantity	unstructured	n/a	n/a
Other details of cartographic content	element	~	quantity / quality	unstructured	n/a	n/a
Other mathematical data	element sub-type	scale / projection / coordinates / geodetic, grid, and vertical measurements (cartographic image/object)	quantity	unstructured	n/a	n/a
Other features of cartographic content	element sub-type	special characteristic (remote sensing image) / +	quality	unstructured	n/a	n/a
Awards	element	critical response to the expression	quality?	unstructured	n/a	n/a
ACQUISITION AND ACCESS						
Terms of availability	element	terms of availability	quantity / quality	unstructured	n/a	n/a
Contact information	element	source for acquisition/access authorization	role?	unstructured	n/a	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Restrictions on access	element	access restrictions on the manifestation / access restrictions on the item	quality	unstructured	n/a	n/a
Restrictions on use	element	access restrictions on the manifestation / access restrictions on the item	quality	unstructured	n/a	n/a
Appraisal and accrual	element	n/a	quality	unstructured	n/a	n/a
PERSONS, FAMILIES, AND CORPORATE BODIES ASSOCIATED WITH A RESOURCE						
Designation of role	element	n/a?	role	unstructured	n/a	n/a
Access points for creators and contributors of content						
Creator	element	is created by	role	structured	RDA controlled access point [person / family / corporate body]	n/a
Contributor	element	is realized by	role	structured	RDA controlled access point [person / family / corporate body]	n/a
Access points for other persons, families, or corporate bodies associated with a resource						
Authorizing body	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Person, family, or corporate body to whom a work has been dubiously or erroneously attributed	element	?	role	structured	RDA controlled access point [person / family / corporate body]	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Other person, family, or corporate body associated with a resource	element	n/a	role	structured	RDA controlled access point [person / family / corporate body]	n/a
Access points for producers, publishers, etc.						
Producer, publisher, distributor	element	is produced by +	role	structured	RDA controlled access point [person / family / corporate body]	n/a
Access points for owners, custodians, etc.						
Owner, custodian, etc.	element	is owned by +	role	structured	RDA controlled access point [person / family / corporate body]	n/a
Access points for persons, families, or corporate bodies associated with legal works						
Jurisdiction governed	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Enacting jurisdiction	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Head of state, chief executive, or ruling executive body	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Legislative body	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Issuing agency or agent	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Promulgating agency or agent	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Court governed	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Body governed	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Issuing jurisdiction	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Signatory to a treaty, etc.	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
International conference resulting in a treaty, etc.	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Court, tribunal, etc.	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Reporter [of court proceedings]	element	n/a	role	structured	RDA controlled access point [person]	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Person or corporate body prosecuted	element	n/a	role	structured	RDA controlled access point [person / corporate body]	n/a
Person or corporate body bringing the action	element	n/a	role	structured	RDA controlled access point [person / corporate body]	n/a
Person or corporate body on the opposing side	element	n/a	role	structured	RDA controlled access point [person / corporate body]	n/a
Judge	element	n/a	role	structured	RDA controlled access point [person]	n/a
Party to a case	element	n/a	role	structured	RDA controlled access point [person / corporate body]	n/a
Lawyer representing a party	element	n/a	role	structured	RDA controlled access point [person]	n/a
Access points for persons and corporate bodies associated with religious works						
Person associated with the sacred scripture	element	n/a	role	structured	RDA controlled access point [person]	n/a
Harmonizer	element	is realized by?	role	structured	RDA controlled access point [person]	n/a
Denominational body associated with the creed, etc.	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Person associated with the creed, etc.	element	n/a	role	structured	RDA controlled access point [person]	n/a
Church or denominational body associated with the liturgical work	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Body within the church, etc., associated with the liturgical work	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Access points for persons and corporate bodies associated with official communications						
Corporate access point for official issuing the communication	element	n/a	role	structured	RDA controlled access point [corporate body]	n/a
Personal access point for the person issuing the communication	element	n/a	role	structured	RDA controlled access point [person]	n/a
Access points for persons associated with academic disputations						
Praeses	element	n/a	role	structured	RDA controlled access point [person]	n/a
Defendant, respondent, etc.	element	n/a	role	structured	RDA controlled access point [person]	n/a
RELATED RESOURCES						
Designation of relationship	element	n/a?	role	unstructured	n/a	n/a
Relationship between a work and an expression of the work						
Links between records for work and expression	element	is realized through / is a realization of	role	structured	n/a	n/a
Controlled access point for expression	element	is realized through / is a realization of	role	structured	RDA controlled access point [expression]	n/a
Composite record	description	is realized through / is a realization of	role	unstructured	n/a	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Relationship between a manifestation and a work or expression embodied in the manifestation						
Links between records for manifestation and work (or expression)	element	embodies/ is embodied in	role	structured	n/a	n/a
Controlled access point for work (or expression)	element	embodies/ is embodied in	role	structured	RDA controlled access point [work / expression]	n/a
Composite record	description	embodies/ is embodied in	role	unstructured	n/a	n/a
Relationship between an item and a manifestation exemplified by the manifestation						
Links between records for manifestation and item	element	is exemplified by / exemplifies	role	structured	n/a	n/a
Composite record	description	is exemplified by / exemplifies	role	unstructured	n/a	n/a
Equivalence relationships						
Equivalent manisfestation	element	reproduction / alternate	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	reproduction / alternate	role	structured	n/a	n/a
Name	element sub-type	reproduction / alternate	role	structured	RDA controlled access point [manifestation]	n/a
Description	description	reproduction / alternate	role	unstructured	n/a	n/a
Equivalent item	element	reproduction	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	reproduction	role	structured	n/a	n/a
Name	element sub-type	reproduction	role	structured	RDA controlled access point [item]	n/a
Description	description	reproduction	role	unstructured	n/a	n/a
Derivative relationships						
Source work (or expression)	element	summarization / adaptation / transformation / imitation / abridgement / revision / translation / arrangement (music)	role	structured / unstructured	see below	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Resource identifier	element sub-type	summarization / adaptation / transformation / imitation / abridgement / revision / translation / arrangement (music)	role	structured	n/a	n/a
Name	element sub-type	summarization / adaptation / transformation / imitation / abridgement / revision / translation / arrangement (music)	role	structured	RDA controlled access point [work / expression]	n/a
Description	description	summarization / adaptation / transformation / imitation / abridgement / revision / translation / arrangement (music)	role	unstructured	n/a	n/a
Derivative work (or expression)	element	summarization / adaptation / transformation / imitation / abridgement / revision / translation / arrangement (music)	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	summarization / adaptation / transformation / imitation / abridgement / revision / translation / arrangement (music)	role	structured	n/a	n/a
Name	element sub-type	summarization / adaptation / transformation / imitation / abridgement / revision / translation / arrangement (music)	role	structured	RDA controlled access point [work / expression]	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Description	description	summarization / adaptation / transformation / imitation / abridgement / revision / translation / arrangement (music)	role	unstructured	n/a	n/a
Descriptive relationships						
Described entity	element	n/a	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	n/a	role	structured	n/a	n/a
Name	element sub-type	n/a	role	structured	RDA controlled access point [work / expression / manifestation / item]	n/a
Description	description	n/a	role	unstructured	n/a	n/a
Describing work (or expression)	element	n/a	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	n/a	role	structured	n/a	n/a
Name	element sub-type	n/a	role	structured	RDA controlled access point [work / expression]	n/a
Description	description	n/a	role	unstructured	n/a	n/a
Whole-part relationships						
Whole work (or expression)	element	whole/part [work-to-work / expression-to-expression]	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	whole/part [work-to-work / expression-to-expression]	role	structured	n/a	n/a
Name	element sub-type	whole/part [work-to-work / expression-to-expression]	role	structured	RDA controlled access point [work / expression]	n/a
Description	description	whole/part [work-to-work / expression-to-expression]	role	unstructured	n/a	n/a
Part of a work (or expression)	element	whole/part [work-to-work / expression-to-expression]	role	structured / unstructured	see below	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Resource identifier	element sub-type	whole/part [work-to-work / expression-to-expression]	role	structured	n/a	n/a
Name	element sub-type	whole/part [work-to-work / expression-to-expression]	role	structured	RDA controlled access point [work / expression]	n/a
Description	description	whole/part [work-to-work / expression-to-expression]	role	unstructured	n/a	n/a
Whole manifestation (or item)	element	whole/part [manifestation-to-manifestation / item-to-item]	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	whole/part [manifestation-to-manifestation / item-to-item]	role	structured	n/a	n/a
Name	element sub-type	whole/part [manifestation-to-manifestation / item-to-item]	role	structured	RDA controlled access point [manifestation / item]	n/a
Description	description	whole/part [manifestation-to-manifestation / item-to-item]	role	unstructured	n/a	n/a
Part of a manifestation (or item)	element	whole/part [manifestation-to-manifestation / item-to-item]	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	whole/part [manifestation-to-manifestation / item-to-item]	role	structured	n/a	n/a
Name	element sub-type	whole/part [manifestation-to-manifestation / item-to-item]	role	structured	RDA controlled access point [manifestation / item]	n/a
Description	description	whole/part [manifestation-to-manifestation / item-to-item]	role	unstructured	n/a	n/a
Accompanying relationships						
Augmenting work (or expression)	element	supplement / complement	role	structured / unstructured	see below	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Resource identifier	element sub-type	supplement / complement	role	structured	n/a	n/a
Name	element sub-type	supplement / complement	role	structured	RDA controlled access point [work / expression]	n/a
Description	description	supplement / complement	role	unstructured	n/a	n/a
Primary work (or expression)	element	supplement / complement	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	supplement / complement	role	structured	n/a	n/a
Name	element sub-type	supplement / complement	role	structured	RDA controlled access point [work / expression]	n/a
Description	description	supplement / complement	role	unstructured	n/a	n/a
Accompanying manifestation (or item)	element	n/a	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	n/a	role	structured	n/a	n/a
Name	element sub-type	n/a	role	structured	RDA controlled access point [manifestation / item]	n/a
Description	description	n/a	role	unstructured	n/a	n/a
Primary manifestation (or item)	element	n/a	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	n/a	role	structured	n/a	n/a
Name	element sub-type	n/a	role	structured	RDA controlled access point [manifestation / item]	n/a
Description	description	n/a	role	unstructured	n/a	n/a
Sequential relationships						
Preceding work (or expression)	element	successor	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	successor	role	structured	n/a	n/a

RDA Scope Analysis

RDA element	Classification	Corresponding FRBR attribute / relationship	<ind ecs> attribute type	RDA value string category	Syntax encoding scheme	Vocabulary encoding scheme
Name	element sub-type	successor	role	structured	RDA controlled access point [work / expression]	n/a
Description	description	successor	role	unstructured	n/a	n/a
Succeeding work (or expression)	element	successor	role	structured / unstructured	see below	n/a
Resource identifier	element sub-type	successor	role	structured	n/a	n/a
Name	element sub-type	successor	role	structured	RDA controlled access point [work / expression]	n/a
Description	description	successor	role	unstructured	n/a	n/a

futurelib | Framework

- QuickStart
- Recent Activity
- SideBar

Create a new page
Create a classroom
Create a syllabus
Create a group project
Use another template

Take a tour

Framework for a Bibliographic Future

Draft for discussion, by Karen Coyle, Diane Hillmann, Jonathan Rochkind, Paul Weiss

How to Comment You must login to comment. After you login you will see the *Comments* tab on this page. To login you must provide the wiki-wide password. The password for this wiki is the last name of the man who invented the Dewey Decimal System, plus the last two digits of the famous year of the founding of the United States. If this isn't enough of a hint, contact Karen Coyle for the actual password.

See other frameworks from other authors

Introduction

Metadata is a generic term for the data that we create about persons, places, things, documents, and anything else about which we wish to communicate or wish to operate on in an electronic environment. Although it is common to hear that "all data is metadata," it is certainly the case that not all metadata is well designed. Good design increases the potential success of a metadata standard.

The design components proposed in this model are not new. Similar components are used to some degree in standards such as the OpenURL Framework (Z39.88), the Semantic web, and the Dublin Core Metadata Initiative. A framework such as this serves many purposes. In particular, we are interested in producing metadata that is both highly extensible and that will promote compatibility between communities and applications that extend the metadata.

The four components that we propose are: a **model** of basic structures and relationships, a **schema** that defines an extensible set of properties, **guidance** for application of the properties, and **encoding**. The model can be used to create one or more schemas, and any schema can be expressed using one or more encodings. The guidance document is a key element that provides both direction to creators but also describes the semantics of the data elements in a human-understandable way. These four components provide a basis for creation of machine-manipulable metadata that has meaning to a community yet it can be defined in a rigorous way to communicate clearly to any users of the data.

Model

This is sometimes called an abstract model or a data model, although it does not define the data itself. Models are high-level views of the structures and relationships that the metadata will address. In the library community, the entity-relationship structure provided by FRBR is a type of model. It includes basic aspects of the information universe that will eventually be defined by metadata (works, expressions, manifestations, and items, plus the entities such as person and concept that will have a relationship with the primary four). We need to consider carefully how the FRBR model works in the context of other models that are used for bibliographic data such as the Dublin Core Abstract Model (DCAM) and OpenURL Framework. The model that results will be independent of any particular schemas or encodings of bibliographic metadata, but will provide a structure that all implementations of metadata derived from the model will have in common.

