Title: Usage Board meeting - agenda
Identifier: /usage/meetings/2006/04/seattle/
Created: 2006-03-30

Seattle, 29-30 April 2006
Venue: /usage/meetings/2006/04/seattle/venue/

Saturday, 29 April 2006 - 9:00 to 12:00

--- Changes to terms in the DCMES namespace (Andrew)
004 /usage/meetings/2006/04/seattle/dcmes-changes/

005 [1] /usageboard/2006/2006-01.definitions/term-changes/
010 [2] /usage/meetings/2006/04/seattle/dcmes-changes/comment-announcement-text.txt
011 [3] /usage/meetings/2006/04/seattle/dcmes-changes/2005-03-23.dc-language.html

--- Changes to terms in the DCTERMS namespace (Diane)
012 /usage/meetings/2006/04/seattle/term-changes/

013 [1] /usageboardwiki/TermChanges?action=print
016 [2] /usage/meetings/2006/04/seattle/term-changes/2006-03-11.educationLevel.html

--- Replicating DCMES terms in the DCTERMS namespace (Andy)
020 /usage/meetings/2006/04/seattle/all-in-dcterms/

021 [1] /usageboardwiki/ReplicatingDCMESINDCTERMS?action=print

Saturday, 29 April 2006 - 13:00 to 17:00

--- DCMI property domains and ranges (Andy) - before dcmitype
023 /usage/meetings/2006/04/seattle/domains-ranges/

025 [1] /architecturewiki/DCPropertyDomainsRanges?action=print
043 [2] /architecturewiki/DCRDFTaskforce/DCRDFExecutiveSummary3?action=print
045 [3] /usage/meetings/2006/04/seattle/domains-ranges/2006-03-24.domain-range-rationale.html
049 [4] /usage/meetings/2006/04/seattle/domains-ranges/2006-03-28.domain-range-comments.html
053 [5] /usage/meetings/2006/04/seattle/domains-ranges/2006-03-24.dcPropertiesRanges.pdf

--- Finalizing DCMI Type Vocabulary (Stuart) - after domains-ranges
063 /usage/meetings/2006/04/seattle/dcmitype/

064 [1] /usage/public-comment/2005/12/type-vocabulary-changes/
068 [2] /usage/public-comment/2006/03/type-vocabulary-comments/

Sunday, 30 April 2006 - 9:00 to 12:00

--- NSDL Registry (Diane)
071 /usage/meetings/2006/04/seattle/nsdl-registry/

--- Review of Application Profiles: Reference documents (Tom)

See separate packet of documentation.

--- Review of Application Profiles - the pipeline
072 /usage/meetings/2006/04/seattle/profile-pipeline/

073 [1] http://homes.ukoln.ac.uk/~lispj/dc-cd/rep200602.html

Sunday, 30 April 2006 - 13:00 to 17:00

--- Article about Dublin Core in Wikipedia (Diane)
075 /usage/meetings/2006/04/seattle/wikipedia/

076 [1] /usage/meetings/2006/04/seattle/wikipedia/dc-in-wikipedia.pdf
079 [2] /usage/meetings/2006/04/seattle/wikipedia/2006-03-21.digest.txt

--- Review of Resource Description and Access (RDA, aka AACR3) (Tom, Diane)
080 /usage/meetings/2006/04/seattle/rda-review/

085 [1] /groups/libraries/rda/
087 [2] /usage/meetings/2006/04/seattle/rda-review/2006-02-06.digest.html
093 [3] /usage/meetings/2006/04/seattle/rda-review/2006-03-28.rda-discussion.html
095 [4] /usage/meetings/2006/04/seattle/rda-review/SCScommentsRDAPart1-excerpts.pdf
108 [5] /usage/meetings/2006/04/seattle/rda-review/RDA_for_who.htm

--- Alternative paths to semantic specificity (Tom)
111 /usage/meetings/2006/04/seattle/semantic-specificity/

113 [1] /usage/meetings/2006/04/seattle/semantic-specificity/2006-03-13.digest.txt

--- DCMI Documentation Roadmap (Tom)
125 /usage/meetings/2006/04/seattle/docn-roadmap/

126 [1] /usage/meetings/2006/04/seattle/docn-roadmap/2006-03-31.roadmap.pdf

--- DCMI Usage Board process (Diane, Stuart)
145 /usage/meetings/2006/04/seattle/ub-process/

146 [1] /usage/documents/process/
151 [2] /usage/meetings/2006/04/seattle/ub-process/2006-02-23.process-ub-voting.txt

--- Backburner issues
153 /usage/meetings/2006/04/seattle/backburner/

154 [1] http://people.opera.com/charlesm/2006/shortdate/
161 [2] http://www.ukoln.ac.uk/metadata/dcmi/date-dccd-odrf/
164 [3] /usage/meetings/2006/04/seattle/backburner/2005-08-10.rebecca-comments.txt
165 [4] /usage/meetings/2006/04/seattle/backburner/2005-08-13.YearMonthDate-profile.txt
167 [5] /usage/meetings/2006/04/seattle/backburner/2005-08-22.douglas-campbell-long.txt
174 [6] /usage/meetings/2006/04/seattle/backburner/2005-05-10.usageboard-profile.txt

---

Date: Mon, 20 Mar 2006 09:34:51 -0800
Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: Stuart Sutton <sasutton@U.WASHINGTON.EDU>
Subject: Seattle UB meeting logistics
Comments: cc: Mike Crandall <mikecran@u.washington.edu>
To: DC-USAGE@JISCMAIL.AC.UK

All, please accept my apologies for being so late getting logistical information to you for the Seattle meeting Saturday and Sunday, April 29-30 (Board meeting on Monday, May 1). Both the Usage Board and Board of Trustees meetings will be held in Mary Gates Hall (MGH) room 310D--one the iSchool's conference rooms on the University of Washington (UW) campus. Here is a link to the campus map with MGH circled: http://www.washington.edu/home/maps/?MGH. We can accommodate direct network access without authentication and wireless with authentication (working on having wireless for you without authentication).

HOTELS:

There are three comparable hotels within easy walking distance (5 minutes) of the UW campus. We will need you to make your own room reservations. I have listed the three below in descending order of expense (althought Watertown and University Tower are priced fairly similarly). The University Inn is reasonably priced and quite satisfactory. Until recently, the Inn was the primary hotel in the area for University guests. The Watertown is the newest and considered one of the best hotels in Seattle. The University Tower was recently renovated. All three have Internet access. All three are approximately 10 minutes from downtown by cab (Watertown has free dowtown shuttle).

University Inn
4140 Roosevelt Way NE=20
Seattle, WA 98105
http://www.universityinnseattle.com/guestrooms.htm
(Free wireless in rooms)

Watertown
4242 Roosevelt Way NE=20
Seattle, WA 98105=20
http://www.watertownseattle.com/guestrooms.htm
(Free internet access)

University Tower Hotel
4507 Brooklyn Ave NE
Seattle, WA 98105
Seattle.com (better rates): http://www.seattle.com/universitytowerhotel/
Direct: http://www.universitytowerhotel.com/
(Free wireless in rooms; DSL in rooms available on request)

DINNER WITH BOARD OF TRUSTEES:

We will be dining with the Board of Trustees the evening of Sunday, April 30.

Stuart A. Sutton, Associate Professor
The Information School Research Commons
University of Washington
4311 11th Ave NE, Suite 400
Seattle, WA 98105
http://www.ischool.washington.edu

Title: Changes to terms in the DCMES namespace
Identifier: /usage/meetings/2006/04/seattle/dcmes-changes/
Created: 2006-03-30

Shepherd: Andrew

Texts:
[1] /usageboard/2006/2006-01.definitions/term-changes/
[2] /usage/meetings/2006/04/seattle/dcmes-changes/comment-announcement-text.txt
[3] /usage/meetings/2006/04/seattle/dcmes-changes/2005-03-23.dc-language.html

Background
[4] /usageboard/2006/2006-01.definitions/2005-09-10.meeting-notes-excerpts.html

The background for this issue is that documentation of any (and all) changes to the original fifteen elements needs to be done on a priority basis in order to prepare for a potential NISO review later this year.

Andrew has compared the draft against what was decided last year in Madrid [3].

Aside from polishing the text explaining the changes in the above, an announcement giving overall context for the comment period on dc-general needs to be finalized, see first draft at [2] (maybe move this draft into the Wiki).

As of 2006-03-30, the proposed changes need to be justified. Each individual term needs to have a justification for the change proposed, even if this is just cut-and-paste from another term. Several sections marked (@@@@TODO) need to be filled in. An announcement giving overall context for the comment period and justification for the changes (especially regarding "content of the resource") needs to be finalized.

ACTION 2006-03-16 Andrew: Wordsmith justifications for the proposed changes, also on a per-element basis (sections marked "@@@TODO"), with particular attention to "content of the resource".

ACTION 2006-03-16 Andrew: Suggest a comment for dc:language along these lines ("Use an encoding scheme...") and post to dc-usage for discussion. (This was done on 2006-03-28 and the proposed comment text has been integrated into [1]).

In February 2006, discussion on the DC-Government and DC-Corporate lists suggested that DCMI re-think its approach to 2- and 3-letter language codes.

In effect, the change in wording proposed in [1] (recommending that one follow RFC 3066 in using 2-letter codes when available) is at odds with the preference of many implementers for 3-letter codes. While this issue may require special attention on the list, the consensus was to try to keep this as part of the current batch of changes.

Diane suggested that the crux of the issue is what "recommendation" means. Why does DCMI need to recommend one or the other? One way to handle this problem is to generalize the comment. From an interoperability point of view, the crucial point is that whichever way an implementer does it, an encoding scheme should be used. For example, in the comment to dc:coverage, we say simply: "Recommended best practice is to select a value from a controlled vocabulary..."

ACTION 2006-03-16 Andrew: Suggest a comment for dc:language along these lines ("Use an encoding scheme...") and post to dc-usage for discussion.

Public comment on changes to DCMI term descriptions

Identifier: http://dublincore.org/usage/public-comment/2006/03/term-changes/
Date: 2006-03-13

Introduction

This document indicates some potential problems with the wording of some of the definitions in the DCMES and proposes some alternatives.

Note on the use of "content of the resource" in definitions

A number of the current terms in the DCMES use "content of the resource" rather than just "resource" in their definitions (contributor, coverage, creator, description, subject and type). Furthermore, language uses "intellectual content of the resource". This appears to arise from an attempt to differentiate those terms that are used to describe the "work" from those that are used to describe the "manifestation". It is primarily a feature of the fact that DCMI was never very clear about what kinds of resources its terms are used to describe, either in the general case or in the specific case of particular descriptions.

I would recommend changing all usage of "[intellectual] content of the resource" to "resource" in the definitions (consider the case where DC is being used to describe abstract concepts, as we do ourselves, or natural objects for example). (Note that this document does not make explicit proposals for doing this, except where the definitions of terms are being changed for other reasons.) Coverage is problematical, because in this case the phrase "content of the resource" is used to indicate that the definition is referring to what the resource is about rather than where it is or when it was created. It might be helpful to consider using the same language in coverage as is used in subject (e.g. "the spatial or temporal topic of the resource") as suggested above?

Changes to http://purl.org/dc/elements/1.1/coverage

Current Label
: Coverage

Current Definition
: The extent or scope of the content of the resource.

Current Comment
: Coverage will typically include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and that, where appropriate, named places or time periods be used in preference to numeric identifiers such as sets of coordinates or date ranges.

Problem
: The use of "extent" in the definition is potentially confusing with respect to the Format element and Extent element refinement. Furthermore, it's not totally clear what 'scope' means?

Proposed definition
: **The spatial or temporal topic of the resource, or the jurisdiction under which the resource is relevant.**

Proposed comment
: **Spatial topic may be a named place or a location specified by its geographic coordinates. Temporal period may be a named period, date, or date range. A jurisdiction may be a named administrative entity. Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and that, where appropriate, named places or time periods be used in preference to numeric identifiers such as sets of coordinates or date ranges.**

Changes to http://purl.org/dc/elements/1.1/description

Current Label

Description

Current Definition

An account of the content of the resource.

Current Comment

Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

Problem

The definition should not refer to the "content of the resource", simply to the "resource" (see the 'Note' above). The comment should not refer to "a reference to a graphical representation", simply to "a graphical representation".

Proposed definition

An account of the resource.

Proposed comment

Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource.