Schema

Metadata schemas (sometimes called 'element sets,' 'metadata formats' or 'data dictionaries') define the actual properties that will carry values in the data set, as well as the relationships between those properties. Data elements can be defined at any relevant level of granularity. They can have hierarchical relationships between them or non-hierarchical relationships. The Dublin Core Element Set is an example of a set of data elements. FRBR defines data elements in its attributes, but they must be restructured in a way that allows the development of different levels of granularity and that promotes extensibility of the schema, both over time and across communities. Ideally, the schema would be expressed in one or more machine-readable formats that facilitate its use by both people and computer applications.

Guidance

Guidance is often desired to aid in the creation or assignment of values to data elements in a consistent way. Guidance may be general or specific, but it usually attempts to address circumstances that users will encounter in the creation of the metadata. Different communities making use of the same data elements may define their own specific best practices that attempt to produce the metadata that is most useful for their purposes, but in general they may not re-define the elements in order to address those needs. The library community has traditionally received its guidance from cataloging rules (such as AACR) and from practices published as part of the encoding of library data using MARC21. Increasingly, specialized guidance for specific communities has been developed that reflects the differences in materials or approach inherent in their tasks: examples are Cataloging Cultural Objects (CCO) for the museum community and Describing Archives: A Content Standard (DACS) from the archival community.

Encoding

We can assume that any metadata being created today will be expressed and exchanged via a machine-readable encoding. The primary requirement for metadata encoding is that it must be able to encode the full detail of the semantics and relationships intended by the metadata creators; and it must expand as the metadata schema grows and changes. The same metadata can be encoded in different data formats and still be fully shareable, as long as the encoding is true to the data elements and to the overall structure of the metadata model.

Discussion

FRBR's entity-relationship model (as defined in Chapters 3-5 of the FRBR Report) is a useful, if not complete or even wholly accurate, analysis of our bibliographic universe. The delineation of the four group 1 entities illuminates an important issue of our legacy: we have been cramming metadata about different bibliographic entities into single descriptions. As just one example, the FRBR report provides an explanation for the ambiguity of dates in bibliographic records: there are at least four dates of creation that apply to each bibliographic resource--those of its work, expression, manifestation, and item. For many resources all these are the same, so there is no need to delve further, but some resources are more complex, and that complexity has led to confusion about dates used in brief displays and search limits.

FRBR, and work by Barbara Tillett, Richard Smiraglia, and others has contributed to an increasingly formalized notion of relationships among bibliographic resources, and between bibliographic resources and associated entities (for instance, FRBR's group 2 entities--persons, corporate bodies--and draft FRAR's families, as well as subject entities). Examining current practices from the perspective of this work on relationships shows great inadequacies in the identification, recording, and utility of relationships.

FRBR does an admirable job of providing one way to analyze the bibliographic universe, though as has been noted by others, it doesn't extend well to museum or archival collections. Although FRBR covers attributes of bibliographic entities, it does not model the metadata itself (that is, none of the entities represents metadata per se).

DCAM

The Dublin Core Abstract Model from the Dublin Core Metadata Initiative (DCMI), on the other hand, takes the next logical step, and models metadata. Its purpose is to "to gain a better understanding of the kinds of descriptions that we are trying to encode and facilitates the development of better mappings and translations between different syntaxes."

The FRBR model and the Dublin Core Abstract Model are not contradictory; in fact, they are complementary. FRBR provides a start at defining properties for RDA and allows the description of resources using specific relationships that can be assigned at the proper level as well as aggregated for better expression to the user. The DCAM helps us to envision the FRBR entities as a package, allowing the discussion about issues like identity and linking to be posed and discussed in a more useful manner.

Metadata Schemas

Even as we validate the use of FRBR as a model, we take issue with its embedded attributes. One of the things the DCAM and the Dublin Core experience generally tells us is that we need to develop our attributes/properties/elements separately from the model as well as from the values used. Separating elements and their definitions from guidance on determining their values (controlled vocabularies, transcription, etc.) is crucial in order to achieve interoperability and extensibility.

As a first step, the FRBR attributes must be carefully generalized. For example, instead of defining separate elements (including their names, definitions, examples, etc.) for title of the work, title of the expression, and title of the manifestation, there should be one title element reused at multiple levels. The declaration of these elements should include clear specification of where in the FRBR Group I they may be used. This increased generalization promotes interoperability, minimizes a tendency toward complexity, and eases machine manipulation and extensibility. It also requires more rigorous consideration of when attributes at the various levels are really the same thing or not, and can point out inconsistencies that can be rectified. Along with the

development of the generalized elements, DCAM should be capable of extension or refinement of those elements, to ensure that appropriate extensions can be made and managed.

Crucial to the proper development of a metadata schema is a clear notion of requirements for technical expression of the attributes, and a plan for maintenance and growth. We have learned much in the library community about the importance of community consensus and how to maintain important standards over time. MARBI is a good example of doing it correctly, and in fact the Dublin Core Usage Board process is based loosely on MARBI.

Guidance for Application

It is critically important that we develop good usage guidance based first on the Metadata Schema attributes in their most generalized form. We must provide this usage guidance in a manner that allows communities of practice to use the general guidance as they extend the basic structure for their own purposes. Traditional library cataloging is just such a community of practice, and should extend the schema and guidance to fit their needs, without the necessity of bringing their special library colleagues along with them. If the general elements, and the guidance attached specifically to them, can be approached as a extensible set, other communities will be encouraged to incorporate them specifically in their metadata and to extend in ways that provide a sound basis for interoperable use and re-use. In this scenario, mapping between library metadata schemas and others, as well as the mix/match capabilities of application profiles, can be made easier. This approach will tend to minimize data loss when information is crosswalked, and improve the ability of machines to act upon the data regardless of its origin.

As part of this development of extended guidance material along specialist lines, we need to recognize that different communities will apply FRBR Group I boundaries differently. Much of the discussion about how decisions will be made about works, expressions and manifestations indicates clearly that specialized communities will tend to make different decisions about where these boundaries lie. This has been seen as a problem, and an impediment to the integration of FRBR principles into actual practice. Part of the rationale for separating traditional library specific instruction from the general RDA, and enabling specific communities to extend from that general base, is that the assumptions and instructions for where these boundaries lie can be made explicit by community, and librarians can get out of the trap of trying to herd everyone else into the same decisions. This will make it easier for the communities as a whole to use each other's work--when differences are not exceptions but can be explicitly expressed as policies and appropriately supported with more detailed extensions to the general framework, everyone enjoys easier and more cost effective machine manipulation of data. So long as the determination of what is a work can be ascribed to the community that made the decision, other communities can predict and cope with the variations.

Encoding

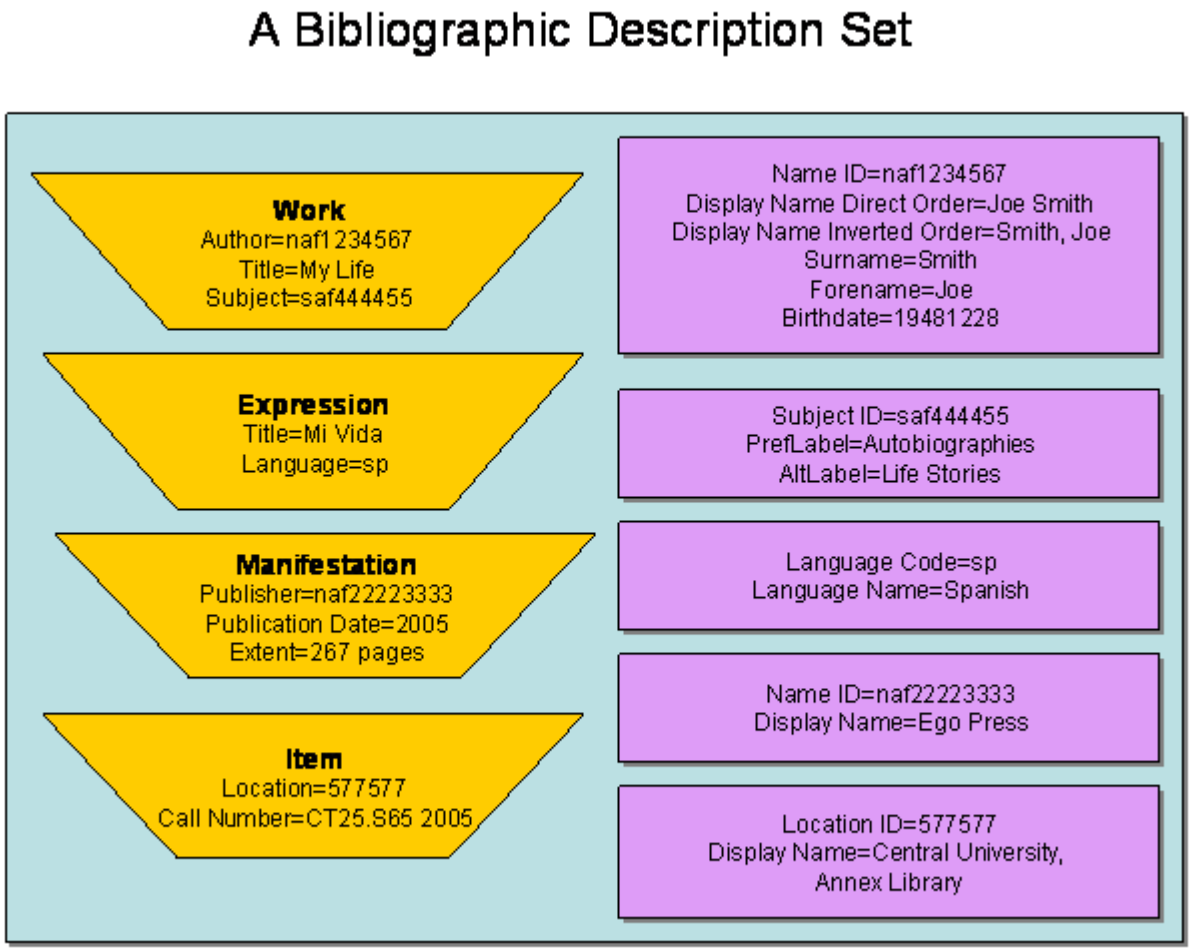
It seems unlikely that MARC21 can be sufficiently remodeled to serve as an encoding for a modern metadata schema, but certainly some of the accumulated wisdom and experience embedded in the MARC21 documentation can be repurposed. One issue is that insofar as it supplies definitions, labels and relationships not necessarily explicit in AACR2, MARC21 itself represents a combination of functions that requires significant attention, and perhaps deconstruction, to prise out what should be included in the metadata schema and what remain as encoding.

It should also be recognized that MARC21 encodes more than bibliographic information, and the formats for classification, authorities and holdings might well be more appropriate for future use, given that they operate where competing data structures are sparse. Where they tend to be problematic is in the area of distinctions at

the statement level, where specification of language of statement, source, and community information may well be necessary.

Encoding for the future must support statement level identification and attribution. Although to a certain extent, this is a 'packaging issue,' it seems important to assert it as a guiding principle, as it supports the notion that the way records will be built in future will be much more iterative, and catalogers are just as likely to start with a re-used description than one created newly for purpose. These catalog records of the future are likely to be aggregations of the work of many catalogers--somewhat like CONSER records are now--and the source and age of particular statements will be critical as we develop applications to make 'decisions' about what statements they will display. Central to this assumption is that, in the shared environment of the future, information may be added, but not subtracted--just ignored if not needed or desired in a particular context.

An Example



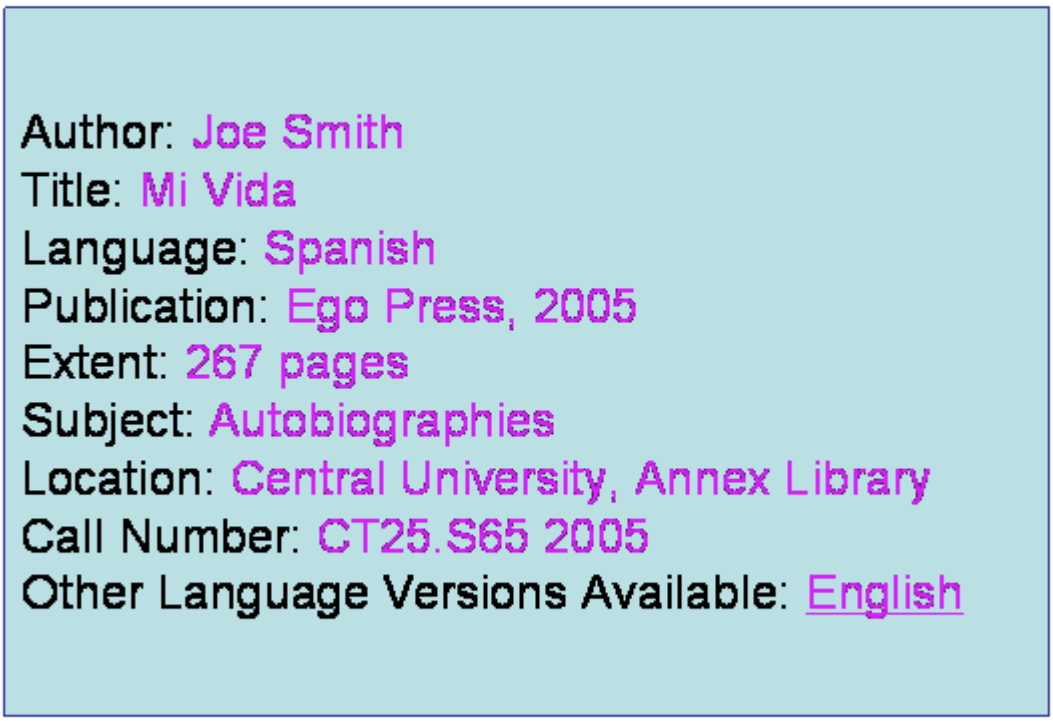
The figure above illustrates some possibilities for a description set based on DCAM that also includes some of the FRBR entities and shows how they would relate. On the left side are the four Group I entities, with a small assortment of generic properties. In the cases where the value of the properties is contained in another description, the relationship between them is conveyed with an identifier, and the identified Group 2 or 3 description is included in full with the description set. Thus, an application using this description set could

presumably look and chose among the available values for the one that suits its purpose best. For instance, in the description of the author, there are two identified possibilities for display text for that particular person, one using direct order, and the other surname first.

Note that the linking techniques are the same regardless of what kind of description, whether author, publisher or subject is related to a particular Group 1 entity. There is both a title in the Work description and another in the Expression--the differences between them and their different functions are conveyed not in the property name, but in where it appears, allowing an application to determine how to display either or both. Grouping of expressions and manifestations can be supported using simple linking and naming strategies, without unnecessary complexity.

Using only the descriptions in this simple example, the following display could be supported:

A Simple Display



Note that the link to an English version is implied by the presence of another description set (not illustrated here) with the same work description and an expression description in English.

Towards an Interoperability Framework for Metadata Standards

Mikael Nilsson
KMR Group, NADA, Royal Institute of Technology, Stockholm, Sweden
mini@nada.kth.se

Pete Johnston
Eduserv Foundation, United Kingdom
pete.johnston@eduserv.org.uk

Ambjörn Naeve
KMR Group, NADA, Royal Institute of Technology, Stockholm, Sweden
amb@nada.kth.se

Andy Powell
Eduserv Foundation, United Kingdom
andy.powell@eduserv.org.uk

1. Abstract

This paper presents a conceptual metadata framework for Dublin Core metadata, intended to support the development of interoperable metadata standards and applications. The model rests on the fundamental concept of an "abstract model" for metadata, as exemplified by the DCMI Abstract Model, and is based on concepts and ideas that have developed over the years within the Dublin Core Metadata Initiative.

The model thus incorporates the concepts of metadata vocabularies, schemas, formats and application profiles into a single framework that can be used to analyse and compare metadata standards, and aid in the process of harmonization of metadata standards. IThe model is used to briefly compare the structures of the Dublin Core metadata specifications and the IEEE LOM standard. Some fundamental differences between the two standards are discussed briefly, and important gaps in the current set of Dublin Core metadata specifications are noted.

Keywords:

Dublin Core, abstract model, semantic interoperability.

2. Background

The publication of the DCMI Abstract model (DCAM) (Powell et al, 2005) in March 2005 marked a major milestone for the Dublin Core community and the DCMI. In developing

the DCAM, the DCMI has shown its intention to gradually move away from dealing primarily with the "core" set of terms, moving instead to dealing primarily with community-specific application profiles, each defined within a common framework (Baker, 2005). Within such a framework, metadata terms from different and independent communities can co-exist, allowing for a controlled mix-and-match of community- and application-specific metadata constructs.

Although the framework used by the the Dublin Core community is still not formalized by the DCMI, considerable experience and documentation regarding the necessary components of such a framework have been collected over the years. It is the intention of this paper to introduce an over-arching model to describe the components of this framework, to serve as a possible basis for further formalization, and to highlight the strong and weak points of the current situation.

The model proposed in this paper is also intended to serve as a guide to understanding the conceptual relationships between the structures of the many different metadata standards currently in use. We will demonstrate this by using the model as a tool to compare the structure of the Dublin Core metadata framework with the IEEE LOM standard. Although the model has its origins in the Dublin Core metadata framework, we believe the model has a substantially more general applicability.

This attempt at designing a framework for Dublin Core metadata shares some features with the Warwick Framework (Lagoze, 1996), although that framework focused more on the packaging of metadata descriptions than on the nature of those metadata descriptions and the interoperability of the standards and specifications on which those metadata descriptions were based. The RDF suite of specifications, however, follow a more similar pattern to the framework presented here.

In other, related contexts, many similar kinds of frameworks have been designed over the years.

- The UML 2.0 specifications in general and the UML Meta Object Facility in particular, share some basic modeling principles with the framework presented here, albeit with a markedly higher level of complexity, and a primary focus on model-driven design.
- The MPEG-7 multimedia metadata framework also contains a complete framework for metadata vocabulary management, but with little emphasis on use in other contexts than multimedia.
- The ISO 11179 framework is of particular significance for describing metadata and metadata models, but is not concrete enough without further specialization to cater for the needs of real-world metadata interoperability.

A fuller analysis would require a much more thorough discussion. However, it can still be concluded that in comparison with these and other related frameworks, the most important distinguishing features of the Dublin Core metadata framework presented here is its relative simplicity, straightforwardness and cross-domain applicability.

3. The DCMI Metadata Framework and its Components

In this section, the metadata framework used by the Dublin Core community is examined and a set of components of that framework for Dublin Core metadata are identified: the

abstract model, *metadata formats*, *metadata vocabularies*, the *vocabulary model*, *application profiles* and the *profile model*. Some of these components correspond to concepts that have been formalized by DCMI (as DCMI recommendations or other documents); other components represent abstractions based on current usage of Dublin Core metadata and on current directions in metadata interoperability.

3.1 The Abstract Model

The abstract model specifies the concepts used in the framework, the nature of terms and how they combine to form an information structure. An early effort to produce such framework for Dublin Core was presented in Bearman, Miller, Rust, Trant and Weibel (1999).

Subsequently the DCMI Usage Board developed the "DCMI Grammatical Principles" (DCMI Usage Board, 2003), as a summary expression of the key concepts underpinning the vocabularies developed by the DCMI. The DCMI Abstract Model, published in March 2005, was a substantial reformulation and clarification of these principles.

The DCMI Abstract Model defines the *description set* as the principal information structure used in Dublin Core metadata. It describes the nature of the components that make up that information structure and it also describes how that composite information structure is to be interpreted.

In summary, a *description set* is described as follows:

- a *description set* is made up of one or more *descriptions*
- a *description* is made up of
  - zero or one *resource URI* and
  - one or more *statements*
- a *statement* is made up of
  - exactly one *property URI* and
  - zero or one reference to a value in the form of a *value URI*
  - zero or more representations of a value, each in the form of a *value representation*
  - zero or one *vocabulary encoding scheme URI*
- a *value representation* is either
  - a *value string* or
  - a *rich representation*
- a value string may have an associated *value string language*
- a value string may have an associated *syntax encoding scheme URI*
- each value may be the subject of a *related description*

A DC metadata description set is to be interpreted as a set of assertions about the resources identified by those URIs, principally about the relationships between the two resources identified by the *resource URI* and the *value URI*.

The abstract model is the key used by a metadata application to unlock the secrets of a metadata expression given in a specific format, thus making it possible for a single standard, though expressed in several different formats, to still be understood in a uniform way by users and applications.

3.2 Metadata Formats

The abstract model describes an abstract information structure. Metadata applications construct and exchange instances of that abstract information structure, and they do so by representing the information structure as a digital object, using the rules specified by one of several *metadata formats* or *bindings*. In the case of Dublin Core, DCMI has published a set of "encoding guidelines" specifications which provide bindings for DC metadata.

A binding is constructed by specifying how each kind of concept in the abstract model is to be encoded in a particular format. Conversely, the binding also specifies how to interpret data given in a specific format in terms of the abstract model. For example, when interpreting a metadata record that uses the Dublin Core XML binding, an XML element called "dcterms:modified" used in a particular place in the XML document represents a property, and the value "dcterms:W3CDTF" of a particular XML attribute represents a syntax encoding scheme for the value string "2001-07-18" occurring as XML content in a particular position.

This fundamental process of *encoding/interpretation* is described in Figure 1. Application A uses the DCMI Abstract Model to represent some metadata about a resource. This

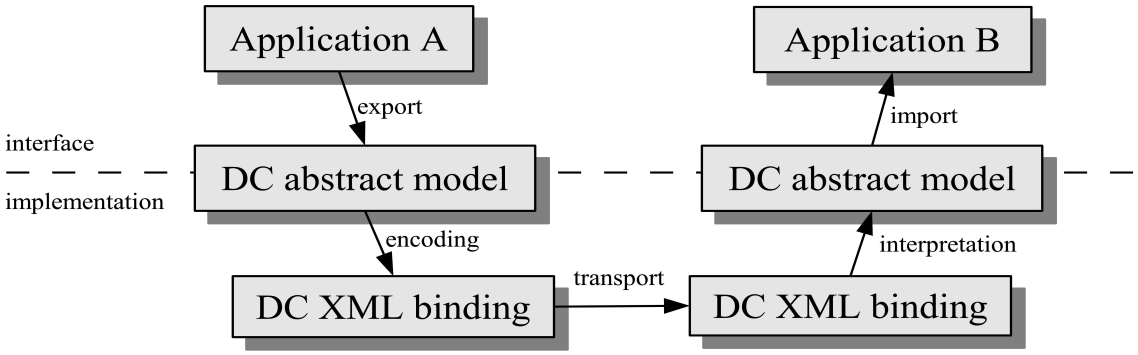


Figure 1. The process of encoding/interpretation of metadata within the framework of an abstract model.

metadata is encoded using the Dublin Core XML binding, and transferred to another application. Application B will use the rules of the Dublin Core XML binding to interpret the XML data in terms of the DCMI Abstract Model. This representation of the metadata can then be used in the application.

When two applications want to exchange Dublin Core metadata, they understand metadata through the lens of the abstract model. The abstract model functions as an opaque interface, an API, to the metadata. In practice, the exchange is realized using one of the Dublin Core bindings, but the details of the formats are of no interest to the applications, which instead analyse the metadata in terms of the interface given by the abstract model.

Note that it is possible to produce applications that process metadata without regard to the abstract model. Such *ad-hoc processing* of metadata records requires that the precise content of the records is well-known in advance, which is the case in many systems where extensibility, modularity and refinements are not design requirements. In contrast, the kind

of *interoperable processing* based on the abstract model described above is necessary when an application needs to be prepared for metadata constructs that do not fall within the limits of such a precise, pre-conceived description. Thus, it should be clear that interoperable processing is a basic prerequisite for metadata interoperability.

3.3 Metadata Vocabularies

Although the abstract model specifies the nature of the terms that are used in a DC metadata description set, it does not list any fixed set of terms to be used. On the contrary, the Dublin Core metadata framework is based on the notion that the *vocabularies* used in DC metadata description sets are created and maintained separately from the abstract model.

Although the initial focus of the DCMI was on building consensus around the use of a small set of metadata terms that could be used to create fairly simple descriptions of a wide range of resources – the fifteen properties (or "elements") of the Dublin Core Metadata Element Set – the experience of implementing DC metadata highlighted that in practice these terms were supplemented with other terms to meet the requirements of some particular community or application context.

In Dublin Core metadata, a vocabulary can be one of two things:

1. A *value vocabulary*, consisting of concepts from a controlled set as specified by a vocabulary encoding scheme. For example, the "dcterms:LCSH" vocabulary encoding scheme refers to the vocabulary formed by the set of Library of Congress subject headings.
2. An *element vocabulary*, consisting of a set of metadata properties together with their definitions. For example, the Dublin Core Element Set, consisting of the 15 original Dublin Core elements (*dc:title*, *dc:subject*, etc.), is such a vocabulary.

Element vocabularies and value vocabularies have fundamentally different characteristics. While value vocabularies are used to construct taxonomies and thesauri that describe relationships between concepts in terms of broader/narrower, containment etc, element vocabularies are used to construct application profiles, schemas and ontologies that describe how metadata instances are to be constructed.

3.4 Vocabulary Model

As the Dublin Core community embraced the notion that DC metadata might utilize multiple metadata vocabularies, they also recognized that specific types of relationship could exist between the metadata terms referenced in DC metadata – both between terms within a single vocabulary and between terms in different vocabularies. An example of such a relationship between terms is that of "element refinement" where one property is described as a specialization of another property.

Consensus on the nature of these relationship types is the basis of an implicit *vocabulary model*. Clearly that vocabulary model is closely related to the DCMI Abstract Model since it is concerned specifically with the types of terms described by the abstract model, and the relationships between terms of those types.

If applications are to be able to act on information about such relationships between metadata terms, then those terms and the relationships between them must be described in

a machine-processable form, i.e. a language for describing metadata vocabularies is necessary. Such a vocabulary language enables the description of element and value vocabularies in a form which enables applications to access information about the nature of the terms and their relationships with other terms in the same or in different vocabularies.

The Dublin Core vocabulary model has not yet been formalized, but embryos such as Baker (2003) exist. DCMI has a history of using RDF Schema (Brickley *et al* 2004) as a basis for its machine-readable term declarations. RDF Schema is useful for describing both element and value vocabularies.

3.5 Application Profiles

The Dublin Core metadata standard emerged from an interest in developing a resource description standard that could be applied across a broad range of communities and domains. Since its inception, the DC community had the expectation that Dublin Core would be deployed alongside other metadata standards. They also learned from experience that implementers tailored the standard to fit the requirements of their own context.

More recently, these two trends have converged in the notion of the DC application profile, and the principle that implementers of metadata standards should be able to assemble the components that they require for some particular set of functions - and if that means drawing on components that are specified within different metadata standards, that should be possible.