Changes to http://purl.org/dc/elements/1.1/contributor

Current Label

Contributor

Current Definition

An entity primarily responsible for making contributions to the content of the resource.

Problem

The definition should not refer to the "content of the resource", simply to the "resource".

Proposed definition

An entity primarily responsible for making contributions to the resource.

Changes to http://purl.org/dc/elements/1.1/creator

Current Label

Creator

Current Definition

An entity primarily responsible for making the content of the resource.

Problem

The definition should not refer to the "content of the resource", simply to the "resource", as per the note above.

Proposed definition

An entity primarily responsible for making the resource.

Changes to http://purl.org/dc/elements/1.1/identifier

Current Label

Resource Identifier

Problem

Only Identifier and Type explicitly use a label containing the word 'Resource'. We don't for example, talk about 'Resource Title' or 'Resource date'. In fact 'resource' is really implied in each property label. The label for Identifier should be changed for the sake of consistency.

Proposed Label

Identifier

Changes to http://purl.org/dc/elements/1.1/format

Current Label

Format

Current Definition

The physical or digital manifestation of the resource.

Current Comment

Typically, Format may include the media-type or dimensions of the resource. Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

Problem

The current definition implies that the value is a "manifestation of the resource", which is not the intention. This issue has previously been raised on one of the the DCMI lists (I forget when and where) which suggests that at least one real end-user has mis-interpreted this wording in this way.

Proposed definition

The medium, Internet Media Type or extent of the resource.

Proposed comment

Format may be used to select the software, hardware or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

Changes to http://purl.org/dc/elements/1.1/language

Current Label

Language

Current Definition
A language of the intellectual content of the resource.
Current Comment
Recommended best practice is to use RFC 3066 [RFC3066], which, in conjunction with ISO 639 [ISO639], defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk" for Akkadian, and "en-GB" for English used in the United Kingdom.
Problem
The UB position generally is that we should not be (or be seen to be) prescriptive about which encoding schemes should be used. Most other comments where it is appropriate, apart from Language and Date, recommend the use of an encoding scheme but do not specify one. It is proposed to amend the comment to reflect this, while retaining RFC3066 as an example. However, RFC3066 says that if both 2- and 3-letter codes exist, then the 2-letter code must be used. So, the 'en/eng' example is therefore wrong. Change to 'fr' to broaden number of examples ('en-GB' already appearing later in the sentence). Also, the use of "intellectual content of the" is unnecessary.
Proposed definition
A language of the resource.
Proposed comment
Recommended best practice is to use a value taken from an encoding scheme, such as RFC 3066 [RFC3066], which, in conjunction with ISO 639 [ISO639], defines two- and three-letter primary language tags with optional subtags. Examples from RFC 3066 include "fr" for French, "akk" for Akkadian, and "en-GB" for English used in the United Kingdom.

Changes to http://purl.org/dc/elements/1.1/relation

Current Label
Relation
Current Definition
A reference to a related resource.
Current Comment
Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.
Problem
As per the DCAM, the value is the related resource, not a reference to the resource.
Proposed definition
A related resource.
Proposed comment
[Unchanged] **Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.**

Changes to http://purl.org/dc/elements/1.1/rights

Current Label
Rights Management
Current Definition
Information about rights held in and over the resource.
Current Comment
Typically, Rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions may be made about any rights held in or over the resource.
Problem
The comment refers both to a 'statement' and a reference to a service that provides a statement'. This is an implementation issue and shouldn't be in the comment. Also, the explicit mention about what applications should do or not do if the Rights element is missing is inappropriate because it implies that the element can only be used (or not) as part of a particular set of properties.
Proposed definition
[Unchanged] **Information about rights held in and over the resource.**
Proposed comment
Typically, Rights information includes a rights management statement for the resource. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights.

Changes to http://purl.org/dc/elements/1.1/source

Current Label
Source
Current Definition
A Reference to a resource from which the present resource is derived.
Current Comment
The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.
Problem
Same as with Relation. Also, the use of "Source resource" is horrible. Use of "described resource" is better than "present resource".
Proposed definition
A resource from which the described resource is derived.
Proposed comment
The described resource may be derived from the referenced resource in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.

Changes to http://purl.org/dc/elements/1.1/subject

Current Label
Subject and Keywords
Current Definition
The topic of the content of the resource.
Current Comment
Typically, a Subject will be expressed as keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.
Problem
Current label is misleading. Use of "content of the" in definition is inappropriate. "Better to use "represented using" rather than "expressed as" in the comment. Need to acknowledge use of Coverage for spatial or temporal topics. Depending on the agreed definition of Coverage, it may be appropriate to add "To describe the spatial or temporal topic of the resource, use the Coverage element" to the comment.
Proposed label
Subject
Proposed definition
The topic of the resource.
Proposed comment
Typically, a Subject will be represented using keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.

Changes to http://purl.org/dc/elements/1.1/type

Current Label
Resource Type
Current Definition
The nature or genre of the content of the resource.
Current Comment
Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCMITYPE]). To describe the physical or digital manifestation of the resource, use the Format element.
Problem
The comment echoes the current definition of Format. If the definition of Format is changed then this should be changed as well. Also, the comment refers to "content", rather than specifically to the resource. Only Identifier and Type explicitly use a label containing the word 'Resource'. We don't for example, talk about 'Resource Title' or 'Resource date'. In fact 'resource' is really implied in each property label. The label for Type should be changed for the sake of consistency.
Proposed Label
Type
Proposed definition
The nature or genre of the resource.
Proposed comment
Type can be used to indicate the general category, function, genre, or aggregation level of the resource. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCMITYPE]). To describe the medium, Internet Media Type or extent of the resource, use the Format element.

Changes to http://purl.org/dc/elements/1.1/date

Current Comment
Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and follows the YYYY-MM-DD format.
Problem
We do not recommend specific schemes in most cases where encoding schemes should be used as a source of values for properties, apart from Language and Date. Our position generally is that encoding schemes should be used but we seem reluctant to be prescriptive (or to be seen to be so). The comment for Date should be changed to reflect this position, while retaining WC3DTF as an example..
Proposed comment
Typically, Date will be associated with the creation or availability of the resource. A date value may be a single date or a date range. Date values may express temporal information at any level of granularity (including time). Recommended best practice is to supply an unambiguous representation of the single date or date range using a syntax encoding scheme, such as the W3CDTF profile of ISO 8601.

Metadata associated with this resource: http://dublincore.org/usageboard/2006/2006-01.definitions/term-changes/index.shtml.rdf

---

2006-03-31

The DCMI Usage Board announces the beginning of a comment period for a number of editorial revisions to labels, definitions and comments for the original 15 elements, plus revisions to definitions and comments for several of the terms in the DCMIType Vocabulary.

Recognizing the problems inherent in a decade-long legacy of practice, the Board has undertaken a general cleanup and reconciliation of terminology, capitalization, and style that had accumulated since the submission of the original 15 elements for standardization within NISO and ISO.

In addition, improvements of definitional conformance to the 1:1 rule were included, in particular, the elimination of the usage of the wording "content of a resource" which appeared to undercut the rule.

Some changes were made to reduce overlap of definitions, based on experiences of implementers and the confusion that had been noted in application as the continuing refinement of terms tended to highlight anomalies and problems in definitions.

In reviewing and suggesting changes in these areas, the Usage Board hopes to create a more solid foundation for moving forward with Application Profiles and improved metadata description as well as to reflect our improved understanding of the needs of current and prospective implementors.

Comments from the community are encouraged, both in general concerning the changes as a whole and on particular changes. To comment on individual changes, please separate the comments into individual messages to enable clearer threading on the list archive.

Date: Wed, 23 Mar 2005 09:37:56 +0100
From: Thomas Baker <thomas.baker@bi.fhg.de>
To: DCMI Usage Board <dc-usage@jiscmail.ac.uk>
Subject: Problem with comment for dc:language?

Dear all,

We received this comment from Martin Dürst of W3C about the Comment for
dc:language.

Tom

On Tue, Mar 22, 2005 at 06:09:08PM -0500, duerst@w3.org wrote:
> Submitter Name: Martin Dürst
> Submitter Email: duerst@w3.org
> Feedback Type: Comment
>
> This is a comment about http://dublincore.org/documents/dces/,
> in particular Element Name: Language. It currently says:
> 'Recommended best practice is to use RFC 3066 .... Examples include "en" or "eng" for English...'. This
should be fixed.
> "eng" is not valid in RFC 3066; RFC 3066 clearly says that
> if there is a two-letter and a three-letter code for a
> language, the two-letter code MUST be used.
>
> Regards, Martin.

Date: Wed, 23 Mar 2005 08:51:33 -0000
Reply-To: "A mailing list for the Dublin Core Metadata Initiative's Usage Board" <DC-
USAGE@JISCMAIL.AC.UK>
Sender: "A mailing list for the Dublin Core Metadata Initiative's Usage Board" <DC-
USAGE@JISCMAIL.AC.UK>
From: Andrew Wilson <andrew.c.wilson@AHDS.AC.UK>
Subject: Re: Problem with comment for dc:language?
To: DC-USAGE@JISCMAIL.AC.UK

Well he's right. RFC 3066 says quite explicitly that you must use a 2 letter
code from ISO 639-1 unless the language only has a 3 letter code. We need to
change the comment .

Title: Changes to terms in the DCTERMS namespace
Identifier: /usage/meetings/2006/04/seattle/term-changes/index.shtml

Discussant: Diane

[1] /usageboardwiki/TermChanges
[2] /usage/meetings/2006/04/seattle/domains-ranges/2006-03-11.educationLevel.html

As decided on 2006-02-23, Tom split off changes to terms of the DCTERMS namespace into a separate document and placed it in the Wiki [1]. Additional errata and unfinished business with regard to DCTERMS terms have since been added by Tom (DDC, alternative) and Diane (date refinements).

We should make sure the document is checked for spelling consistency -- organisation/organization (to prefer the latter).

Also, dcterms:educationLevel needs a better definition [2].

Editorial changes to terms in the DCTERMS Namespace

This document is part of the DC Usage Board Wiki.

IMPORTANT: Do **not** cite materials in this Wiki other than for the purposes of collaborating on document creation. This Wiki is intended to be used to work on draft copies of documents. Finished documents will be published, in a persistent and citable form, on the dublincore.org Web site (or elsewhere in some cases).

Please announce any changes to this Wiki via the DC-USAGE@jiscmail.ac.uk list.

Changes to Refinement: conformsTo

- URI: http://purl.org/dc/terms/conformsTo
- Current Label: Conforms To
- Current Definition: A reference to an established standard to which the resource conforms.
- Current Comment: None.
- Problem: The use of "A reference to" in the definition is problematic.
- Proposed definition: ""An established standard to which the resource conforms.""
- Proposed comment: [Unchanged] None.

Changes to Syntax Encoding Scheme: URI

- URI: http://purl.org/dc/terms/uri
- Current Label: URI
- Current "See" reference: http://www.ietf.org/rfc/rfc2396.txt
- Problem: RFC 2396 ("Uniform Resource Identifiers: Generic Syntax") has been superseded by RFC 3986.
- Proposed "See" reference: http://www.ietf.org/rfc/rfc3986.txt

Changes to Refinement: Alternative

- URI: http://purl.org/dc/terms/alternative
- Current Label: Alternative
- Current Definition: Any form of the title used as a substitute or alternative to the formal title of the resource.
- Current Comment: This qualifier can include Title abbreviations as well as translations.
- Problem: The property definition refers to the notion of an "alternative title", so it should be labeled as such. The comment refers to a "qualifier", using terminology that is no longer consistent with the DCMI Abstract Model as of 2005.
- Proposed Label: Alternative Title
- Proposed Comment: An abbreviation or translation of a title may be considered an alternative title.

Changes to Refinement: Created

- URI: http://purl.org/dc/terms/created
- Current Label: Created
- Current Definition: Date of creation of the resource.