Duval et al (2002) employ the metaphor of the Lego set to describe this process: an application designer should be able to "snap together" selected "building blocks" drawn from the "kits" provided by different metadata standards to build the construction that meets their requirements, even if the kits that provide those blocks were created quite independently.

Heery and Patel (2000) present a compelling vision of metadata implementers "mixing and matching" "data elements", constructing application profiles by selecting from the sets of "data elements" provided by metadata standards and by other implementers. Such application profiles are fundamental to a modern metadata framework.

Just as the description set construct defined by the DCMI Abstract Model embraces the description of a number of related resources, so too a DC application profile may specify the construction of the related descriptions of several kinds of related resources, such as a collection, the items it consists of and the associated contributors. Thus, such a specification is a multi-layered structure of some complexity, that can not, in general, be captured by a flat list of properties.

3.6 Profile Model

Although the concept of the DC application profile has gained general acceptance within DCMI and the DC implementer community, it has not yet been formalized by DCMI in the form of a model for a DC application profile.

Like the vocabulary model, the profile model is closely related to the abstract model, because it is concerned with specifying the creation of the particular information structures described by the abstract model – in the case of Dublin Core, description sets, as defined by the DCMI Abstract Model.

Any such model must not be tied to a specific metadata format, but must operate at the level of the abstract model, so that the application profile can be applied independently of the metadata format in which metadata instances are encoded.

Promising work on machine-processable DC application profiles can be seen in, e.g.,"Guidelines" (2005)

3.7 The DCMI Metadata Framework

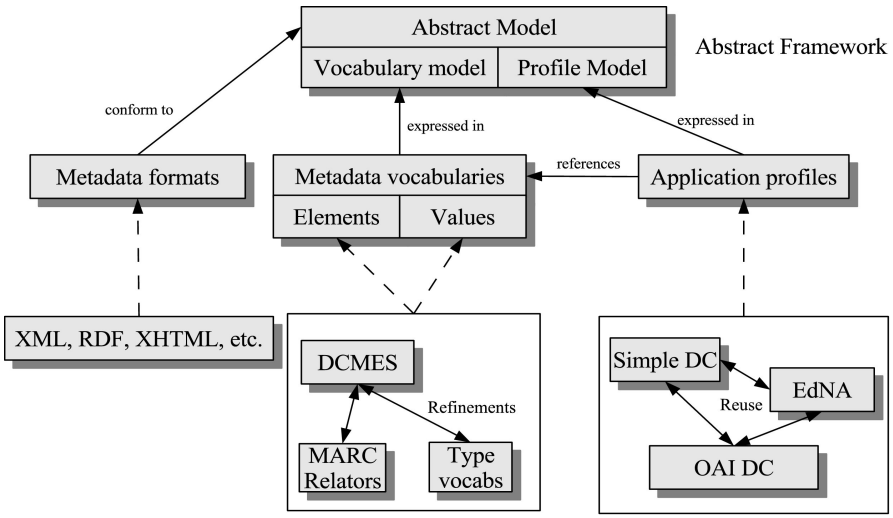


Figure 2. A model of the Dublin Core metadata framework.

This brief survey of DCMI specifications and DC metadata usage highlights the existence of a number of inter-related and inter-dependent features, which when viewed together can be seen, implicitly at least, as components of a larger framework. The relationships between these component parts of the Dublin Core metadata framework are depicted in Figure 2.

The diagram highlights the close relationship between the DCMI Abstract Model and the DC vocabulary model and DC application profile model.

4. A comparative view: applying the framework model to LOM, Dublin Core and the Semantic Web

This section seeks to generalize this model of a metadata framework and to apply it to the analysis of two other metadata standards, and to identify the corresponding components within the frameworks deployed by those standards.

The following table presents a summary view of the framework components as they are present within the IEEE LOM standard and within the Semantic Web suite of specifications, and indicates the extent to which each component is formally distinguished from other components within the framework. Note that by "Dublin Core framework" we refer to the complete set of DCMI specifications, and similarly for LOM.

Framework concept	*Dublin Core framework*	*LOM framework*	*Semantic Web framework*
Abstract Model	DCMI Abstract Model	Implicit in LOM Data Model	RDF Concepts and Abstract Syntax
Metadata Formats	XML, RDF and HTML bindings	XML binding	RDF/XML syntax, N-triples, etc.
Metadata Element Vocabularies	DCMES, large set of external properties and encoding schemes	LOM Data Model includes element vocabulary, various extensions to LOM	Many external element vocabularies
Metadata Value Vocabularies	DCMIType vocabulary. Many external value vocabularies	LOM Data Model includes several basic value vocabularies, many external vocabularies	Many external value vocabularies
Vocabulary Model	Not formalized, but see Baker (2003)	Not formalized	RDF Vocabulary Description Language
Application Profiles	Some published by DCMI, many external application profiles	LOM Data Model includes basic application profile, many external application profiles.	Many in the form of ontologies
Profile Model	Not formalized, but cf "Guidelines".	Not formalized	Possibly OWL, the Web Ontology Language

A few comments on this table:

- Not all parts are formalized. The DCMI is slowly progressing towards formalizing the complete abstract framework, including abstract model, vocabulary model and profile model. Similar efforts are not under way in LOM.
- The most mature parts are certainly value vocabularies, where many external sources exist. Dublin Core metadata element vocabularies are also relatively mature. To some extent, and to some extent application profiles have some maturity, even though there is still a certain amount of confusion in the community regarding the precise nature of an application profile.
- In spite of the existence of many application profiles and metadata vocabularies, no formal model is usually followed in their design.
- LOM has a relatively weak notion of element vocabularies, as noted in Nilsson et al (2006), that does not support URI identification of elements.
- The LOM Data Model defines, in a single standard, both an abstract model (implicitly, at least), a metadata element vocabulary, a set of metadata value vocabularies, and a basic application profile. This is one way of expressing the well known "monolithic" nature of the LOM standard.
- Further comparison with e.g. MODS, MPEG-7 etc. remains the subject of a future article.

In short, the above table can be used to analyse and compare metadata standards, and understand how they relate to different aspects of the Dublin Core universe.

5. Interoperability across metadata frameworks

Although the use of the model has enabled us identify the corresponding components within the frameworks of the different standards, significant differences may still exist between the corresponding components in the different frameworks. For example, although both the Dublin Core metadata standard and the LOM metadata standard incorporate the notion of an abstract model (either explicitly or implicitly), those two

abstract models are quite different: the conceptual information structures that they describe, and the nature of the terms used in those conceptual information structures, are quite different – and those differences carry over into the corresponding vocabulary models and the profile models. In the cases of Dublin Core and the Semantic Web specifications, again there are differences between the two abstract models, but they are more similar than in the case of Dublin Core and LOM. In the case of Dublin Core and the Semantic Web specifications, the two abstract models are broadly compatible, and this is reflected in DCMI's use of the RDF Vocabulary Description Language to describe its vocabularies.

Such differences become critical when we begin to consider interoperability across different metadata standards constructed within their own metadata frameworks. A significant part of the motivation for the development of the profile models within both the Dublin Core and LOM frameworks was precisely to facilitate the (re)use of metadata vocabularies across the boundaries of the two corresponding frameworks. While those models have certainly increased interoperability *within* the respective frameworks, interoperability between the different frameworks remains a difficult problem.

With a similar aim in mind, the CORES Resolution (Baker and Dekkers, 2002), which has been signed by both the IEEE LTSC and the Dublin Core Metadata Initiative, encouraged the owners of metadata standards to assign URI references to their "elements", the "units of meaning comparable and mappable to elements of other standards". The assignment of a URI to an "element" means that it can be unambiguously cited in a global context, and this is a necessary condition for the sort of mixing and matching foreseen by Heery and Patel. However the assignment of a URI to an "element" does not change the nature of that "element": and it does not make it meaningful to use the URI of a LOM data element as, e.g., a property URI in a Dublin Core metadata description. Similar incompatibilities have been noted between, e.g., RDF and MPEG-7 (van Ossenbruggen, Nack and Hardman, 2004 and Nack, van Ossenbruggen and Hardman, 2005).

The analysis in Nilsson et al (2006) shows that we must not confuse the components used in a metadata format and the constructs in the abstract model. The components in a metadata format, such as "element URIs" may seem to be similar and compatible, but in reality they belong to completely different frameworks that might not be compatible. Thus, according to the analysis in Nilsson et al (2006), the notion of reusing "elements" between metadata standards and formats using incompatible frameworks is fundamentally flawed. While assigning URIs for the component parts of a metadata standard is clearly a worthwhile effort in other ways, this does not really address the fundamental issue when creating interoperable metadata standards, namely the compatibility of their respective frameworks, and in particular, their abstract models.

Basing metadata on a compatible abstract models carries a number of important benefits

- Clear guidelines on how to create and maintain customized metadata vocabularies. There is currently some confusion on how to best produce vocabularies, much due to the differing fundamental principles for vocabularies in the different metadata standards.
- Fine-grained control over relationships between terms from different standards, including refinement and partial mappings. Automation of interoperable metadata management systems will be greatly improved, and metadata vocabularies will be

able to build upon each other.

- A single set of format bindings. Contrast this with the current situation, which requires every metadata standard to have its own set of format bindings. This will make life easier not only for metadata standardization bodies, but also for applications that will only need to support one format.
- A single framework for extending and combining metadata from different standards. This will enable standardized principles for the construction of interoperable application profiles.
- A single storage and query model for very different types of data and vocabularies. For example, storing metadata from different specifications in the same database will become more straightforward. Implementing searching that includes dependencies between metadata expressed in different schemas will be simplified.

6. The word "Metadata Standard"

In light of the model presented here, it seems clear that the current use of the term "metadata standard" or "metadata schema" will need refinement. These terms are often used interchangeably to describe one of the following:

- The over-arching abstract model standard. This will also include a specification for how to express the semantics of vocabularies adhering to the abstract model (the vocabulary model) as well as a specification for how to express application profiles in a machine-processable way (the profile model).
- Metadata format specifications. These will include bindings of the abstract model to a set of formats and systems, including XML, database layouts, programming languages, etc., as well as translations or mappings to other knowledge representation systems such as RDF. Such specification are closely tied to the abstract model.
- Metadata vocabularies. These will include metadata terms from different communities. The Dublin Core terms, the LOM elements and so on are examples of metadata element vocabularies, and a large set of value vocabularies also fit into this category.
- Application profiles. These will specify usages of metadata vocabularies in complex combinations.

Clarification of the underlying framework can hopefully contribute to better terminology in this domain.

7. Looking forward

We have presented an overarching framework for Dublin Core metadata, based on the implicit structure of current Dublin Core metadata standardization and practise.

The authors believe that the Dublin Core Metadata Initiative would be greatly helped by applying this understanding to improve its documentation and vision of metadata interoperability. In particular, a high-level framework for Dublin Core metadata has not

been proposed since the Warwick framework, and it is now time to revisit the overall structure of metadata standardization. Luckily, as the analysis shows, there is some coherency in the current set of DCMI specifications, though much of it remains implicit. Making the overall structure explicit has the advantages of increasing coherency of terminology, making it easier to communicate the relative significance of each specification, simplifying for users to understand how metadata constructs may be used and reused, and more.

Another issue is that of interoperability with other metadata standards. By reinterpreting the framework in terms of LOM and the Semantic Web, we learn about differences between the metadata standards and deficiencies in their respective frameworks. The authors have little hope that deep integration between metadata standards can be made a reality unless they adhere to a single common framework. Unfortunately, a thorough analysis shows (Nilsson *et al,* 2006) that there are fundamental incompatibilities between frameworks such as the LOM framework and that of Dublin Core. On the other hand, the framework of RDF and the semantic web share many features with Dublin Core, and advanced interoperability between those frameworks has already been demonstrated.

The authors therefore argue that the long-term solution is to proceed towards a *shared* metadata framework. Having all metadata standards expressed using a common abstract model, or at least using compatible abstract models, would greatly increase interoperability in several ways. It would also create a natural separation between the specification of the structure of metadata descriptions and the declaration of metadata terms used within that structure, so that both LOM vocabularies and Dublin Core vocabularies would appear as metadata vocabularies within that one structure. Great care must be taken to ensure that such an abstract model does not conflict with the emerging metadata format for the Web: RDF.

There are already initiatives to develop a common abstract model that covers both LOM and Dublin Core, but unfortunately it seems to be impossible to arrive at such a model without re-engineering at least one standard to retrofit it to the new abstract model, which naturally is a major undertaking. An alternative approach is to produce “compatibility layers” that allow one metadata standard to be described and used in a different framework based on a common abstract model. An example of this is the development of a mapping of LOM to the DCMI Abstract Model (See “Joint DCMI/IEEE LTSC Task Force”). Reaching out to embrace the other important metadata standards, such as MODS, MPEG-7 and the IMS set of standards is then the logical next step.

The basis of the envisioned metadata standardization framework is the abstract model. The incompatibilities of abstract models are the most significant stumbling blocks for metadata interoperability. The development of a common abstract model for metadata is therefore of central importance if we are ever going to experience true metadata interoperability.

8. References

Baker, T. (2003), DCMI Usage Board Review of Application Profiles. Retrieved July 1 2005, from http://dublincore.org/usage/documents/profiles/

Baker, T. & Dekkers, M., (2002), CORES Standards Interoperability Forum Resolution on Metadata Element Identifiers. Retrieved July 1, 2005, from http://www.cores-eu.net/interoperability/cores-resolution/

Baker, T. (2003), Usage Board Application Profile (draft). Retrieved April 30, 2006, from http://dublincore.org/usage/meetings/2003/06/Usage-ap.html

Baker, T. (2005), Diverse Vocabularies in a Common Model: DCMI at ten years, Keynote speech, DC-2005, Madrid, Spain. Retrieved April 30, 2006, from http://dc2005.uc3m.es/program/presentations/2005-09-12.plenary.baker-keynote.ppt

Bearman, D., Miller, E., Rust, G., Trant, J. & Weibel, S. (1999), A Common Model to Support Interoperable Metadata, *D-Lib Magazine*, January 1999. Retrieved July 1, 2005, from http://www.dlib.org/dlib/january99/bearman/01bearman.html

Brickley, D. & Guha, R. V. (2004), RDF Vocabulary Description Language 1.0: RDF Schema, *W3C Recommendation 10 February 2004*. Retrieved July 1, 2005, from http://www.w3.org/TR/rdf-schema/

DCMI Usage Board (2003), DCMI Grammatical Principles. Retrieved April 30, 2006 from http://dublincore.org/usage/documents/principles/

Dublin Core Application Profile Guidelines (2003), CEN Workshop Agreement CWA 14855. Retrieved July 1, 2005, from ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/MMI-DC/cwa14855-00-2003-Nov.pdf

Duval, E., Hodgins, W., Sutton, S. & Weibel, S. L. (2002), Metadata Principles and Practicalities, *D-Lib Magazine*, April 2002. Retrieved July 1, 2005, from http://www.dlib.org/dlib/april02/weibel/04weibel.html

Friesen, N., Mason, J. & Ward, N. (2002), Building Educational Metadata Application Profiles, *Dublin Core - 2002 Proceedings: Metadata for e-Communities*: *Supporting Diversity and Convergence*. Retrieved July 1, 2005, from http://www.bncf.net/dc2002/program/ft/paper7.pdf

Godby, C. J., Smith, D. & Childress, E. (2003), Two Paths to Interoperable Metadata, *Proceedings of DC-2003: Supporting Communities of Discourse and Practice – Metadata Research & Applications,* Seattle, Washington (USA). Retrieved July 1, 2005, from http://www.siderean.com/dc2003/103_paper-22.pdf

Guidelines for machine-processable representation of Dublin Core Application Profiles (2005), CEN Workshop Agreement CWA 15248. Retrieved July 1, 2005, from ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/MMI-DC/cwa15248-00-2005-Apr.pdf

Heery, R. & Patel, M. (2000), Application Profiles: mixing and matching metadata schemas, *Ariadne Issue 25*, September 2000. Retrieved July 1, 2005, from http://www.ariadne.ac.uk/issue25/app-profiles/

Heflin, J. (2004), OWL Web Ontology Language – Use Cases and Requirements, *W3C Recommendation 10 February 2004*. Retrieved July 1, 2005, from http://www.w3.org/TR/webont-req/

IMS Global Learning Consortium (2004), IMS Meta-data Best Practice Guide for IEEE 1484.12.1-2002 Standard for Learning Object Metadata. Retrieved July 1, 2005, from http://www.imsglobal.org/metadata/mdv1p3pd/imsmd_bestv1p3pd.html

Johnston, P., (2005a), XML, RDF, and DCAPs. Retrieved July 1, 2005, from http://www.ukoln.ac.uk/metadata/dcmi/dc-elem-prop/

Johnston, P., (2005b), Element Refinement in Dublin Core Metadata. Retrieved July 1, 2005, from http://dublincore.org/documents/dc-elem-refine/

Joint DCMI/IEEE LTSC Task Force. http://dublincore.org/educationwiki/DCMIIEEELTSCTaskforce

Klyne, G. & Carroll, J. J. (2004), Resource Description Framework (RDF): Concepts and Abstract Syntax, *W3C Recommendation 10 February 2004*. Retrieved July 1, 2005, from http://www.w3.org/TR/rdf-concepts/

Lagoze, C. (1996), The Warwick Framework – A Container Architecture for Diverse Sets of Metadata, D-Lib Magazine, July/August 1996. Retrieved April 30, 2006 from http://www.dlib.org/dlib/july96/lagoze/07lagoze.html

Manola, F. & Miller, E. (2004), RDF Primer, *W3C Recommendation 10 February 2004*. Retrieved July 1, 2005, from http://www.w3.org/TR/rdf-primer/

Miles, A. J. & Brickley, D. (2005), SKOS Core Guide, *W3C Working Draft 10 May 2005*. Retrieved July 1, 2005, from http://www.w3.org/TR/swbp-skos-core-guide

Memorandum of Understanding between the Dublin Core Metadata Initiative and the IEEE Learning Technology Standards Committee (2000). Retrieved July 1, 2005, from http://dublincore.org/documents/2000/12/06/dcmi-ieee-mou/

Nack, F., van Ossenbruggen, J. & Hardman, L. (2005), That obscure object of desire: multimedia metadata on the Web, part 2, *IEEE Multimedia 12* (1) 54-63. Retrieved July 1, 2005, from http://ieeexplore.ieee.org/iel5/93/30053/01377102.pdf?arnumber=1377102

Nilsson, M. (2001), The Semantic Web: How RDF will change learning technology standards, Feature article, Centre for Educational Technology Interoperability Standards (CETIS). Retrieved July 1, 2005, from http://www.cetis.ac.uk/content/20010927172953/viewArticle

Nilsson, M., Palmér, M. & Naeve, A. (2002), Semantic Web Metadata for e-Learning - Some Architectural Guidelines, *Proceedings of the 11th World Wide Web Conference (WWW2002)*, Hawaii, USA. Retrieved July 1, 2005, from http://kmr.nada.kth.se/papers/SemanticWeb/p744-nilsson.pdf

Nilsson, M., Palmér, M. & Brase, J. (2003), The LOM RDF Binding - Principles and Implementation, *Proceedings of the Third Annual ARIADNE conference*. Retrieved July 1, 2005, from http://kmr.nada.kth.se/papers/SemanticWeb/LOMRDFBinding-ARIADNE.pdf

Nilsson, M., Johnston, P., Naeve, A., Powell, A. (2006), The Future of Learning Object Metadata Interoperability, in Koohang A. (ed.) *Learning Objects: Standards, Metadata, Repositories, and LCMS*, in press.

Powell, A., Nilsson, M., Naeve, A., Johnston, P. (2005), DCMI Abstract Model, *DCMI Recommendation.* Retrieved July 1, 2005, from http://dublincore.org/documents/abstract-model/

Uschold, M. & Gruninger, M. (2002), Creating Semantically Integrated Communities on the World Wide Web, Invited Talk, Semantic Web Workshop, Co-located with WWW 2002, Honolulu, HI, May 7 2002. Retrieved July 1, 2005, from http://semanticweb2002.aifb.uni-karlsruhe.de/USCHOLD-Hawaii-InvitedTalk2002.pdf

van Ossenbruggen, J., Nack, F. & Hardman, L. (2004), That obscure object of desire: multimedia metadata on the Web, part 1, *IEEE Multimedia 11* (4) 38-48. Retrieved July 1, 2005, from http://ieeexplore.ieee.org/iel5/93/29587/01343828.pdf?arnumber=1343828

ARIADNE

Issue 50 January 2007

Main Articles

A Dublin Core Application Profile for Scholarly Works

Julie Allinson, **Pete Johnston** and **Andy Powell** describe a Dublin Core application profile for describing scholarly works that makes use of FRBR and the DCMI Abstract Model.

Main Contents | Section Menu | Email Ariadne | Search Ariadne

Introduction

In May 2006, the Joint Information Systems Committee (JISC) [1] approached UKOLN [2] and the Eduserv Foundation [3] to collaborate on the development of a metadata specification for describing eprints (alternatively referred to as scholarly works, research papers or scholarly research texts) [4]. A Dublin Core (DC) [5] application profile was chosen as the basis of the specification given the widespread use of DC in existing repositories, the flexibility and extensibility of the DCMI Abstract Model [6] and its compatibility with the Semantic Web [7]. The main driver for this work was the establishment of a three-year project to aggregate content from repositories and offer cross-searching and other added-value services [8]. Drawing on the conclusions of the ePrints-UK Project [9] and the findings of the ongoing PerX Project [10], JISC was quick to identify that the quality and consistency of metadata would be a critical success factor for this project.

The work was carried out over a three-month period from May to July 2006, before moving in early August into what we have termed our community acceptance period. A working group of invited experts was assembled to contribute to the development of the application profile both in person and through an active email discussion list and the project deliverables were developed in the open, collaborative arena of the UKOLN Repositories Research Team wiki [11]. The core deliverables were a functional requirements specification, the application model, application profile and usage guidelines, eprints XML schema and 'dumb-down' guidelines. The ensuing article offers a whistle-stop tour of the development process that led to the production of these deliverables and the application profile as a whole.

Identifying Metadata Requirements for Describing Scholarly Works

Identifying the functionality that we need to support is an important first step if the profile is going to be fit for its primary purpose. Current practice for repositories is to expose simple DC records over OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) [12] as mandated by that protocol. However, it is widely agreed that simple DC has limitations that pose problems for repository developers and aggregator services. Issues relating to normalised names, use of controlled subject vocabularies or other authority lists, dates and identifiers are common and many were identified in the course of our functional requirements gathering.

From the work specification supplied by JISC, we defined our primary use case in developing an application profile for scholarly works as: supporting the Intute repository search project to aggregate richer, more consistent, metadata from repositories. Through liaison with that project, a review of existing standards and previous

project findings, plus consultation with our working group, we established a set of scenarios from which we derived an extensive list of functional requirements [13].

Principal amongst these were the following:

- Provision of richer, more consistent metadata.
- Facilitate search, browse or filter by a range of elements, including journal, conference or publication title, peer-review status and resource type.
- Enable identification of the latest, or most appropriate, version and facilitate navigation between different versions.
- Support added-value services, particularly those based on the use of OpenURL ContextObjects [14].
- Implement an unambiguous method of identifying the full text(s).
- Enable identification of the research funder and project code.
- Facilitate identification of open access materials.

Note that by 'open access' we mean "*free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself*." [15]

The Application Model

In order to build up a DC application profile for scholarly publications we first need to develop an application model. This model shows the entities that we want to describe in DC and the key relationships between those entities. It is critical to undertake this modelling step in the development of any application profile. Without it, users of the application profile may become confused about which entity is being described by any given metadata property.

As a simple example, imagine developing an application profile to describe a personal audio CD collection. One might choose to model the following set of entities: the *collection* and its *owner*, each CD, the *recording artist* and the *record label*. Each of these entities could then be described separately, using a specific set of properties for each entity. We refer to such a model as the application model.

The application model for scholarly publications presented here [16] is based on the Functional Requirements for Bibliographic Records (FRBR) [17. FRBR is an entity-relationship model developed by the library community for the entities that bibliographic records are intended to describe. FRBR models the bibliographic world using four key entities: *work, expression, manifestation* and *item*. This article does not attempt to summarise the FRBR model in any detail. Readers that are not familiar with it are encouraged to consult the FRBR documentation [17].

In the context of this model an eprint is defined to be a *scientific* or *scholarlyresearch text* (as defined by the Budapest Open Access Initiative [18]), for example a peer-reviewed journal article, a preprint, a working paper, a thesis, a book chapter, a report, etc.

The Model

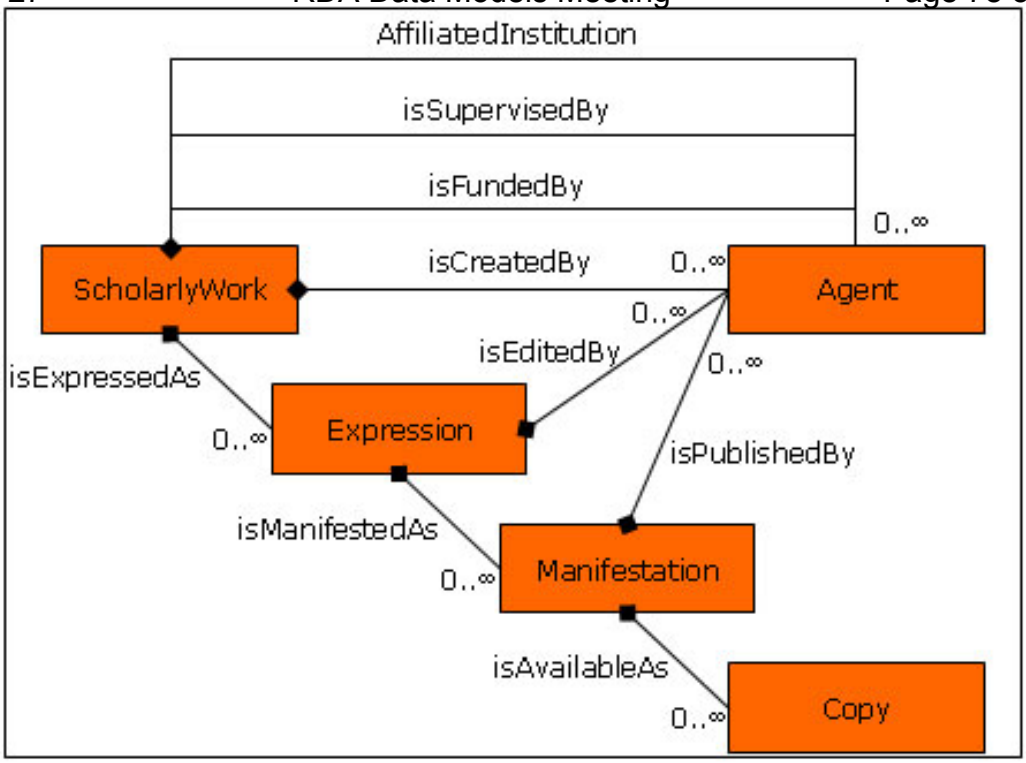


Figure 1 shows the Eprints application model

Although FRBR is used as the basis of the model, some of the entity and relationship labels used in FRBR have been modified for this model, in order to make them more intuitive to those dealing with eprints and to align them with the terminology used in DC:

DC	**FRBR**
ScholarlyWork	Work
Copy	Item
Agent	Corporate Body
isExpressedAs relationship	'is realized through'
isManifestedAs relationship	'is embodied in'
isAvailableAs relationship	'is exemplified by'
isCreatedBy relationship	'is created by'
isPublishedBy relationship	'publisher' attribute of a Manifestation

A ScholarlyWork is a distinct intellectual or artistic scholarly creation.