- Current Comment: none
- Problem: Later Date refinements have included "Date" in the Label, and this form is now preferred.
- Proposed Label: Date Created
- Proposed Comment: [unchanged] none

Changes to Refinement: Valid

- URI: http://purl.org/dc/terms/valid
- Current Label: Valid
- Current Definition: Date (often a range) of validity of a resource.
- Current Comment: none
- Problem: Later Date refinements have included "Date" in the Label, and this form is now preferred.
- Proposed Label: Date Valid
- Proposed Comment: [unchanged] none

Changes to Refinement: Available

- URI: http://purl.org/dc/terms/available
- Current Label: Available
- Current Definition: Date (often a range) that the resource will become or did become available.
- Current Comment: none
- Problem: Later Date refinements have included "Date" in the Label, and this form is now preferred.
- Proposed Label: Date Available
- Proposed Comment: [unchanged] none

Changes to Refinement: Issued

- URI: http://purl.org/dc/terms/issued
- Current Label: Issued
- Current Definition: Date of formal issuance (e.g., publication) of the resource.
- Current Comment: none
- Problem: Later Date refinements have included "Date" in the Label, and this form is now preferred.
- Proposed Label: Date Issued
- Proposed Comment: [unchanged] none

Changes to Refinement: Modified

- URI: http://purl.org/dc/terms/modified
- Current Label: Modified
- Current Definition: Date on which the resource was changed.
- Current Comment: none
- Problem: Later Date refinements have included "Date" in the Label, and this form is now preferred.
- Proposed Label: Date Modified
- Proposed Comment: [unchanged] none

Changes to Encoding Scheme: DDC

- URI: http://purl.org/dc/terms/DDC
- Current "See": http://www.oclc.org/dewey/index.htm
- Problem: 404 not found
- Proposed "See": http://www.oclc.org/dewey/

On 2005-06-03, Pete Johnston wrote:

Just to bump this point I raised a while ago:

On Fri, 11 Mar 2005 17:46:08 -0000, Pete Johnston <p.johnston@UKOLN.AC.UK>
wrote:

>OK... if the intent is that the value of dcterms:educationLevel really
>_is_ a class of entity e.g. a group of learners (which was what I
>half-suspected, and which _would_ make sense of the subproperty
>relation), then
>
>(a) the current name/label are slightly confusing ("education level"
>implies (to me, at least) an abstract notion rather than a group of
>learners)
>(b) the current definition literal is wrong
>
>Also, I suspect (though I'm guessing!) implementers have been using this
>property with values that are "statements" (as the current definition
>states) or "levels" of some sort (e.g. "qualification levels") (as the
>current name/label suggest), both of which are different things from
>"classes of entity", and neither of which are compatible with the
>subproperty assertion. :-(

I had a message from Thomas Fischer at Goettingen a couple of
days ago (part of which I forward here with his permission),
in which he made exactly the same point:

<quote>
1.
Audience: A class of entity for whom the resource is intended or useful.

educationLevel: A general statement describing the education or training
context. Alternatively, a more specific statement of the location of the
audience in terms of its progression through an education or training context.

I don't understand "class of entity", but doubt that "A general statement
describing the education or training context." fits this description.
</quote>

I can't remember if this was one of the things covered under the general
clean up of labels/definitions but I think it probably wasn't? And I'm not
sure whether fixing this can be done in a way that would still qualify as
"editorial errata"?

Date: Wed, 21 Sep 2005 07:58:39 -0700
From: Stuart Sutton <sasutton@U.WASHINGTON.EDU>
Subject: Re: [DC-USAGE] Is dcterms:educationLevel a subproperty of dcterms:audience?
To: DC-USAGE@JISCMAIL.AC.UK

Well, I continue to assert that the values of educationLevel are
"classes of entities" and that when I assert an instance of the property
to be "grade 5", I am meaning the category of students (AgentGroup) that
are situated (in the US and Canada) at a certain point in their
progressing through the "system". That is why it was conceived as a
subproperty of audience and not some new property. However, I am afraid
that changing the definition will probably not help at all in changing
practice in the real world.

Date: Wed, 21 Sep 2005 15:44:40 +0100
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: [DC-USAGE] Is dcterms:educationLevel a subproperty of dcterms:audience?
To: DC-USAGE@JISCMAIL.AC.UK

I'm not sure where to put this just now, but seeing as I started the
thread here a while ago....

I think the work of the DC-RDF-Taskforce looking at property ranges and
domains

http://www.ukoln.ac.uk/twiki/bin/view/Metadata/DCPropertyDomainsRanges

highlights this problem quite clearly.

Andy has tentatively assigned an rdfs:range of some:AgentGroup to the property dcterms:audience

```
  dcterms:audience rdfs:range some:AgentGroup .
```

which seems quite reasonable, and this implies that whenever I find

```
  resource:X dcterms:audience _:x .
```

I can conclude that

```
  _:x rdf:type some:AgentGroup .
```

or: the value is an AgentGroup (an instance of the class some:AgentGroup).

I note we haven't yet suggested a range for dcterms:educationLevel, but the current definition

"A general statement describing the education or training context. Alternatively, a more specific statement of the location of the audience in terms of its progression through an education or training context."

suggests that the range should be the class of Statements of some sort (cf dc:rights, which has a range of some:RightsStatement)

So for the sake of argument let's say

```
  dcterms:educationLevel rdfs:range some:EdLevelStatement .
```

This implies that whenever I find

```
  resource:X dcterms:educationLevel _:x .
```

I can conclude that

```
  _:x rdf:type some:EdLevelStatement .
```

But also from the subproperty relation,

```
  dcterms:educationLevel rdfs:subPropertyOf dcterms:audience .
```

I can conclude

```
  resource:X dcterms:audience _:x .
```

and so

```
  _:x rdf:type some:AgentGroup .
```

i.e. the value is both an AgentGroup and an Education Levels Statement. I find it hard to imagine a case where that makes sense :-(

So I think either

(a) the subproperty assertion is wrong and should be withdrawn or
(b) the definition of educationLevel needs to be amended so that its values are "classes of entity....", which Stuart suggested was the intent [1]

However as I said back at [2] I suspect that current use of dcterms:educationLevel probably includes values that are either "statements" or "levels" and not just "classes of entity..." :-(

Pete

[1] http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0503&L=dc-usage&P=5801
[2] http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0503&L=dc-usage&P=6068

Date: Fri, 11 Mar 2005 11:09:21 -0000
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Is dcterms:educationLevel a subproperty of dcterms:audience?

Having had a few queries about element refinements recently
(partly as a result of the document currently out for review
on the AB list), I've been looking at some of the current
DCMI subproperty assertions, and I've ended up asking myself:

Is it correct to describe dcterms:educationLevel as a
subproperty/refinement of dcterms:audience (N.B. as both of
these properties are currently defined) ?

dcterms:educationLevel
======================

Definition : A general statement describing the education or
training context. Alternatively, a more specific statement
of the location of the audience in terms of its progression
through an education or training context.

dcterms:audience
================

Definition: A class of entity for whom the resource is intended
or useful.

Comment: A class of entity may be determined by the creator
or the publisher or by a third party.

Is it the case that:

for every statement (1) x dcterms:educationLevel y

it is also true that (2) x dcterms:audience y

?

I really don't think it _is_ the case, and I'm really
uneasy about that subproperty assertion. If the value
in statement (1) is a "general statement...." or a "more
specific statement...." (which is what the definition of
dcterms:educationLevel says it should be), I can not see how
that value can also be "a class of entity...".

For example, I think the definition of dcterms:educationLevel
permits me to say

x dcterms:educationLevel "This resource is for beginner
guitarists."

The subproperty assertion then allows me to infer

x dcterms:audience "This resource is for beginner guitarists."

But that value is not "a class of entity...."; it's a
statement.

Now, yes, it may be a statement which includes
some information _about_ a class of entity, but that is
not the same thing. I know, I know, that must seem like the
most awful nit-picking, and a human reader might manage the
distinction, but subproperty assertions are designed to be
used by software applications, and an application which is
expecting the value to be "A class of entity" will _not_
manage that inferred statement correctly.

(And in fact, the decision _not_ to assert that
dcterms:rightsHolder was a subproperty of dc:rights was taken
precisely to avoid ending up with the reverse situation,
and supporting inferences that an entity - a value of
dcterms:rightsHolder - was "information about rights held...",
which is what dc:rights says its values should be.)

Now then, it may be that the intent with dcterms:educationLevel
really was/is that values are "classes of entity" and
indeed the examples in the term proposal [1] ("Grade 5" =
the class of students at that Grade; "Preschool" = the class
of preschool children) do seem to hint this is the case.

But (IMHO) the current definitions of the properties do make

the subproperty assertion problematic. :-(

(Sorry!)

Title: Replicating ELEMENTS1.1 terms in the DCTERMS namespace
Identifier: /admin/www/usage/meetings/2006/04/seattle/all-in-dcterms/
Created: 2006-03-30

Discussion leader: Andy

Reading:
-- /usageboardwiki/ReplicatingDCMESINDCTERMS

In Madrid (September 2005), the Architecture Working Group discussed the possibility of adding the fifteen DCMES terms [1] to the DCTERMS namespace [2] (in addition to them being in the existing DCMES namespace) -- i.e., to coin [3] in addition to [4]. This would mean that many DC users would only need to use a single DCTERMS namespace. In Madrid, there was no clear agreement about whether this should be done or not.

Such a change, if approved, would require further changes to the namespace policy [5,6]. In fact, the change would have a cascading effect on all DCMI documentation that cites DCMES URIs.

In the 2006-03-23, we agreed on the need for a small document summarizing reasons pro and con. We need a coherent line on why we would bother to do this. One possibility from a process point of view would be for Architecture to ask UB to take this on as a work item.

ACTION 2006-03-23: Andy - post a discussion for the rationale to Architecture list for replicating Elements 1.1 in DCTERMS This input on the list would become input at UB in Seattle.

[1] http://purl.org/dc/elements/1.1/
[2] http://purl.org/dc/terms/
[3] http://purl.org/dc/terms/title
[4] http://purl.org/dc/elements/1.1/title
[5] /documents/dcmi-namespace/
[6] /architecturewiki/NamespacePolicy

Replicating DCMES properties in the DCTERMS namespace

This document is part of the DC Usage Board Wiki.

IMPORTANT: Do **not** cite materials in this Wiki other than for the purposes of collaborating on document creation. This Wiki is intended to be used to work on draft copies of documents. Finished documents will be published, in a persistent and citable form, on the dublincore.org Web site (or elsewhere in some cases).

Please announce any significant changes to this Wiki via the DC-USAGE@jiscmail.ac.uk list.

Proposal

- The 15 DCMES properties that are currently assigned URIs in the DC namespace (http://purl.org/dc/elements/1.1/) will be replicated in the DCTERMS namespace (http://purl.org/dc/terms/), i.e. 15 new terms will be added to the DCTERMS namespace.
- The existing terms in the DC namespace will continue to be valid, however implementors will be encouraged to use the newer terms.
- The RDFS/OWL machine-readable declarations for the 15 existing terms will be modified to indicate the equivalences between the existing terms and the newer terms.
- The RDFS/OWL machine-readable declarations for the 15 new terms will indicate the equivalences between the newer terms and the existing terms.
- The human-readable declarations of both sets of terms will indicate the equivalences between the older and newer terms and will recommend usage of the newer terms.
- The DCMI namespace policy and the various encoding guidelines (XHTML, XML and RDF) will be updated to use the newer terms.

Background

The Namespace Policy for the Dublin Core Metadata Initiative (DCMI) defines how term URIs are assigned to all DCMI metadata terms and what guarantees DCMI makes about their persistence.

The 15 terms (elements) that make up the "Dublin Core Metadata Element Set, Version 1.1" are assigned URIs in the DC namespace (i.e. they are assigned URIs that begin with http://purl.org/dc/terms/1.1/). All other DCMI terms (including new elements, element refinements, encoding schemes and terms in DCMI-maintained controlled vocabularies) are either assigned URIs in the DCTERMS namespace or in a vocabulary-specific namespace.

Implementors of metadata systems using DC elements and element refinements that go beyond the use of the 15 DCMES elements therefore need to build knowledge about two different namespaces into their applications. This causes some confusion about the correct URI for terms in the different namespaces. For example, there is some evidence of people using 'DC' URIs for terms in the DCTERMS namespace and vice versa. There is also some evidence that the perceived hurdle of having to deal with two namespaces rather than one is sufficient to cause implementors to ignore Dublin Core metadata in favour of their own (or third-party) solutions.