The isExpressedAs, isManifestedAs and isAvailableAs relationships can be thought of as 'vertical' relations between the ScholarlyWork and its Expressions, between an Expression and its Manifestation and between a Manifestation and its Copies. There are also 'horizontal' relationships between different Expressions of the same ScholarlyWork (e.g. the 'has a translation' relationship in FRBR), different Manifestations of the same Expression (e.g. the 'has an alternative' relationship in FRBR) and so on. These 'horizontal' relationships have not been included in this model. Software applications may be able to infer some of these 'horizontal' relationships by navigating up and down the 'vertical' relationships.

In natural language, what the above model says is:

A ScholarlyWork may be expressed as one or more Expressions. Each Expression may be manifested as one or more Manifestations. Each Manifestation may be made available as one or more Copies. Each ScholarlyWork may have one or more creators, funders and supervisors. Each Expression may be have one or more editors. Each Manifestation may have one or more publishers.

The most common forms of Expression of an eprint are the various 'revisions' that it goes through (draft, pre-print, ..., final published version, etc.) and its different translations. Therefore, the most important Expression to Expression relationships required are isVersionOf/hasVersion and isTranslationOf/hasTranslation.

Attributes

A critical part of developing the application model is to identify the key attributes that will be used to describe each entity in the model. Initially, this can be done in a fairly generic way, noting for example that we want to capture the 'title' of the ScholarlyWork but not worrying about whether we are going to use DC Title or some other kind of title. The key attributes for each of the entities in our application profile are listed below.

Attributes of a ScholarlyWork

- title
- subject
- abstract
- grant number
- has adaptation
- identifier (URI)

Attributes of an Expression

- title
- description
- date available
- status
- version number or string
- language
- genre / type
- copyright holder
- has version
- has translation
- bibliographic citation
- references
- identifier (URI)

Attributes of a Manifestation

- format
- date modified
- identifier (URI)

Attributes of a Copy

- date available
- access rights
- licence
- is part of
- identifier/locator (URI)

Attributes of an Agent

- name
- family name
- given name
- type of agent
- workplace homepage
- mailbox
- homepage
- identifier (URI)

A Note on Implementing This Model Using DC Metadata

Many of the above relationships and attributes can be implemented fairly easily using metadata terms already defined by the Dublin Core Metadata Initiative (DCMI) [5]. DC metadata is sometimes only considered capable of describing flat, single-entity, constructs - a Web page, a document, an image, etc. However, the DCMI Abstract Model [6] introduces the notion of a *description* set, a group of related *descriptions*, which allows it to be used to capture metadata about more complex sets of entities, using application models like the one described here.

DCMI is currently developing a revised set of encoding guidelines for XML and RDF/XML, which will allow these more complex, multi-description, *description set* constructs to be encoded and shared between software applications.

The Application Profile and Vocabularies

The application profile provides a way of describing the attributes and relationships of each of the five entities as part of a *description set*. The profile also identifies mandatory elements, provides usage guidelines and offers illustrative examples. Note that for this application profile, we have made very few elements mandatory. Indeed, all that a minimal description set *must* include is **either**:

- a single ScholarlyWork *description* with at least one dc:title *statement* and one dc:type *statement* indicating that this is a ScholarlyWork entity,

or:

- a single ScholarlyWork *description* with one dc:type *statement* indicating that this is a ScholarlyWork and one eprints:isExpressedAs *statement* linking to a single Expression *description* with at least one dc:title *statement* and one dc:type *statement* indicating that this is an Expression.

All other aspects of the application profile are optional.

It is not the intention of this article to offer a full analysis of the different metadata properties and readers should refer to the documentation for further information [4]. Briefly, the profile makes use of properties from a number of schemes: the DC Metadata Element Set (simple DC) [19], DC Metadata Terms (includes qualified DC terms) [20] and the MARC relator codes [21] all provide terms. Properties from the Friends of a Friend (FOAF) Scheme [22] introduce some semantic web flavour and only five new properties have been created from scratch: grant number, affiliated institution, status, version and copyright holder.

Where existing dc:relation qualifiers have been used, the relationships being documented have been clearly defined alongside five new properties:

- has adaptation
- has translation
- is expressed as
- is manifested as
- is available as

To aid fulfilment of several of the functional requirements further, four vocabularies have been defined for:

- access rights (Open, Restricted or Closed)
- entity type (ScholarlyWork, Expression, Manifestation, Copy or Agent)
- status (Peer Reviewed or Non Peer Reviewed)
- resource type

Figure 2 shows the resource type vocabulary as an extension of the value 'Text' in the DCMI Type scheme [23].

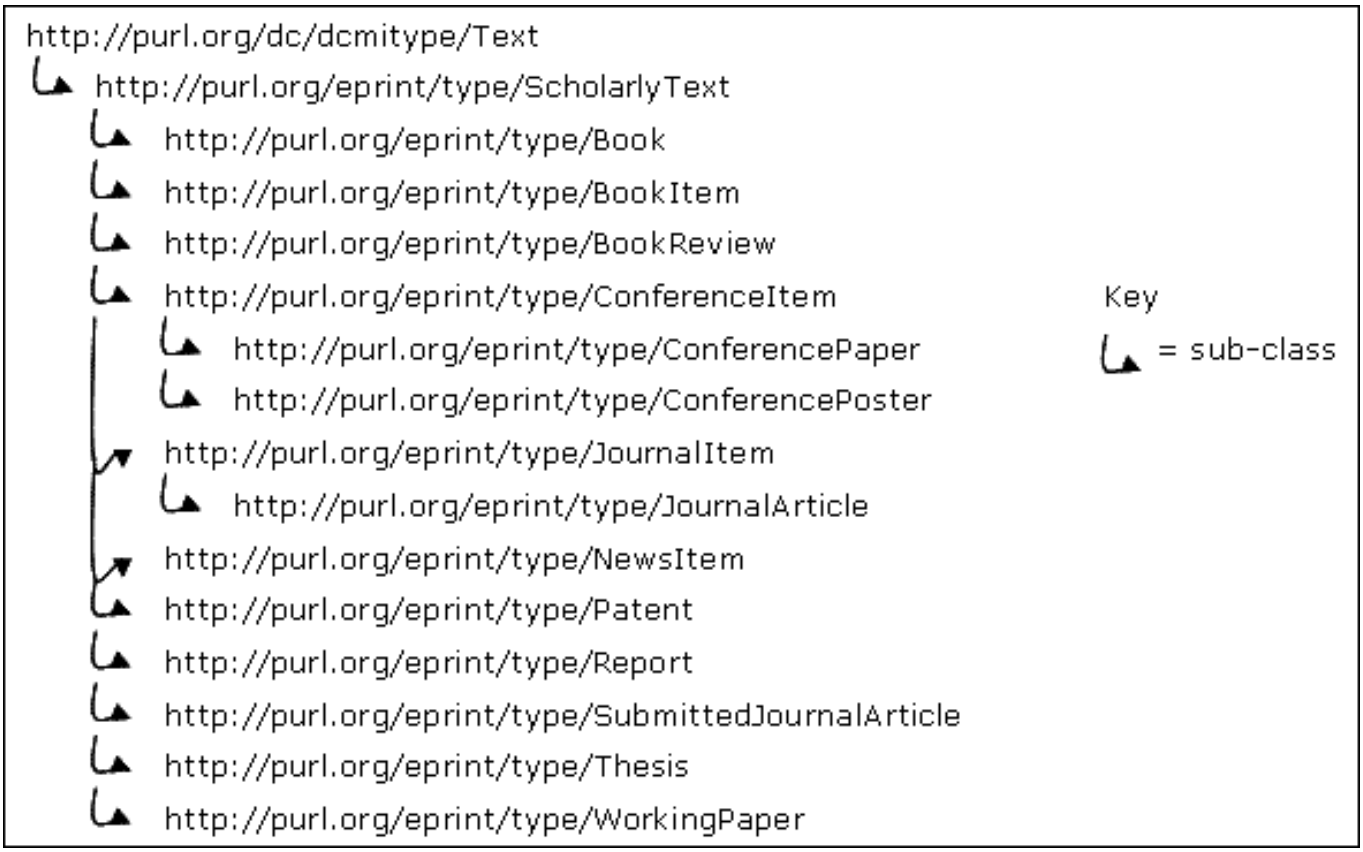


Figure 2: Eprints Type Vocabulary

Eprints DC XML: An XML Format for the Eprints Application Profile

At the time of writing (January 2007), the DCMI does not define an XML format to support the serialisation of DC description sets as described by the DCMI Abstract Model (DCAM). The existing DCMI recommendation *Guidelines for implementing Dublin Core in XML* [24] pre-dated the development of the DCAM and is based on two simpler 'abstract models' for DC metadata which are described in that specification itself. The DCMI Architecture Community [25] is currently considering a working draft [26] that describes a new XML format which is based on the DCAM, with the intention of producing a new DCMI recommendation, probably in early/mid-2007.

Since the Eprints application profile makes use of the full range of features of the DCAM, the serialisation of description sets based on that application profile requires a format which supports those features, so the working group has defined an XML format known as Eprints DC XML [27]. The format is based very closely on the latest drafts being considered by the DCMI Architecture Community, although it does make use of different XML Namespace Names from those used in the DCMI drafts.

Figure three shows an example instance of the Eprints DC XML format:

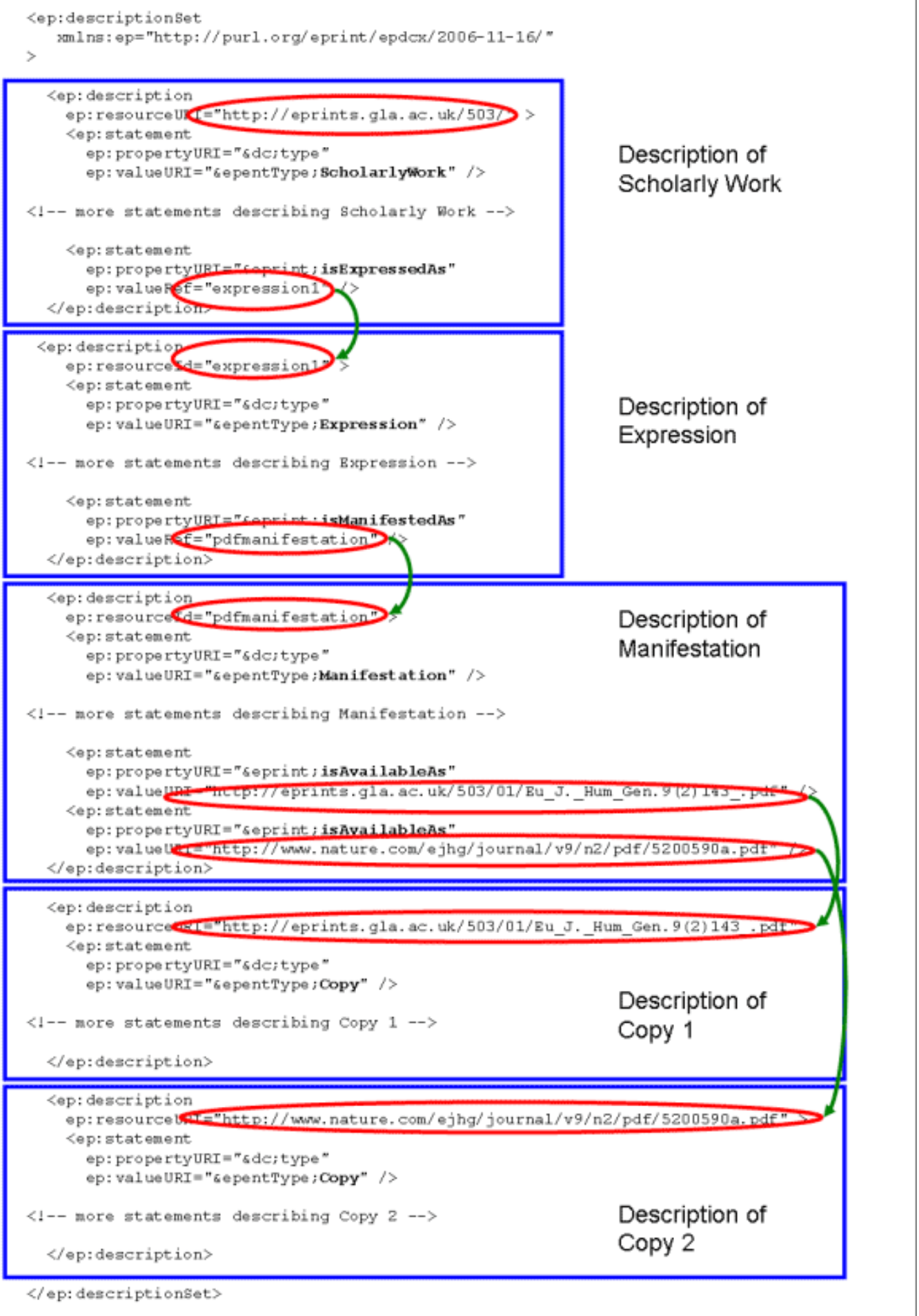

Description of Scholarly Work
Description of Expression
Description of Manifestation
Description of Copy 1
Description of Copy 2

Figure 3: Eprints DC XML instance

A W3C XML Schema and a RELAX NG [28] Schema for Eprints DC XML are available.

The Eprints Application Profile and 'dumb-down'

One of the requirements of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) specification [12] is that, for each "item" in a repository, the repository must support the dissemination of metadata records in the "oai_dc" "metadata format". "oai_dc" is a format defined by the OAI-PMH specification to serialise "Simple DC" description sets. Simple DC is an application profile in which:

- the *description set* comprises a single *description*
- each *statement* within that *description* references one of the 15 *properties* of the Dublin Core Metadata Element Set [19]
- each of those 15 *properties* may be referenced in multiple *statements*
- each *statement* has a single *value string*
- each *value string* may have an associated *language tag*
- there is no use of *resource URIs, vocabulary encoding scheme URIs, syntax encoding scheme URIs, value URIs* or *rich representations*

The Simple DC profile is used by many systems as a 'lowest common denominator' for basic interoperability, and the process of transforming description sets based on some richer application profile into description sets based on the Simple DC profile is sometimes referred to as 'dumb-down', reflecting the fact that such a transformation involves a loss of information content. The working group provided a mapping from the Eprints application profile to the Simple DC application profile [29]. Because a description set based on the Eprints application profile typically contains multiple descriptions, each of a single resource, the mapping generates multiple Simple DC description sets from a single Eprints application profile description set.

In the proposed mapping, the resulting Simple DC description sets describe the eprint only at the ScholarlyWork and Copy levels. The Simple DC description set for the ScholarlyWork complies with the guidelines specified by the ePrints-UK Project [30]. This is not the only possible approach to mapping the Eprints application profile to Simple DC. For example, it would also be possible to map to a group of Simple DC description sets, one for each entity in the model or to a single Simple DC description set only about the ScholarlyWork. However, the working group felt that the chosen mapping offered the most useful set of simple DC descriptions with minimal loss of information.

Conclusion: Towards Community Acceptance

This application profile represents a relatively innovative approach to metadata, taking as it does the FRBR model and applying it to scholarly works. By making use of the benefits afforded by the DCMI Abstract Model, the profile is able to group descriptions of multiple entities into a single description set. Overall this approach is guided by the functional requirements identified above and the primary use case of richer, more functional, metadata. It also makes it easier to rationalise 'traditional' citations between 'expressions' and 'modern' hypertext links between 'copies', as well as supporting navigation between different versions and the identification of appropriate, and, we hope, open access, full-text copies. In practice, this seemingly complex model may be manifest in relatively simple metadata and/or end-user interfaces. Furthermore, it is likely that many repositories already capture the metadata properties identified in the profile, but are prevented from usefully exposing this metadata to other services by the limitations imposed by simple DC.

Yet the application profile alone cannot bring about interoperability or provide Intute and other aggregators with the metadata necessary to offer rich functionality. For this we need community uptake by repositories and repository software developers, agreement on common approaches and most of all, Eprints DC XML metadata being generated and exposed. There are growing signs that our community acceptance and dissemination activities to date are generating momentum, with support built into the newly released GNU Eprints version 3, alongside statements of support from DSpace and Fedora [31] developers, interest from European colleagues and lively discussions at the recent DC - 2006 and Open Scholarship 2006 conferences in Mexico and Glasgow.

References

1. The Joint Information Systems Committee (JISC) http://www.jisc.ac.uk/
2. UKOLN, University of Bath http://www.ukoln.ac.uk/
3. The Eduserv Foundation http://www.eduserv.org.uk/foundation/
4. Eprints Application Profile http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Application_Profile
5. The Dublin Core Metadata Initiative (DCMI) http://dublincore.org/
6. Powell, Andy, Nilsson, Mikael, Naeve, Ambjörn and Johnston, Pete, DCMI Abstract Model. DCMI Recommendation, May 2005 http://dublincore.org/documents/abstract-model/
7. W3C Semantic Web Activity http://www.w3.org/2001/sw/
8. Intute Repository Search Project http://www.intute.ac.uk/projects.html
9. Eprints UK Project http://eprints-uk.rdn.ac.uk/project/
10. PerX Project http://www.icbl.hw.ac.uk/perx/
11. Repositories Research Team wiki http://www.ukoln.ac.uk/repositories/digirep/
12. Lagoze, Carl, Van de Sompel, Herbert, Nelson, Michael and Warner, Simeon. *The Open Archives Initiative Protocol for Metadata Harvesting*. Protocol Version 2.0 of 2002-06-14. http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm
13. Eprints Application Profile Functional requirements specification http://www.ukoln.ac.uk/repositories/digirep/index/Functional_Requirements
14. ANSI/NISO Z39.88 -2004: The OpenURL Framework for Context-Sensitive Services http://www.niso.org/standards/standard_detail.cfm?std_id=783
15. Budapest Open Access Initiative: Frequently Asked Questions http://www.earlham.edu/~peters/fos/boaifaq.htm#openaccess
16. Eprints Application Model http://www.ukoln.ac.uk/repositories/digirep/index/Model
17. IFLA, *Functional Requirements for Bibliographic Records*, 1998 http://www.ifla.org/VII/s13/frbr/frbr.pdf
18. Budapest Open Access Initiative http://www.soros.org/openaccess/
19. *Dublin Core Metadata Element Set*, Version 1.1. DCMI Recommendation. April 2006. http://dublincore.org/documents/dces/
20. DCMI Usage Board, *DCMI Metadata Terms*, DCMI Recommendation, December 2006 http://dublincore.org/documents/dcmi-terms/
21. Library of Congress Network Development and MARC Standards Office, MARC Code Lists for Relators, Sources, Description Conventions, January 2007 http://www.loc.gov/marc/relators/
22. Brickley, Dan and Miller, Libby, *FOAF Vocabulary Specification*, January 2006 http://xmlns.com/foaf/0.1/
23. DCMI Usage Board, *DCMI Type Vocabulary*, DCMI Recommendation, August 2006 http://dublincore.org/documents/dcmi-type-vocabulary/
24. Powell, Andy and Johnston, Pete. *Guidelines for implementing Dublin Core in XML*. DCMI Recommendation. April 2003. http://dublincore.org/documents/dc-xml-guidelines/
25. DCMI Architecture Community. http://dublincore.org/groups/architecture/
26. Johnston, Pete and Powell, Andy. *Expressing Dublin Core metadata using XML*. DCMI Working Draft. May 2006. http://dublincore.org/documents/2006/05/29/dc-xml/
27. Johnston, Pete. *Eprints DC XML*. November 2006. http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_DC_XML
28. OASIS RELAX NG Technical Committee http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=relax-ng
29. Powell, Andy. *Mapping the Eprints Application Profile to Simple DC*. August 2006. http://www.ukoln.ac.uk/repositories/digirep/index/Mapping_the_Eprints_Application_Profile_to_Simple_DC
30. Powell, Andy, Day, Michael and Cliff, Peter. *Using simple Dublin Core to describe eprints*. Version 1.2. http://eprints-uk.rdn.ac.uk/project/docs/simpledc-guidelines/
31. Eprints Application Profile Community Acceptance Plan http://www.ukoln.ac.uk/repositories/digirep/index/Community_Acceptance_Plan

Author Details

Julie Allinson

Repositories Research Officer

UKOLN

University of Bath

Email: j.allinson@ukoln.ac.uk

Web site: http://www.ukoln.ac.uk/

Pete Johnston

Technical Researcher

Eduserv Foundation

Email: pete.johnston@eduserv.org.uk

Web site: http://www.eduserv.org.uk/foundation/

Andy Powell

Head of Development

Eduserv Foundation

Email: andy.powell@eduserv.org.uk

Web site: http://www.eduserv.org.uk/foundation/

Article Title: "A Dublin Core Application Profile for Scholarly Works"

Author: Julie Allinson, Pete Johnston and Andy Powell

Publication Date: 30-January-2007 Publication: Ariadne Issue 50

Originating URL: http://www.ariadne.ac.uk/issue50/allinson-et-al/

Quick Guide to Publishing a Thesaurus on the Semantic Web

W3C Working Draft 17 May 2005

This version:
http://www.w3.org/TR/2005/WD-swbp-thesaurus-pubguide-20050517
Latest version:
http://www.w3.org/TR/swbp-thesaurus-pubguide
Previous version:
http://www.w3.org/TR/2005/WD-swbp-thesaurus-pubguide-20050510
Editors:
Alistair Miles, CCLRC

Abstract

This document describes in brief how to express the content and structure of a thesaurus, and metadata about a thesaurus, in RDF. Using RDF allows data to be linked to and/or merged with other RDF data by semantic web applications. The Semantic Web, which is based on the Resource Description Framework (RDF), provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries.

Status of this Document

This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the latest revision of this technical report can be found in the W3C technical reports index at http://www.w3.org/TR/.

This document is an updated First Public Working Draft published by the Semantic Web Best Practices and Deployment Working Group, part of the W3C Semantic Web Activity. The Working Group intends the Quick Guide to Publishing a Thesaurus on the Semantic Web to become a W3C Working Group Note.

This Quick Guide accompanies the SKOS Core Vocabulary Specification and SKOS Core Guide.

We encourage public comments. Please send comments to public-esw-thes@w3.org [archive] and start the subject line of the message with "comment:".

Publication as a Working Draft does not imply endorsement by the W3C Membership. This is a draft document and may be updated, replaced or obsoleted by other documents at any time. It is inappropriate to cite this document as other than work in progress.

Change History: The initial (previous) version of this document was published with incorrect URIs in the UKAT examples. The URIs used in the previous version resolved to textual descriptions of UKAT concepts and were not intended to be the names of the concepts themselves. This last-minute publication change is corrected in this version.

Contents

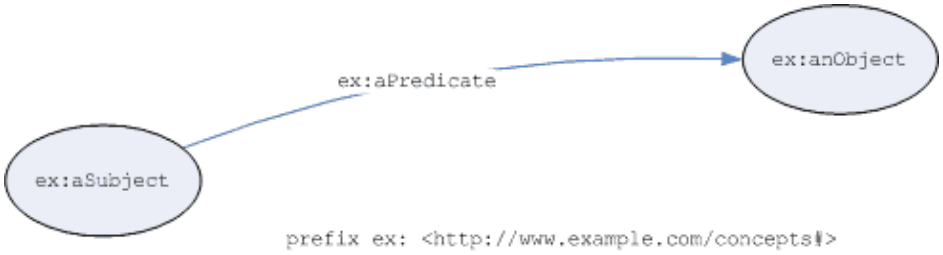
Introduction

The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries [Semantic Web Activity]. It is based on the Resource Description Framework (RDF) [RDF], which provides a simple data formalism for talking about things, their properties, inter-relationships, and categories (classes). For an overview of RDF, see [RDF Concepts]

This document describes in brief how to express the content and structure of a thesaurus, and metadata about a thesaurus, in RDF. Using RDF allows your data to be linked to and/or merged with other RDF data by semantic web applications. In practice, this means that data sources can be distributed across the web in a decentralised way, but still be meaningfully composed and integrated by applications, often in novel and unanticipated ways.

The examples in this guide are given as a visualisation of the RDF graph, e.g.

example



An RDF graph can be serialised (i.e. encoded as a series of characters) according to any of three currently defined RDF syntaxes: RDF/XML [RDF Syntax], N3/Turtle [Turtle][N3], N-Triple [N-Triple]. All examples in this guide use the RDF/XML syntax for consistency.

The examples in this document use the SKOS Core Vocabulary, which is a set of properties and classes that can be used to express the conceptual content of a thesaurus as an RDF graph. SKOS Core is designed to be used with not only thesauri, but also other types of 'concept scheme', such as classification schemes, subject heading systems, controlled vocabularies, glossaries, taxonomies etc. For a complete description of SKOS Core, see [SKOS Core Guide].

The examples in this document also use the DCMI Metadata Terms, which are properties and classes for describing resource metadata. For more about DCMI Terms, see [DCMI Terms].