According to the current namespace policy, the following XHTML example use of two DCMI elements is correct:

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />

<meta name="DC.title" content="_the title_" />
<meta name="DCTERMS.audience" content="_the audience_" />
```

Under the proposed change, this would become:

```
<link rel="schema.DC" href="http://purl.org/dc/terms/" />

<meta name="DC.title" content="_the title_" />
<meta name="DC.audience" content="_the audience_" />
```

Impact of this proposal

On metadata producers

Systems that expose metadata (e.g. for harvesting or in search results) will not need to be changed in any way, though the implementors of such systems will be encouraged to move to using the 15 new terms in the DCTERMS namespace rather than the older terms in the DC namespace.

On metadata consumers

Systems that consume metadata (e.g. by harvesting it or receiving it in search results) will need to be modified to treat the 15 new terms in the DCTERMS namespace as equivalents for the 15 older terms in the DC namespace. This may be done by hardwiring such knowledge into the software, or by extracting this knowledge automatically from the machine-readable schema declarations of the DCMI terms (i.e. the RDFS/OWL schemas that are available at the two namespace URIs).

On DCMI documentation

The DCMI namespace policy and the XHTML, XML and RDF encoding guidelines will all need to be updated.

The machine-readable RDFS/OWL schemas available at the DCMI namespace URIs will need to be updated.

The DCMI-maintained XML schemas will need to be updated.

Arguments in favour of the proposal

- Replicating the 15 DCMES terms in the DCTERMS namespace means that implementors need only deal with a single namespace for all DC elements and element refinements.

Arguments against the proposal

- Although in the long term this proposal should result in less confusion amoungst implementors, it may result in greater confusion in the short term.
- The use of separate DC and DCTERMS namespaces may be perceived to be helpful in highlighting the 'core' natures of the 15 DCMES properties.

Note on the continued use of PURLs as metadata term identifiers

This proposal would result in the coining of fifteen new PURL-based identifiers for DCMI metadata terms. It should be noted that the continued use of PURLs as the base URI for DCTERMS identifiers will consolidate a situation in which DCMI cannot comply with a June 2005 W3C TAG resolution on the use of http URIs for non-information resources such as metadata terms. As explained in the W3C Working Draft Best Practice Recipes for Publishing RDF Vocabularies:

> When the central PURL server was originally developed in the 1990s, the standard response of an HTTP to a request against a PURL was to return a response code of 302 ("temporarily moved"). Web architecture has evolved since then, and the Technical Architecture Group (TAG) of W3C has resolved that, for the purpose of such redirects, the response code 303 ("see other") should be returned. As PURL servers use a 302 response code and there is currently no way to configure them to use 303 response codes, existing vocabularies with http://purl.orgslash namespaces servers do not strictly conform to the current TAG recommendations.

If the PURL server were modified to support the selection of a response code by the maintainer of a PURL, this would solve the problem. However, it is unclear whether the maintainers of the PURL server software would want to do this.

Title: DCMI property domains and ranges
Identifier: /admin/www/usage/meetings/2006/04/seattle/domains-ranges/
Created: 2006-03-30

Shepherd: Andy

Readings:
[1] /architecturewiki/DCPropertyDomainsRanges
[2] /architecturewiki/DCRDFTaskforce/DCRDFExecutiveSummary3
[3] /usage/meetings/2006/04/seattle/domains-ranges/2006-03-24.domain-range-rationale.html
[4] /usage/meetings/2006/04/seattle/domains-ranges/2006-03-28.domain-range-comments.html
[5] /usage/meetings/2006/04/seattle/domains-ranges/2006-03-24.dcPropertiesRanges.pdf
[6] /usage/meetings/2006/04/seattle/domains-ranges/2005-02-26.Encoding-scheme-types.html

See also:
[7] /usage/meetings/2006/04/seattle/domains-ranges/2006-03-11.educationLevel.txt
[8] /usage/meetings/2006/04/seattle/dcmitype/
[9] /usage/public-comment/2006/03/type-vocabulary-comments/
[10] /usage/public-comment/2005/12/type-vocabulary-changes/

In his draft DCPropertyDomainsRanges [1], Andy has proposed domains and ranges for all DCMI terms. In Seattle, our goal should be to decide whether this is a reasonable thing to do. We should weigh whether we want to undertake this exercise at all and assess its implications.

Many of the domain and range proposals are probably uncontroversial. Some notable exceptions:

-- dcterms:educationalLevel
-- dc:creator and dc:creator (see discussion DCRDFExecutiveSummary3 [2])

One important implication is that many of the 30-40 "possible classes" proposed as domains and ranges would need to be defined, approved, given URIs, formally declared, and maintained. Should DCMI do this? Giving them URIs would involve either expanding the DCMI Type Vocabulary or creating a new Vocabulary. Creating a new Vocabulary would involve revising the DCMI Namespace Policy.

In a conference call, Diane voiced concern with legacy implementations. Adding domains and ranges would seem to introduce more complexity in terms of what is considered "appropriate data" and suggests that we should not just assume that because it is a good thing technically, we should actually go ahead and do it.

Giving definitions to these classes would also involve deciding on a style for definitions. Ideally, this style would be consistent between this new Domain-Range Vocabulary and the existing DCMI Type Vocabulary. Since we happen to be finalizing a set of changes to the DCMI Type Vocabulary right now [7], we have an opportunity to ensure that a common style is adopted. Possibilities include:

-- The style currently proposed for the DCMI Type Vocabulary [10], e.g.:

  -- Collection: An aggregation of items.
  -- Dataset: Information encoded in a defined structure.
  -- Image: A primarily symbolic visual representation other than text.

-- DCMI Type Style, Renaud style [9], e.g.:

  -- Collection: A resource which is an aggregation of items.
  -- Dataset: A resource in which data is encoded in a defined structure.
  -- Image: A resource which is a visual, non-textual representation.

-- Domain-Range Vocabulary style [1], e.g.:

```
      -- DigitalResource: The class of all digital resources.