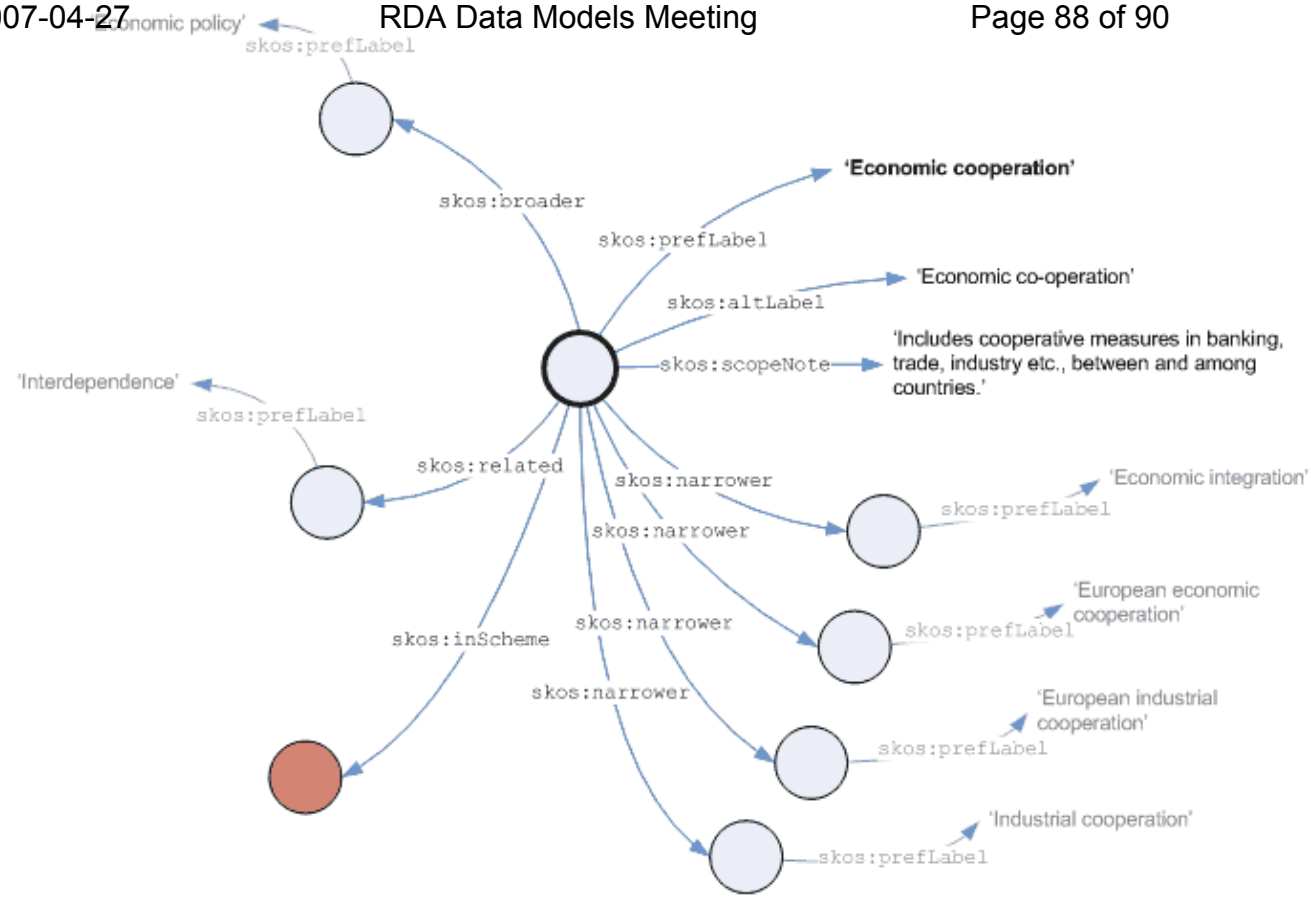
Note that the prefix `skos:` in this document stands for `http://www.w3.org/2004/02/skos/core#` - so for example `skos:prefLabel` is an abbreviation of `http://www.w3.org/2004/02/skos/core#prefLabel`.

Expressing a Thesaurus in RDF

Below is an extract from the UK Archival Thesaurus (UKAT) [UKAT]:

```
Term: Economic cooperation

Used For:
    Economic co-operation

Broader terms:
    Economic policy

Narrower terms:
    Economic integration
    European economic cooperation
    European industrial cooperation
    Industrial cooperation

Related terms:
    Interdependence

Scope Note:
Includes cooperative measures in banking, trade, industry etc., between and among countries.
```

This example, expressed as an RDF graph using the SKOS Core Vocabulary, looks like:



Each of the blue circles in the image above represents a concept from the UKAT. The red circle represents the UKAT itself.

Each concept from the UKAT has an allocated URI. URIs are globally unique identifiers that may be used to refer to a resource unambiguously from any context. Anything can be a 'resource', not just web documents, therefore URIs can be used as identifiers for anything. For example, the URI:

`http://www.ukat.org.uk/thesaurus/concept/1750`

... denotes the concept from the UKAT whose preferred term is 'Economic cooperation'. (The URIs are not shown in the visualisation of the graph above for the sake of readability.)

Allocating URIs to the concepts in a thesaurus allows anybody to refer to them unambiguously from any context.

For a complete description of considerations relevant to allocating URIs, see [WEBARCH]. See also the section 'HTTP URIs for Concepts' in [SKOS Core Guide].

An RDF/XML serialisation of the RDF description of the 'Economic cooperation' concept from the UKAT is below:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

    <skos:Concept rdf:about="http://www.ukat.org.uk/thesaurus/concept/1750">
        <skos:prefLabel>Economic cooperation</skos:prefLabel>
        <skos:altLabel>Economic co-operation</skos:altLabel>
        <skos:scopeNote>Includes cooperative measures in banking, trade, industry etc.,
          between and among countries.</skos:scopeNote>
        <skos:broader rdf:resource="http://www.ukat.org.uk/thesaurus/concept/4382"/>
        <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/2108"/>
        <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/9505"/>
        <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/15053"/>
        <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/18987"/>
        <skos:related rdf:resource="http://www.ukat.org.uk/thesaurus/concept/3250"/>
        <skos:inScheme rdf:resource="http://www.ukat.org.uk/thesaurus"/>
    </skos:Concept>

</rdf:RDF>
```

Note that, in expressing the content of a thesaurus such as the UKAT in RDF using SKOS Core, each descriptor (preferred term) becomes a *preferred label for a concept*, and each non-descriptor (non-preferred term) becomes an *alternative label for a concept*.

Note also the `skos:inScheme` property - this property may be used to assert a link between a concept and the concept scheme(s) in which it participates. SKOS Core places no restriction on the number of schemes in which a concept may participate.

Creating and Maintaining an RDF Representation of a Thesaurus

Most thesauri are managed via a thesaurus management system. Where the thesaurus management system stores its data in a relational database, or in an XML or structured text file format, or where the standard output of the thesaurus management system is an XML or structured text format, it is usually possible to create an RDF representation of the thesaurus via an automated procedure (e.g. database report, text parsing program, XSLT transformation).

A full discussion of conversion techniques and best practice is beyond the scope of this document. However, note that when using an automated procedure, care must be taken to ensure that the output generated is sensible and conforms to the recommended usage of the SKOS Core Vocabulary. For example, if an XML format contains an XML element whose name is 'scopenote' it should not be automatically assumed that the textual content of these elements should be converted to the value of a `skos:scopeNote` property. Perhaps these 'scopenote' elements actually contain definitions, in which case the `skos:definition` property should be used; or perhaps the 'scopenote' elements have been used very loosely and contain all kinds of note types, in which case the more general `skos:publicNote` property would be more appropriate.

Also note that maintaining an RDF representation of a thesaurus requires clear policies for versioning and change management. For example, users need to know if the meaning associated with a URI is stable, and if it isn't they need to know how and when it may change. A fuller discussion of management best practice is currently being undertaken by the Vocabulary Management Task Force of the Semantic Web Best Practices and Deployment Working Group.

Expressing Thesaurus Metadata in RDF

RDF can also be used to express metaproperties of a thesaurus, such as it's title, description, date of modification and so on. The DCMI Metadata Terms [DCMI Terms] include a number of useful properties for this purpose. For example, below is an RDF/XML serialisation of the UKAT metadata:

```
<rdf:RDF
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:skos="http://www.w3.org/2004/02/skos/core#"
    xmlns:dc="http://purl.org/dc/elements/1.1/">

    <skos:ConceptScheme rdf:about="http://www.ukat.org.uk/thesaurus">
        <dc:title>The UK Archival Thesaurus</dc:title>
        <dc:description>A subject thesaurus produced to support indexing in the UK
          archive sector.</dc:description>
        <dc:creator>UK Archival Thesaurus project</dc:creator>
        <dc:date>2004-08-22</dc:date>
        <dc:language>en</dc:language>
        <dc:rights>All rights reserved. Data in the UK Archival Thesaurus may be freely
          used and copied, without prior permission, for educational and other non-commercial
          purposes. These purposes include (but are not limited to) the incorporation of UKAT
          data into indexes, thesauri and finding aids created by organisations and projects
          in the archive sector and the wider heritage sector, in the UK and elsewhere. Under
          no circumstances may copies of UKAT data be sold without prior written permission
          from the UKAT Project (support@ukat.org.uk).</dc:rights>
        <skos:hasTopConcept rdf:resource="http://www.ukat.org.uk/thesaurus/field/1"/>
        <skos:hasTopConcept rdf:resource="http://www.ukat.org.uk/thesaurus/field/2"/>
        <skos:hasTopConcept rdf:resource="http://www.ukat.org.uk/thesaurus/field/3"/>
        <skos:hasTopConcept rdf:resource="http://www.ukat.org.uk/thesaurus/field/4"/>
        <skos:hasTopConcept rdf:resource="http://www.ukat.org.uk/thesaurus/field/5"/>
        <skos:hasTopConcept rdf:resource="http://www.ukat.org.uk/thesaurus/field/6"/>
        <skos:hasTopConcept rdf:resource="http://www.ukat.org.uk/thesaurus/field/8"/>
    </skos:ConceptScheme>

</rdf:RDF>
```

See [DCMI Terms] for a description of the recommended usage of these properties.

To make statements about a thesaurus in RDF, the thesaurus must have an allocated URI. For example, the URI:

`http://www.ukat.org.uk/thesaurus`

... denotes the UKAT itself.

Note that SKOS Core models a thesaurus as a 'concept scheme'. For more about this, see the section 'Concept Schemes' in [SKOS Core Guide].

Note also that, in the example above, a link has been asserted between the UKAT thesaurus and the top concepts in the UKAT thesaurus (in the UKAT they are known as 'fields') using the `skos:hasTopConcept` property. Using this property is recommended, as it gives applications an efficient way of locating the top concepts for a given scheme.

Publishing RDF Data

The simplest way to publish RDF data is to create one or more RDF documents containing your data, and publish them on the web via a normal HTTP server.

Note that, although the examples above all use the RDF/XML serialisation syntax (i.e. file format), there are two other alternative syntaxes for RDF: N3/Turtle [N3][Turtle] and N-Triples [N-Triples]. For documents containg RDF data in the RDF/XML format, the 'content-type' field in the HTTP header for that document should be 'application/rdf+xml'.

You can also publish your RDF data on the web via a dedicated RDF server such as Joseki [Joseki] or Sesame [Sesame]. Publishing via an RDF server allows anyone to query the thesaurus over the web via an RDF query language such as SPARQL [SPARQL].

Current practice suggests that if you use HTTP URIs as identifiers for your vocabulary and its concepts, then those URIs should do double duty to an RDF description of the thing they identify.

Further Reading

- SKOS Core Guide
- RDF Primer
- Architecture of the World Wide Web, Volume One

References

[DCMI Terms]
DCMI Metadata Terms. Dublin Core Metadata Initiative, 2004. (See http://dublincore.org/documents/dcmi-terms/)

[Joseki]
Joseki Jena RDF Server. Sourceforge. (See http://www.joseki.org/)

[N3]
Tim Berners-Lee. *Primer: Getting into RDF & Semantic Web using N3*. World Wide Web Consortium, 2004. (See http://www.w3.org/2000/10/swap/Primer)

[N-Triples]
Jan Grant, Dave Beckett, editors. *RDF Test Cases (Section 3. N-Triples)*. World Wide Web Consortium, 2004. (See http://www.w3.org/TR/rdf-testcases/#ntriples)

[RDF]
Resource Description Framework (RDF). (See http://www.w3.org/RDF/)

[RDF Concepts]
Graham Klyne, Jeremy Carroll, editors.*Resource Description Framework (RDF): Concepts and Abstract Syntax*. World Wide Web Consortium, 2004. (See http://www.w3.org/TR/rdf-concepts/)

[RDF Syntax]
Dave Beckett, editor. *RDF/XML Syntax Specification (Revised)*. World Wide Web Consortium, 2004. (See http://www.w3.org/TR/rdf-syntax-grammar/)

[Semantic Web Activity]
Semantic Web Activity Statement. World Wide Web Consortium, 2001. (See http://www.w3.org/2001/sw/Activity)

[Sesame]
Sesame RDF Database. (See http://www.openrdf.org/)

[SKOS Core Guide]
Alistair Miles, Dan Brickley, editors. *SKOS Core Guide*. World Wide Web Consortium, 2004. (See http://www.w3.org/2004/02/skos/core/guide/)

[SPARQL]
Eric Prud'hommeaux, Andy Seaborne, editors. *SPARQL Query Language for RDF*. World Wide Web Consortium, 2004. (See http://www.w3.org/TR/rdf-sparql-query/)

[Turtle]
Dave Beckett. *Turtle - Terse RDF Triple Language*. ILRT, University of Bristol, 2004. (See http://www.ilrt.bris.ac.uk/discovery/2004/01/turtle/)

[UKAT]
The UK Archival Thesaurus. See (http://www.ukat.org.uk/)

[WEBARCH]
Ian Jacobs, Norman Walsh, editors. *Architecture of the World Wide Web, Volume One*. World Wide Web Consortium, 2004. (See http://www.w3.org/TR/webarch/)