      Accordingly...
      -- Collection: The class of all aggregations of items.
      -- Dataset: The class of all data encoded in a defined
         structure.
      -- Image: The class of all visual, non-textual
         representations.
```

These alternatives are, of course, not just "stylistic" changes. Some of them involve making explicit what is currently implicit in a definition.

Mikael has obtained statistical indexing data from Swoogle [5], included in the packet as input to this discussion.

It should also be noted that currently, the Web page at /documents/dcmi-terms/ asserts itself to be "an up-to-date, authoritative specification of all metadata terms maintained by DCMI".

DCMI currently makes no such claims for the RDF schema representation of its terms. Indeed, the only policy statement on the subject, at /schemas/rdfs/, says that "users of RDF guidelines and schemas posted on the DCMI Web site need to be aware that these resources may be subject to change based on the results of further discussions within DCMI and W3C" -- a situation that can hopefully be remedied by the work of the DC RDF Task force.

If in addition to the "natural-language" definitions currently provided in the Web documents, DCMI were also to provide "definitive" RDF schemas, then DCMI would be saying, in effect, that its terms are defined not just by natural-language definitions, but also by the sum of formal assertions and relations, within which the terms are embedded, as expressed in the RDF schema.

We would need to consider whether it would be realistic for DCMI to claim that both the Web document and the RDF schema are "authoritative" -- raising the bar for keeping the documents not only in synch, but for expressing formal assertions adequately in the Web documents -- or whether one should be definitive while the other is considered to be derived.

The finalization of authoritative domain and range declarations has implications for DCMI process, as they would presumably be subject to review and maintenance by the Usage Board.

As of the 2006-03-23 telecon, this document now belongs to the Usage Board.

As one piece of unfinished business in this regard, we should recall that the existing encoding schemes are still designated as "encoding schemes" in a generic sense and are not yet differentiated into "vocabulary encoding schemes" and "syntax encoding schemes" [6].

ACTION 2006-03-23. Andy Consider removing the FRBR-related classes (see discussion in [4]).

DC property domains and ranges

This document is currently under development. It is being worked on by the DC RDF Taskforce. Comments should be sent to the dc-rdf-taskforce@jiscmail.ac.uk mailing list.

Introduction

This document indicates the domains and ranges that apply to all DCMI properties.

The domain of a property indicates the class (or classes) of resources that the property can be used to describe. The range of a property indicates the class (or classes) of resources that can be used as values of the property.

Possible classes

This section lists some classes that we are likely to need in order to assign domains and ranges to DCMI properties.

Resource

The class of everything.

Agent

The class of all things that are a Person, Organization or Service.

Person

The class of all people (living or dead).

Organization

The class of all organizations (and other corporate bodies).

Service

The class of all services (as defined in DCMIType).

AgentGroup

The class of all things that are groups of Agents (e.g. students, women, charities, lecturers).

NonAgentResource

The class of everything that is not an Agent.

ConceptualResource

The class of all concepts.

DigitalResource

The class of all digital Resources.

PhysicalResource

The class of all physical Resources.

Collection

The class of everything that is an aggregation of one or more DigitalResources or PhysicalResources.

Literal

The class of all literal strings.

Location

The class of all places and geographical regions.

Period

The class of all dates and date ranges.

Jurisdiction

The class of all administrative entities.

LocationPeriodOrJurisdiction

The class of things that are a Location, Period or Jurisdiction.

FileFormat

The class of all DigitalResource formats (e.g. as defined by the list of Internet Media Types).

PhysicalMedium

The class of all things that are a material or physical carrier of a PhysicalResource (e.g. paper, canvas, etc.).

MediaType

The class of all things that are a FileFormat or PhysicalMedium.

Extent

The class of all things that are a size of a DigitalResource or PhysicalResource (e.g. length/width/breadth, number of pages, etc.).

Duration

The class of all things that are a time taken to 'play' a DigitalResource or PhysicalResource (e.g. in hours/minutes/seconds).

Dimensions

The class of things that are an Extent or Duration.

MediaTypeOrDimensions

The class of things that are a MediaType or Dimensions.

Reference

The class of all things that are an identifier for a Resource that is unambiguous in a given context (e.g. a URI).

Language

The class of all human languages.

RightsStatement

The class of all things that are a rights statement about a NonAgentResource.

Topic

The class of all subjects.

Class

The class of all classes.

Frequency

The class of all things that are statements about the rate at which something recurs.

AccrualMethod

The class of all things that are methods by which Resources are added to a Collection.

Policy

The class of all things that are a plan or course of action, as of a government, political party, or business, intended to influence and determine decisions, actions, and other matters.

InstructionalMethod

The class of all things that are a process that is used to engender knowledge, attitudes and skills.

ProvenenceStatement

The class of all things that are a statement of any changes in ownership and custody of a NonAgentResource since its creation that are significant for its authenticity, integrity and interpretation.

Domains and ranges of DC properties

The Dublin Core Metadata Element Set

contributor

URI: http://purl.org/dc/elements/1.1/contributor

Definition: An entity responsible for making contributions to the content of the resource.

Comment: Examples of a Contributor include a person, an organisation, or a service. Typically, the name of a Contributor should be used to indicate the entity.

Domain: NonAgentResource

Range: Agent

coverage

URI: http://purl.org/dc/elements/1.1/coverage

Definition: The extent or scope of the content of the resource.

Comment: Coverage will typically include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and that, where appropriate, named places or time periods be used in preference to numeric identifiers such as sets of coordinates or date ranges.

References: [TGN] http://www.getty.edu/research/tools/vocabulary/tgn/index.html

Domain: NonAgentResource

Range: LocationPeriodOrJurisdiction

creator

URI: http://purl.org/dc/elements/1.1/creator

Definition: An entity primarily responsible for making the content of the resource.

Comment: Examples of a Creator include a person, an organisation, or a service. Typically, the name of a Creator should be used to indicate the entity.

Domain: NonAgentResource

Range: Agent

date

URI: http://purl.org/dc/elements/1.1/date

Definition: A date associated with an event in the life cycle of the resource.

Comment: Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and follows the YYYY-MM-DD format.

References: [W3CDTF] http://www.w3.org/TR/NOTE-datetime

Domain: NonAgentResource

Range: Period

description

URI: http://purl.org/dc/elements/1.1/description

Definition: An account of the content of the resource.

Comment: Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

Domain: NonAgentResource

Range: NonAgentResource

format

URI: http://purl.org/dc/elements/1.1/format

Definition: The physical or digital manifestation of the resource.

Comment: Typically, Format may include the media-type or dimensions of the resource. Format may be used to determine the software, hardware or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

References: [MIME] http://www.iana.org/assignments/media-types/

Domain: Manifestation

Range: MediaTypeOrDimensions

identifier

URI: http://purl.org/dc/elements/1.1/identifier

Definition: An unambiguous reference to the resource within a given context.

Comment: Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Example formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).

Domain: Resource

Range: Reference

language

URI: http://purl.org/dc/elements/1.1/language

Definition: A language of the intellectual content of the resource.

Comment: Recommended best practice is to use RFC 3066 [RFC3066], which, in conjunction with ISO 639 [ISO639], defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk" for Akkadian, and "en-GB" for English used in the United Kingdom.

References: [RFC3066] http://www.ietf.org/rfc/rfc3066.txt

References: [ISO639] http://www.loc.gov/standards/iso639-2/

Domain: NonAgentResource

Range: Language

publisher

URI: http://purl.org/dc/elements/1.1/publisher

Definition: An entity responsible for making the resource available

Comment: Examples of a Publisher include a person, an organisation, or a service. Typically, the name of a Publisher should be used to indicate the entity.

Domain: NonAgentResource

Range: Agent

relation

URI: http://purl.org/dc/elements/1.1/relation

Definition: A reference to a related resource.

Comment: Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.

Domain: Resource

Range: Resource

rights

URI: http://purl.org/dc/elements/1.1/rights

Definition: Information about rights held in and over the resource.

Comment: Typically, a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource.

Domain: NonAgentResource

Range: RightsStatement

source

URI: http://purl.org/dc/elements/1.1/source

Definition: A reference to a resource from which the present resource is derived.

Comment: The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.

Domain: NonAgentResource

Range: NonAgentResource

subject

URI: http://purl.org/dc/elements/1.1/subject

Definition: The topic of the content of the resource.

Comment: Typically, a Subject will be expressed as keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.

Domain: NonAgentResource

Range: Topic

title

URI: http://purl.org/dc/elements/1.1/title

Definition: A name given to the resource.

Comment: Typically, a Title will be a name by which the resource is formally known.

Domain: NonAgentResource

Range: Literal

type

URI: http://purl.org/dc/elements/1.1/type

Definition: The nature or genre of the content of the resource.

Comment: Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCMITYPE]). To describe the physical or digital manifestation of the resource, use the Format element.

References: [DCMITYPE] http://dublincore.org/documents/dcmi-type-vocabulary/

Domain: NonAgentResource

Range: Class

Other Elements and Element Refinements

abstract

URI: http://purl.org/dc/terms/abstract

Definition: A summary of the content of the resource.

Refines: http://purl.org/dc/elements/1.1/description

Domain: NonAgentResource

Range: Literal

accessRights

URI: http://purl.org/dc/terms/accessRights

Definition: Information about who can access the resource or an indication of its security status.

Comment: Access Rights may include information regarding access or restrictions based on privacy, security or other regulations.

Refines: http://purl.org/dc/elements/1.1/rights

Domain: NonAgentResource

Range: RightsStatement

accrualMethod

URI: http://purl.org/dc/terms/accrualMethod

Definition: The method by which items are added to a collection.

Comment: Recommended best practice is to use a value from a controlled vocabulary.

Domain: Collection

Range: AccrualMethod

accrualPeriodicity

URI: http://purl.org/dc/terms/accrualPeriodicity

Definition: The frequency with which items are added to a collection.

Comment: Recommended best practice is to use a value from a controlled vocabulary.

Domain: Collection

Range: Frequency

accrualPolicy

URI: http://purl.org/dc/terms/accrualPolicy

Definition: The policy governing the addition of items to a collection.

Comment: Recommended best practice is to use a value from a controlled vocabulary.

Domain: Collection

Range: Policy

alternative

URI: http://purl.org/dc/terms/alternative

Definition: Any form of the title used as a substitute or alternative to the formal title of the resource.

Comment: This qualifier can include Title abbreviations as well as translations.

Refines: http://purl.org/dc/elements/1.1/title

Domain: NonAgentResource

Range: Literal

audience

URI: http://purl.org/dc/terms/audience

Definition: A class of entity for whom the resource is intended or useful.

Comment: A class of entity may be determined by the creator or the publisher or by a third party.

Domain: NonAgentResource

Range: AgentGroup

available

URI: http://purl.org/dc/terms/available

Definition: Date (often a range) that the resource will become or did become available.

Refines: http://purl.org/dc/elements/1.1/date

Domain: NonAgentResource

Range: Period

bibliographicCitation

URI: http://purl.org/dc/terms/bibliographicCitation

Definition: A bibliographic reference for the resource.

Comment: Recommended practice is to include sufficient bibliographic detail to identify the resource as unambiguously as possible, whether or not the citation is in a standard form.

Refines: http://purl.org/dc/elements/1.1/identifier

Domain: NonAgentResource

Range: Reference

conformsTo

URI: http://purl.org/dc/terms/conformsTo

Definition: A reference to an established standard to which the resource conforms.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

created

URI: http://purl.org/dc/terms/created

Definition: Date of creation of the resource.

Refines: http://purl.org/dc/elements/1.1/date

Domain: NonAgentResource

Range: Period

dateAccepted

URI: http://purl.org/dc/terms/dateAccepted

Definition: Date of acceptance of the resource (e.g. of thesis by university department, of article by journal, etc.).

Refines: http://purl.org/dc/elements/1.1/date

Domain: NonAgentResource

Range: Period

dateCopyrighted

URI: http://purl.org/dc/terms/dateCopyrighted

Definition: Date of a statement of copyright.

Refines: http://purl.org/dc/elements/1.1/date

Domain: NonAgentResource

Range: Period

dateSubmitted

URI: http://purl.org/dc/terms/dateSubmitted

Definition: Date of submission of the resource (e.g. thesis, articles, etc.).

Refines: http://purl.org/dc/elements/1.1/date

Domain: NonAgentResource

Range: Period

educationLevel

URI: http://purl.org/dc/terms/educationLevel

Definition: A general statement describing the education or training context. Alternatively, a more specific statement of the location of the audience in terms of its progression through an education or training context.

Refines: http://purl.org/dc/terms/audience

Domain: NonAgentResource

Range: AgentGroup

Note: proposed range is in line with desire for educationLevel to be a sub-property of audience BUT it doesn't fit with the current wording of the definition.

extent

URI: http://purl.org/dc/terms/extent

Definition: The size or duration of the resource.

Refines: http://purl.org/dc/elements/1.1/format

Domain: NonAgentResource

Range: Extent

hasFormat

URI: http://purl.org/dc/terms/hasFormat

Definition: The described resource pre-existed the referenced resource, which is essentially the same intellectual content presented in another format.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

hasPart

URI: http://purl.org/dc/terms/hasPart

Definition: The described resource includes the referenced resource either physically or logically.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

hasVersion

URI: http://purl.org/dc/terms/hasVersion

Definition: The described resource has a version, edition, or adaptation, namely, the referenced resource.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

instructionalMethod

URI: http://purl.org/dc/terms/instructionalMethod

Definition: A process, used to engender knowledge, attitudes and skills, that the resource is designed to support.

Comment: Instructional Method will typically include ways of presenting instructional materials or conducting instructional activities, patterns of learner-to-learner and learner-to-instructor interactions, and mechanisms by which group and individual levels of learning are measured. Instructional methods include all aspects of the instruction and learning processes from planning and implementation through evaluation and feedback.

Domain: NonAgentResource

Range: InstructionalMethod

isFormatOf

URI: http://purl.org/dc/terms/isFormatOf

Definition: The described resource is the same intellectual content of the referenced resource, but presented in another format.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

isPartOf

URI: http://purl.org/dc/terms/isPartOf

Definition: The described resource is a physical or logical part of the referenced resource.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

isReferencedBy

URI: http://purl.org/dc/terms/isReferencedBy

Definition: The described resource is referenced, cited, or otherwise pointed to by the referenced resource.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

isReplacedBy

URI: http://purl.org/dc/terms/isReplacedBy

Definition: The described resource is supplanted, displaced, or superseded by the referenced resource.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

isRequiredBy

URI: http://purl.org/dc/terms/isRequiredBy

Definition: The described resource is required by the referenced resource, either physically or logically.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

issued

URI: http://purl.org/dc/terms/issued

Definition: Date of formal issuance (e.g., publication) of the resource.

Refines: http://purl.org/dc/elements/1.1/date

Domain: NonAgentResource

Range: Period

isVersionOf

URI: http://purl.org/dc/terms/isVersionOf

Definition: The described resource is a version, edition, or adaptation of the referenced resource. Changes in version imply substantive changes in content rather than differences in format.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

license

URI: http://purl.org/dc/terms/license

Definition: A legal document giving official permission to do something with the resource.

Comment: Recommended best practice is to identify the license using a URI. Examples of such licenses can be found at http://creativecommons.org/licenses/.

Refines: http://purl.org/dc/elements/1.1/rights

Domain: NonAgentResource

Range: RightsStatement

mediator

URI: http://purl.org/dc/terms/mediator

Definition: A class of entity that mediates access to the resource and for whom the resource is intended or useful.

Comment: The audiences for a resource are of two basic classes: (1) an ultimate beneficiary of the resource, and (2) frequently, an entity that mediates access to the resource. The mediator element refinement represents the second of these two classes.

Refines: http://purl.org/dc/terms/audience

Domain: NonAgentResource

Range: AgentGroup

medium

URI: http://purl.org/dc/terms/medium

Definition: The material or physical carrier of the resource.

Refines: http://purl.org/dc/elements/1.1/format

Domain: NonAgentResource

Range: Medium

modified

URI: http://purl.org/dc/terms/modified

Definition: Date on which the resource was changed.

Refines: http://purl.org/dc/elements/1.1/date

Domain: NonAgentResource

Range: Period

provenance

URI: http://purl.org/dc/terms/provenance

Definition: A statement of any changes in ownership and custody of the resource since its creation that are significant for its authenticity, integrity and interpretation.

Comment: The statement may include a description of any changes successive custodians made to the resource.

Domain: NonAgentResource

Range: ProvenanceStatement

references

URI: http://purl.org/dc/terms/references

Definition: The described resource references, cites, or otherwise points to the referenced resource.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

replaces

URI: http://purl.org/dc/terms/replaces

Definition: The described resource supplants, displaces, or supersedes the referenced resource.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

requires

URI: http://purl.org/dc/terms/requires

Definition: The described resource requires the referenced resource to support its function, delivery, or coherence of content.

Refines: http://purl.org/dc/elements/1.1/relation

Domain: NonAgentResource

Range: NonAgentResource

rightsHolder

URI: http://purl.org/dc/terms/rightsHolder

Definition: A person or organization owning or managing rights over the resource.

Comment: Recommended best practice is to use the URI or name of the Rights Holder to indicate the entity.

Domain: NonAgentResource

Range: Agent

spatial

URI: http://purl.org/dc/terms/spatial

Definition: Spatial characteristics of the intellectual content of the resource.

Refines: http://purl.org/dc/elements/1.1/coverage

Domain: NonAgentResource

Range: Location

tableOfContents

URI: http://purl.org/dc/terms/tableOfContents

Definition: A list of subunits of the content of the resource.

Refines: http://purl.org/dc/elements/1.1/description

Domain: NonAgentResource

Range: Literal

temporal

URI: http://purl.org/dc/terms/temporal

Definition: Temporal characteristics of the intellectual content of the resource.

Refines: http://purl.org/dc/elements/1.1/coverage

Domain: NonAgentResource

Range: Period

valid

URI: http://purl.org/dc/terms/valid

Definition: Date (often a range) of validity of a resource.

Refines: http://purl.org/dc/elements/1.1/date

Domain: NonAgentResource

Range: Period

Dublin Core metadata in RDF: implications of new guidelines for legacy implementations

About this note

DCMI is currently considering the assignment of domains and ranges to DCMI metadata terms. Such a step would have important implications for the interpretation of legacy metadata. This note presents a high-level view of the issue and its implications. ***No such changes will be undertaken by DCMI until their impact has been well understood and discussed in a public comment period.*** Implementers with an opinion about the issues presented here are invited to participate in discussion on the DCMI Architecture Working Group mailing list (http://www.jiscmail.ac.uk/lists/dc-architecture.html).

The addition of domains and ranges would help clarify the semantics of DCMI properties in a formal sense. It should be noted, however, that this would have practical consequences only for the creation and interpretation of Dublin Core metadata in RDF. Metadata creators would need to add a few extra angle brackets to ensure that RDF-consuming applications interpret value strings as properties of nodes; and metadata consumers might need to "special-case" the processing of value strings associated directly with Dublin Core properties (i.e., without intervening nodes). The generation of Dublin Core metadata in RDF would become slightly more complex for anyone producing metadata by hand. However, these measures would eliminate the current ambiguity, enabling metadata that is mappable more consistently to the DCAM and improved support by tools thanks to machine-processable restrictions.

The expression of Dublin Core metadata in the other recommended formats recommended by DCMI -- i.e., "Expressing Dublin Core in HTML/XHTML meta and link elements" (http://dublincore.org/documents/dcq-html) and the existing "Guidelines for implementing Dublin Core in XML" (http://dublincore.org/documents/dc-xml-guidelines) -- would not be negatively affected by these developments. The rules for interpreting metadata in these syntaxes in terms of the DCAM are simpler than for RDF, as they are not bound by the semantics of RDF.

Historical context

Since 1997, the "Dublin Core data model" has evolved in a process of mutual influence with W3C's Resource Description Framework (RDF). This process has resulted in the DCMI Abstract Model, which was published in 2005 as a DCMI Recommendation ([http://dublincore.org/documents/abstract-model/ (1)]. The DCMI Abstract Model now provides a reference model on the basis of which particular DC encoding guidelines ((2)) can be defined.

DCMI currently has two specifications for expressing Dublin Core metadata in RDF. The first, "Expressing Simple Dublin Core in RDF/XML", or "DC-Simple-in-RDF" for short ((3)) became a DCMI Recommendation in 2002. The second, "Expressing Qualified Dublin Core in RDF/XML", or "DC-Qualified-in-RDF" ((4)), has been a DCMI Proposed Recommendation since 2002.

Current developments

The RDF Task Force of the DCMI Architecture Working Group is currently drafting a document which is intended to replace these two legacy specifications with a single consolidated and updated DCMI Recommendation for expressing Dublin Core in RDF ([http://dublincore.org/architecturewiki/DCRDFGuidelines (5)]).

This process has important implications for how DCMI "defines" its metadata terms. DCMI metadata terms have hitherto been defined entirely in natural language; the RDF expression of the DCMI term set (e.g., (6)) served essentially to convey these English-language definitions in a form ingestable by RDF applications. As part of the process of clarifying the RDF expression for Dublin Core metadata, the RDF Task Force has recommended that DCMI supplement these English-language definitions with machine-understandable definitions of the "domain" and "range" of DCMI metadata terms ([http://dublincore.org/architecturewiki/DCPropertyDomainsRanges (5)]). Such additional, machine-understandable precision is necessary as Dublin Core is deployed in the context of inference engines and ontology-based solutions.

For most DCMI metadata terms, the process of clarifying domains and ranges machine-understandably is straightforward and unambiguous. However, one problem with regard to legacy metadata usage is serious enough to bear closer consideration. In the early years, the Dublin Core community distinguished between Simple and Qualified Dublin Core -- a distinction which was reflected in the difference between the specifications "DC-Simple-in-RDF" and the "DC-Qualified-in-RDF".

The two legacy specifications differ with regard to whether properties such as dc:creator and dc:date have values that are resources (e.g., a Person or a Date, seen as entities), or strings representing the resources (i.e., a value string). In "DC-Simple-in-RDF", a dc:creator is a name:

```
<http://www.example.com>   dc:creator  "John Smith".
```

In "DC-Qualified-in-RDF", in contrast, a dc:creator is an entity, as in:

```
<http://www.example.com>   dc:creator  <http://www.example.org/person32>
```

or

```
<http://www.example.com>   dc:creator         _:xxx.
_:xxx                      rdf:type           foaf:Person
_:xxx                      dcrdf:valueString  "John Smith"
```

These two contrasting approaches may be pictured as follows:

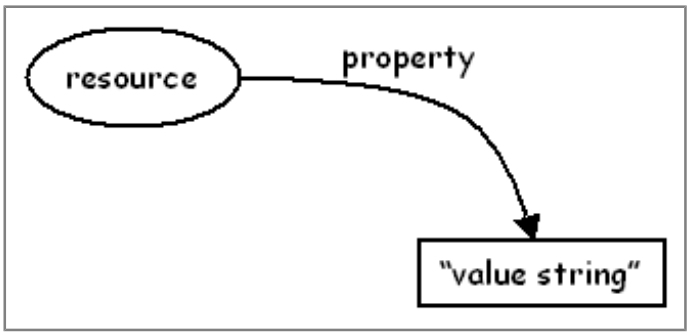

Figure 1

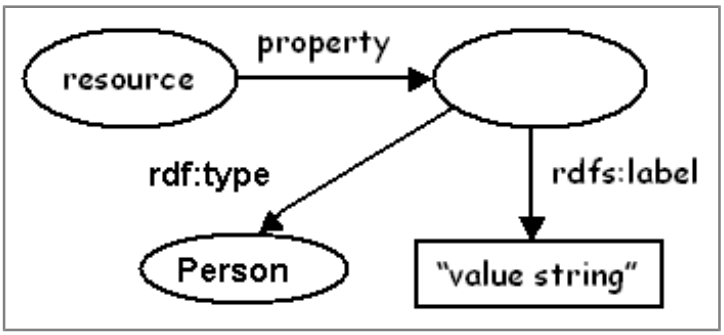

Figure 2

The current draft DC-in-RDF specification under development follows the latter approach -- dc:creator refers to an entity which can be identified (e.g., in an authority file) and described in its own right (e.g., with a name, an affiliation, and a birth date). The English-language definitions of these terms bear out this interpretation; dc:creator is "an entity primarily responsible for making the content of the resource", examples being "a person, an organization, or a service". However, the usage comments associated with these definitions also reflect the ambiguity: "Typically, the name of a Creator should be used to indicate the entity".

In accordance with the current approach, the DCMI Usage Board would assign a range of "Agent" to dc:creator and dc:contributor, where "Agent" would be defined as "the class of all things that are a Person, Organization, or Service". If it is used at all, the range "Literal" would apply only to metadata terms which are typically associated with value strings, such as dc:title. In most cases, the ranges to be defined are reasonably obvious given usage patterns in practice.

Due to the ambiguous usage of dc:creator and dc:contributor over the years, however, the assignment of any range would make one or another part of the legacy metadata appear invalid in the context of machine processing. Declaring "Agent" as the range of dc:creator will mean that inferencing applications will expect to treat the value of the dc:creator property as an entity. Where metadata represents names as literal values for dc:creator, applications will need to treat these as "special cases" in order to merge them with metadata which associate those names with the expected entity constructs.

The existing specifications from the DCMI have not taken these applications into account, which has resulted in an unknown amount of Dublin Core-based RDF data that is inconsistent with the definitions of the Dublin Core properties. The DC-RDF taskforce has judged that the mentioned changes are necessary, albeit painful, to ensure the long-term viability of Dublin Core in RDF.

References

1. Expressing Simple Dublin Core in RDF/XML, DCMI Recommendation, 2002-07-31
2. Expressing Qualified Dublin Core in RDF / XML, DCMI Proposed Recommandation, 2002-05-15
3. Expressing Dublin Core Metadata in RDF, Working Draft
4. Release Notes for above draft.

Date: Thu, 23 Mar 2006 17:21:12 -0000
Reply-To: DCMI Architecture Group <DC-ARCHITECTURE@JISCMAIL.AC.UK>
Sender: DCMI Architecture Group <DC-ARCHITECTURE@JISCMAIL.AC.UK>
From: Andy Powell <andy.powell@EDUSERV.ORG.UK>
Subject: FW: Domains and ranges of DC properties
To: DC-ARCHITECTURE@JISCMAIL.AC.UK

This topic was discussed briefly at the f2f meeting in Madrid - and has trundled on slowly behind the scenes since then, as part of the DC-RDF Taskforce activity.

The proposed list of domain and range classes are in the DC-Architecture wiki at

http://dublincore.org/architecturewiki/DCPropertyDomainsRanges

Given that this discussion is largely about semantics, we agreed at todays Usage Board teleconf, that discussion about this list would move into the remit of the UB. However, that doesn't mean that we aren't interested in people's views. So, if you have comments on the above document, please share them here.

(As an aside, we're going to attempt to bring back DC-RDF Taskforce discussion onto the main list - since the current use of a sub-list seems to have fragmented some our conversations!).

Date: Thu, 23 Mar 2006 23:44:55 +0000
From: Rachel Heery <r.heery@UKOLN.AC.UK>
Subject: Re: FW: Domains and ranges of DC properties
To: DC-ARCHITECTURE@JISCMAIL.AC.UK

On Thu, 23 Mar 2006, Andy Powell wrote:
> http://dublincore.org/architecturewiki/DCPropertyDomainsRanges
>
> Given that this discussion is largely about semantics, we agreed at
> todays Usage Board teleconf, that discussion about this list would move
> into the remit of the UB. However, that doesn't mean that we aren't
> interested in people's views. So, if you have comments on the above
> document, please share them here.

I think it would be useful to clarify to this list the underlying purpose of this exercise and the benefits. Or is this being done in a spirit of 'enabling' unknown future benefits??

As I understand it the aim of the exercise is to make precise distinctions about the "value space" of a property in machine-processable definitions. And that this is being done to indicate what can be inferred when a particular property is used in a triple.

Can someone perhaps give some use cases of how this would be beneficial??

My concern would be that the accuracy of the values in manually crafted DC metadata would not support inference. DC metadata, as I see it, is intended to provide 'cheap and cheerful' minimal level descriptive metadata. Would any application want to start making inferences over such metadata?

On the other hand if the intention is to be able to merge data from different sources on a 'partial understanding' basis I would be more convinced of the benefit...

Date: Fri, 24 Mar 2006 13:24:36 +0100
From: Mikael Nilsson <mini@NADA.KTH.SE>
Subject: Re: FW: Domains and ranges of DC properties
To: DC-ARCHITECTURE@JISCMAIL.AC.UK

tor 2006-03-23 klockan 23:44 +0000 skrev Rachel Heery:
> I think it would be useful to clarify to this list the underlying purpose
> of this exercise and the benefits.

That's very true, we do need to make clear to ourselves and the world why this is important.

> Or is this being done in a spirit of
> 'enabling' unknown future benefits??

No, certainly not. The benefits are tangible and somewhat understood by at least a few...

> As I understand it the aim of the exercise is to make precise distinctions
> about the "value space" of a property in machine-processable definitions.
> And that this is being done to indicate what can be inferred when a
> particular property is used in a triple.

It is not a case of trying to restrict the definitions, but rather to make clear what the definitions *already* imply. Some of the definitions are less than clearly formulated, for the benefit of no one... I'm pretty sure the changes introduced are not meant to be semantic changes at all.

The origin of these considerations is the DCMI Abstract Model, which finally took a position on the question of "What is a 'value' anyway?". We now know the exact distinction DCMI makes between a value and a value string. Again, I don't believe that the DCAM really introduces anything new, but the DCAM made it impossible to escape the issue of "values" by blurring distinctions. Some of that blurring is present in the definitions of the DCMI terms themselves, and we're trying to fix that.

> Can someone perhaps give some use cases of how this would be beneficial??

A reasonable request, again, even though it should be obvious that being clear and consistent are virtues on their own.

It seems clear at this point that the main beneficiary at this point in time is the Semantic Web community. The reason is simply that this community do the most advanced machine-processing of metadata.

If metadata is crafted and consumed purely by humans, precise ranges and domains are generally unnecessary, as the meaning is (often) clear given context, knowledge of the domain, and a good deal of good-will.

This is not always true, mind you. A non-native english speaker not knowing the context might have *huge* problems, for example. And there's no way a machine can help in this case.

However, when machines generate and consume metadata we cannot assume them to be that smart. You already know this, of course. But look at the following:

http://swoogle.umbc.edu/index.php?option=com_frontpage&service=relation&queryType=rel_swd_instance_range_p2c&searchString=http%3A%2F%2Fpurl.org%2Fdc%2Felements%2F1.1%2Fcreator

which is a list of the Classes used for values of dc:creator (not a list of values).

Where's the commonality? To a human, it's clear that most values (except for some anomalies like foaf:maker) are actually entities like persons or organizations. It's a fair guess that even the literal values denote such entities. But how does a piece of software know that these disparate classes denote the same *kind* of thing? The only way is through a common definition - namely the definition of dc:creator. And the way to tell a machine that "values of dc:creator are always Agents" is to give it a range. That's the whole thing, really.

This information would clearly be useful for Swoogle, as it would be able to improve its search engine based in this information.

> My concern would be that the accuracy of the values in manually crafted DC
> metadata would not support inference. DC metadata, as I see it, is
> intended to provide 'cheap and cheerful' minimal level descriptive
> metadata. Would any application want to start making inferences over such
> metadata?

All RDF data is potentially subject to inference. You won't know what people will do with your data, and in the case of RDF, the machine-semantics (the basis for inference) is specified in the RDF specs themselves.

In other words: expressing metadata in RDF makes it subject to the interpretations of *others* according to the rules of RDF. Thus, if you're producing invalid RDF manually, you're not violating the rules of DC but the rules of RDF. Maybe RDF shouldn't be used in those cases? Or software should be used to make the data abide by the RDF rules?

> On the other hand if the intention is to be able to merge data from
> different sources on a 'partial understanding' basis I would be more
> convinced of the benefit...

Well, this is the Swoogle argument. Swoogle is just a single example of a SW-based service that relies on the semantics of RDF, including domains and ranges. *Any* generic RDF service would need to rely on that information, and withholding them that information for DC properties, even though it's there in the definitions for human consumption, wouldn't be very meaningful.

I'm not sure the above helped at all, but that's my (incomplete) view of the situation. Feel free to correct me! :-)

Date: Fri, 24 Mar 2006 13:05:52 +0000
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: FW: Domains and ranges of DC properties
To: DC-ARCHITECTURE@JISCMAIL.AC.UK

Rachel Heery wrote:
> I think it would be useful to clarify to this list the underlying purpose
> of this exercise and the benefits. Or is this being done in a spirit of
> 'enabling' unknown future benefits??
>
> As I understand it the aim of the exercise is to make precise distinctions
> about the "value space" of a property in machine-processable definitions.
> And that this is being done to indicate what can be inferred when a
> particular property is used in a triple.
>
> Can someone perhaps give some use cases of how this would be beneficial??

Mikael beat me to a reply, but as I've been writing this for the last hour or so, I'll send it anyway. I think I'm repeating some of what Mikael said, but I've had a stab at a more concrete example of how the rdfs:range data might be used.

It seems to me the bottom line is, as Mikael said, that this information is _already_ part of the definition of the DCMI-owned properties, but it is currently accessible only to a human reader of the term description/definition (and even in the human-readable definitions sometimes it isn't as clear/unambiguous as it should be - though the DCMI Usage Board has gone to some pains to try to address that, I think).

Including the RDFS assertions makes the information which is already there _explicit_ and accessible to an _application_.

Why is that useful?

One attempt at a "Use Case" :

Problem: to resolve clearly the ambiguity over whether the value of dc:creator (etc) is the agent or the name of the agent.

The DCMI definition says the value is the agent. Some metadata providers follow this definition; some metadata providers interpret the definition as allowing them to use the name of the agent, a literal, as the value. (And to make matters worse, DCMI has published specifications which license both approaches.)

To an RDF application, those two data providers are "saying" two different things. The data providers think they are "saying" the same thing in two different ways.

Consider a service provider consuming metadata and working across both those sets of data that they have harvested.

Case 1: The service provider may not even be aware that these two patterns are in use. They set up a query to answer the question "Which resources were created by the agent called 'John Smith'?" and that query

only returns half the results because they are catering only for one pattern. Not very satisfactory for the user of the service. Cheap, certainly, and maybe cheerful for the service provider in the short term, but probably not for the user who only sees half the results.

Case 2: The service provider may be aware from analysing their harvested data that both these patterns are in use, but not that DCMI and the data providers expects them to be treated as equivalent. The service provider might choose to reject/ignore one of the two cases. How do they decide? Different service providers might make different choices. Again, not very satisfactory for the user of those services, who gets contradictory results. Still cheap-ish and moderately cheerful for the service provider, but the user gets less cheerful as they get different results from different services (and eventually vents their spleen on the service providers, who then aren't cheerful either).

Case 3: The service provider may be aware that both these patterns are in use, and also that DCMI and the data providers expect them to be treated as equivalent. So the provider introduces some processing - processing specific to this group of properties - which treats the two cases as the same or maps one to the other. This introduces a cost to the service provider. They can't apply the generic rule, the rule they use for all the other RDF data they have harvested: they have to introduce "special case" rule for this specific property or set of properties. And every new service provider that comes along and wants to process the data has to (a) find out that these idiosyncrasies specific to this group of properties exist (how do they do that?) and (b) implement this special-case rule in their application. And the more options left open to the data providers - not just two patterns, but three, four, five etc - the more rules the service providers have to find out about and handle in their application, and the more complex and costly those applications become. OK, the user may now be fairly cheerful, but it's neither cheap nor cheerful for our service providers, who probably decide that working with DC metadata is just more hassle than it's worth!

By introducing, say, an rdfs:range constraint for the dc:creator that says the range is a class of some:Agent, which is disjoint - "non-overlapping" - with the class rdfs:Literal, DCMI says unambiguously to my two metadata providers and - perhaps more importantly - to their two metadata creation applications that the value is the agent, not the name of the agent. It allows those two metadata creation applications to "say" the same thing when they create the metadata record.

And it says unambiguously to the service provider and to their metadata processing application that the value is the agent not the literal name of the agent. So when that application processes some harvested data, it can detect contradictions between the data and DCMI's description of the property.

In my case 2 above the best the application could do was say to the human administrator of the service "Look, I've got two (three, four five) sorts of thing as values for dc:creator here" - the Swoogle case, as Mikael pointed out - leaving the human administrator to go to read about dc:creator and try to work out what to do.

With the rdfs:range information, the application can recognise "DCMI says the range of dc:creator is some:Agent which is disjoint with rdfs:Literal, so all these sorts of thing are also instances of some:Agent, that's fine - but ooh, look, Admin-Person, this set of data has literal values and contradicts that assertion". Our providing the rdfs:range data provides the basis for more clarity and consistency. It makes the work of service providers easier, and cheaper, and hopefully the end result is more cheerfulness all round! ;-)

From: DCMI Architecture Group [mailto:DC-ARCHITECTURE@JISCMAIL.AC.UK]
On Behalf Of Andy Powell
Sent: Thursday, March 23, 2006 5:21 PM
To: DC-ARCHITECTURE@JISCMAIL.AC.UK
Subject: [DC-ARCHITECTURE] FW: Domains and ranges of DC properties

This topic was discussed briefly at the f2f meeting in Madrid - and has trundled on slowly behind the scenes since then, as part of the DC-RDF Taskforce activity.

The proposed list of domain and range classes are in the DC-Architecture wiki at

http://dublincore.org/architecturewiki/DCPropertyDomainsRanges

Given that this discussion is largely about semantics, we agreed at todays Usage Board teleconf, that discussion about this list would move into the remit of the UB. However, that doesn't mean that we aren't interested in people's views. So, if you have comments on the above document, please share them here.

(As an aside, we're going to attempt to bring back DC-RDF Taskforce discussion onto the main list - since the current use of a sub-list seems to have fragmented some our conversations!).

Date: Tue, 28 Mar 2006 15:53:26 +0100
From: Ann Apps <ann.apps@MANCHESTER.AC.UK>
Organization: University of Manchester
Subject: Re: FW: Domains and ranges of DC properties
To: DC-ARCHITECTURE@JISCMAIL.AC.UK

Andy, and All,

A few comments after a very quick scan of this list.

Location: Is this just physical and geographic locations? If so maybe it sh
ould be more explicit, ie define 'place'. Or does is include machine/intern
et location?

RightsStatement: I think one may also want to make a rights statement about
 a Service. But Service is defined as being of Agent class (which I have so
me reservations about), and so outside the scope of a rightsStatement.

Topic: I think 'subject' also needs defining. I assume it means 'keyword' o
r classification as in dc:subject. But the word 'subject' has other meaning
s, eg the 'subject' of a sentence, or a 'subject of the Queen'.

I assume that 'DigitalResource' can include Service, and that 'PhysicalReso
urce' includes Person - the definitions seem to allow that. So I could have
 a Collection of Services or a Collection of Persons.

Being really picky - can an animal be an Agent? Or does Agent imply some in
telligence or resource creation?

Should the list include 'event'? Or is that covered by ConceptualResource?

And a 'Class' is the 'class of all classes'... Right...

Date: Tue, 28 Mar 2006 17:33:31 +0100
Sender: DCMI Architecture Group <DC-ARCHITECTURE@JISCMAIL.AC.UK>
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: FW: Domains and ranges of DC properties
To: DC-ARCHITECTURE@JISCMAIL.AC.UK

Ann Apps wrote:

> Topic: I think 'subject' also needs defining. I assume it
> means 'keyword' or classification as in dc:subject. But the
> word 'subject' has other meanings, eg the 'subject' of a
> sentence, or a 'subject of the Queen'.

I guess I find the "Topic" class a bit problematic too. What is a

"Topic"? I can understand the notion of a relationship/property
capturing the notion that resource:y is a topic of resource:x

resource:x has-topic resource:y

That's what the dc:subject property does. But I'm not really clear what
the class of "Topics" is - apart from the class of things which might be
values of dc:subject! Is a Topic a "Conceptual Resource"? Or is the
Topic class broader?

What _is_ the range of dc:subject? Can a person be the value of
dc:subject? A place? An event? Certainly I can have a book which has as
its topic Che Guevara (person) or the Sierra Maestra (place) or the
Cuban Revolution (event). If all these things can be values for
dc:subject, I think a resource of any type can be a value of dc:subject,
and the range of dc:subject is just the class Resource.

If on the other hand we are saying that the value of dc:subject is
always a "conceptual resource" - so in my examples above the values
would be "the concept of the person Che Guevara", "the concept of the
place the Sierra Maestra", "the concept of the event the Cuban
Revolution", rather than the person/place/event. There may be a
person/place/event related to the concept, but they are distinct
resources. That would be a consistent approach too, I think (even if I
find it somewhat slippery at times).

I think, the introduction of the "Topic" class is suggesting the latter
approach, but I'm not sure, and I think it needs clarification. And if
it is the case then I think maybe the range of dc:subject is just the
class of Conceptual Resources, not some separate class of Topics.

> And a 'Class' is the 'class of all classes'... Right...

Those phrases are all intended to be descriptions/definitions of
classes, so it's not saying 'a Class is....'. Rather, it's saying 'the
class labelled "Class" (in practice it would have a URI) is'.... i.e.
that class is the class of all classes, or the class of resources which
are classes. It does make sense, I think.

Date: Thu, 23 Mar 2006 17:12:47 -0000
From: Andy Powell <andy.powell@EDUSERV.ORG.UK>
Subject: Domains and ranges
To: DC-USAGE@JISCMAIL.AC.UK

As discussed in today's call, I've re-worked the list of domains and
ranges to remove the use of the FRBR-like Item, Work and Manifestation.

Instead, I just use Resource, PhysicalResource and DigitalResource.
(Yes, the definitions of the latter two need some work!).

I suspect that I mainly added the FRBR stuff to cope with the use of
'item' in the 'collection description' properties. As redefined, you
can't have a collection of concepts or a collection of people, only a
collection of PhysicalResources or a collection of DigitalResources.

I think that is probably what is inteded by the DC-CD WG, but I might be
wrong.

Date: Thu, 23 Mar 2006 17:42:26 +0000
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Domains and ranges
To: DC-USAGE@JISCMAIL.AC.UK

Andy Powell wrote:

> I suspect that I mainly added the FRBR stuff to cope with the use of
> 'item' in the 'collection description' properties. As redefined, you
> can't have a collection of concepts or a collection of people, only a
> collection of PhysicalResources or a collection of DigitalResources.
>
> I think that is probably what is inteded by the DC-CD WG, but I might be
> wrong.

It was what I intended, based on what Mike Heaney says

====
Collection: An aggregation of physical and/or electronic Items.
====

and

====
Item: The concrete (incorporating physical and electronic) realisation
of Content.
Note: In so far as this analysis is concerned with collections, the
entities Content and Item will be considered only to the extent that
their types and attributes impinge upon Collection Description. In the
vast majority of cases, too, the Items will coincide with what FRBR
calls Items, not Manifestations. 'Item' has been chosen as the most
neutral term in preference to other terms which have been used such as
'Document' or 'Document-like Object'. 'Item' can most easily embrace all
of the concepts of physical and electronic, text and non-text, and human
and natural creations.
====

And I argued, therefore, that an aggregation of events (definitely) and
an aggregation of services (probably) were not collections as defined by
Heaney and were out of scope for the DC CD AP, and such classes should
not be part of the collection type vocabulary.

Others in the DC CD WG argued against that - I think really just on the
basis of "intuitive" notions of "collection" and/or on the basis that if
we were basing our collection classes on the DCMI Type list then we had
to take the whole list, rather than on a reading of Heaney - and (in a
moment of weakness) I caved in to the tyranny of democracy.

Though I have continued to think they were wrong and I should have stood
my ground, so I may go back and fight that out again in the WG ;-)

[1] http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0411&L=dc-collections&P=172
[2] http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0412&L=dc-collections&P=60

Date: Thu, 23 Mar 2006 17:55:22 +0000
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Domains and ranges
To: DC-USAGE@JISCMAIL.AC.UK

Andy Powell wrote:
> I suspect that I mainly added the FRBR stuff to cope with the use of
> 'item' in the 'collection description' properties. As redefined, you
> can't have a collection of concepts or a collection of people, only a
> collection of PhysicalResources or a collection of DigitalResources.

This also raises the interesting question of whether the new class
describes/defined in

http://dublincore.org/architecturewiki/DCPropertyDomainsRanges

as

Collection: The class of everything that is an aggregation of one or
more DigitalResources or PhysicalResources.

is the same as the existing class dcmitype:Collection ("A collection is
an aggregation of items. The term collection means that the resource is
described as a group; its parts may be separately described and navigated")

Or whether this new class is in fact a subclass of dcmitype:Collection.
I don't know the answer: it all depends on the definition of "item" in
the description/definition of dcmitype:Collection. ;-)

2006-03-23 Discussion on Usage Board telecon

   Diane: in looking through possible classes, I see that 3 out of
   4 use FRBR -- I was trying to see where these were assigned.
   Andy: I thought all these classes were used somewhere but
   need to check. Some may only be used in definitions of other
   classes - so not directly assigned. For example, "work" is
   there in order to define "manifestation". Need to double-check

which classes are actually used. -- check to make sure there are no "hanging classes" that do not get used anywhere.

Diane: Problems arise with FRBR expressions: often, "manifestations" relate to expressions, not necessarily to works. Eg, translation as an expression. Manifestation of that translation skips a level in terms of FRBR. Difficult to always distinguish btw manifestation and an Item; things can be both Manifestation and Item in the digital context. Most work on FRBR has come from a library context. Joe: in the archival community, everything is a "copy". Resources -- digital resources and physical resources -- but we do not necessarily need to talk about items and manifestations. What are the consequences about being explicit about domains and ranges? Diane: good to discuss but agree with Tom -- one step at a time. Andy: the minimal aim - if we cannot agree on actual classes - is to decide where this document is going.

UMBC Swoogle data 3/24/2006

DC property	range of value	docs	triples	rank
dc:Author	rdfs:Literal	6	84	3372.21
dc:BAD	rdfs:Literal	1	1	3372.21
dc:Contact.Email	rdf:Alt	2	2	1309.40
dc:Contributor	rdf:Bag	18	18	1191.79
dc:Creator	rdf:Bag	5	5	1191.79
dc:Creator	rdfs:Literal	3	3	3372.21
dc:Creator	rdf:Seq	1	1	2426.25
dc:Creator	http://description.org/schema/Person	1	1	0.57
dc:Creator	foaf:Person	1	1	12377.71
dc:Creator.Address	http://lsdis.cs.uga.edu/~kaarthikaddress	1	1	0.01
dc:Date	rdfs:Literal	10	92	3372.21
dc:Date	rdf:Bag	2	2	1191.79
dc:Description	rdfs:Literal	14	22	3372.21
dc:Description	http://purl.org/dc/dcmitype/InteractiveResource	1	1	0.82
dc:Description	rdf:Alt	1	1	1309.40
dc:ID	rdfs:Literal	2	30	3372.21
dc:Identifier	rdfs:Literal	6	10	3372.21
dc:Language	rdfs:Literal	102	103	3372.21
dc:Publisher	rdf:Bag	1	1	1191.79
dc:Related	rdf:Bag	22	22	1191.79
dc:Rights	rdfs:Literal	98	98	3372.21
dc:Subject	rdf:Bag	16	42	1191.79
dc:Subject	rdf:bag	1	1	1.77
dc:Title	rdfs:Literal	9	17	3372.21
dc:Title	rdf:Bag	2	2	1191.79
dc:alternative	rdfs:Literal	12	12	3372.21
dc:apply_by	rdfs:Literal	3	70	3372.21
dc:arrow	rdfs:Literal	1	50	3372.21
dc:audience	rdfs:Literal	1	1	3372.21
dc:author	rdfs:Literal	53	634	3372.21
dc:author	foaf:Person	6	188	12377.71
dc:available	rdfs:Literal	3	524	3372.21
dc:booktitle	rdfs:Literal	3	3	3372.21
dc:category	rdfs:Literal	3	22	3372.21
dc:collection	rdfs:Literal	1	8	3372.21
dc:comment	rdfs:Literal	8	69	3372.21
dc:contact	http://www.w3.org/2001/sw/EO/usecases/vocab#Person	1	1	0.21
dc:content	rdfs:Literal	1	15	3372.21
dc:contributor	rdfs:Literal	7613	70097	3372.21
dc:contributor	foaf:person	71	7777	1.97
dc:contributor	foaf:Person	52	449	12377.71
dc:contributor	rdf:Bag	26	26	1191.79
dc:contributor	http://www.imsproject.org/rdf/imsmd_rootv1p2#Entity	18	1740	0.08
dc:contributor	http://xmlns.com/wordnet/1.6/Person	9	121	173.95
dc:contributor	http://www.iwi-iuk.org/material/RDF/Schema/Class/iwi#Person	8	8	5.36
dc:contributor	http://cain.nbii.gov/schemas/biodiv_resources#Person	4	4	2.02
dc:contributor	foaf:Organization	2	2	673.94
dc:contributor	http://ilrt.org/discovery/2001/06/schemas/ical-full/hybrid.rdf#CAL-ADDRESS	2	34	0.20
dc:contributor	http://tbc.sk/RDFS/person.rdf#Person	1	22	0.04

UMBC Swoogle data 3/24/2006

dc:contributor	http://web.resource.org/cc/Agent	1	1	159.74
dc:copyright	rdfs:Literal	15	16	3372.21
dc:corp	rdfs:Literal	1	20	3372.21
dc:corp_type	rdfs:Literal	1	20	3372.21
dc:coverage	rdfs:Literal	1186	6438	3372.21
dc:coverage	dct:W3CDTF	7	7	113.91
dc:coverage	dct:Box	6	6	4.30
dc:coverage	dct:TGN	6	6	6.58
dc:coverage	dct:Point	6	6	4.28
dc:coverage	dct:ISO3166	6	6	4.22
dc:coverage	dct:Period	6	6	5.91
dc:coverage	http://www.alliknow.net/ontology#River	2	8	0.00
dc:coverage	http://www.alliknow.net/ontology#Route	2	2	0.00
dc:coverage	http://www.alliknow.net/ontology#Area	2	6	0.00
dc:coverage	http://www.alliknow.net/ontology#Sight	2	2	0.00
dc:coverage	http://www.alliknow.net/ontology#Hostel	2	2	0.00
dc:coverage	http://prismstandard.org/namespaces/pcv/1.0/Descriptor	2	2	0.19
dc:coverage	http://www.alliknow.net/ontology#Quarter	2	2	0.00
dc:coverage	http://www.alliknow.net/ontology#Canyon	2	4	0.00
dc:coverage	http://www.alliknow.net/ontology#Region	2	2	0.00
dc:coverage	http://www.alliknow.net/ontology#Ocean	2	8	0.00
dc:coverage	rdf:Bag	2	2	1191.79
dc:coverage	http://www.alliknow.net/ontology#Village	2	20	0.00
dc:coverage	http://www.alliknow.net/ontology#Canal	2	2	0.00
dc:coverage	http://www.alliknow.net/ontology#City	2	102	0.02
dc:coverage	http://www.alliknow.net/ontology#Convent	2	2	0.00
dc:coverage	http://www.alliknow.net/ontology#Country	2	28	0.01
dc:coverage	http://www.alliknow.net/ontology#Volcano	2	4	0.00
dc:coverage	http://www.alliknow.net/ontology#Ruins	2	4	0.00
dc:coverage	http://www.alliknow.net/ontology#Continent	2	6	0.00
dc:coverage	http://www.alliknow.net/ontology#Street	2	6	0.00
dc:coverage	http://www.alliknow.net/ontology#Lake	2	6	0.00
dc:coverage	http://www.alliknow.net/ontology#Sea	2	12	0.00
dc:coverage	http://www.alliknow.net/ontology#Island	2	6	0.00
dc:coverage	http://nurl.org/0/geography/map/1.0/Point	1	6	0.01
dc:coverage	http://www.kanzaki.com/ns/music#Concert	1	1	0.02
dc:coverage	http://nurl.org/0/geom2d/1.0/Point	1	4	0.01
dc:coverage	dct:spatial	1	2	24.01
dc:coverage	rdf:value	1	1	1731.73
dc:coverage	http://www.w3.org/2003/01/geo/wgs84_pos#Point	1	1	468.46
dc:coverage	http://purl.org/rss/1.0/modules/adminPeriod	1	1	0.00
dc:create	rdfs:Literal	1	1	3372.21
dc:created	rdfs:Literal	63	513	3372.21
dc:created	dct:W3C-DTF	4	4	0.03
dc:created	dct:W3CDTF	1	1	113.91
dc:creater	rdfs:Literal	32	648	3372.21
dc:creator	rdfs:Literal	234655	2477665	3372.21
dc:creator	http://web.resource.org/cc/Agent	4090	6359	159.74
dc:creator	http://xmlns.com/wordnet/1.6/Person	2714	1138250	173.95
dc:creator	foaf:Person	2281	5969	12377.71
dc:creator	foaf:Agent	1723	3234	1066.99
dc:creator	rdf:Bag	792	877	1191.79

dc:creator	http://www.w3.org/2000/10/swap/pim/contact#Person	228	243	111.14
dc:creator	rdf:Seq	126	153	2426.25
dc:creator	http://usefulinc.com/ns/doap#Project	83	83	3.02
dc:creator	http://cain.nbii.gov/schemas/biodiv_resources#Organization	62	146	3.72
dc:creator	http://cain.nbii.gov/schemas/biodiv_resources#Person	37	47	2.02
dc:creator	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#Person	22	22	1.46
dc:creator	x-urn:flickr:user	14	75	0.09
dc:creator	http://ilrt.org/discovery/2001/06/schemas/ical-full/hybrid.rdf#CAL-ADDRESS	13	336	0.20
dc:creator	x-urn:freebsd:user	9	27	0.02
dc:creator	http://xmlns.com/wot/0.1/User	7	2140	0.34
dc:creator	http://musicbrainz.org/mm/mm-2.1#Artist	6	37	0.13
dc:creator	http://www.aifb.uni-karlsruhe.de/WBS/pha/rdf-query/sample.rdf#Person	5	5	0.09
dc:creator	foaf:maker	5	7	1967.23
dc:creator	foaf:person	4	6	1.97
dc:creator	http://web.resource.org/cc/Work	4	9	2242.33
dc:creator	http://www.w3.org/2003/01/geo/wgs84_pos#SpatialThing	4	13	324.14
dc:creator	http://www.iwi-iuk.org/material/RDF/Schema/Class/iwi#Person	4	4	5.36
dc:creator	http://www.ukoln.ac.uk/metadata/education/regproj/reg/Agency	4	4	0.01
dc:creator	http://ltsc.ieee.org/2002/09/lom-base#entity	4	6	0.00
dc:creator	http://purl.org/rss/1.0/Agent	3	3	0.02
dc:creator	http://www.cs.umd.edu/projects/plus/DAML/onts/general1.0.daml#Person	3	6	0.82
dc:creator	http://xmlns.com/wot/0.1/identify	3	3	0.05
dc:creator	rdf:Alt	3	3	1309.40
dc:creator	foaf:Project	2	2	328.12
dc:creator	http://purl.org/rslp/terms#Agent	2	2	0.01
dc:creator	http://purl.org/dc/types-example#Agent	2	2	0.04
dc:creator	http://purl.org/rss/1.0/img	2	2	0.16
dc:creator	http://eulersharp.sourceforge.net/2004/04test/sampleP#Person	2	2	0.06
dc:creator	http://tbc.sk/RDFS/person.rdf#Person	2	2	0.04
dc:creator	http://eulersharp.sourceforge.net/2004/04test/sample#Person	2	2	0.11
dc:creator	http://www.alliknow.net/ontology#User	2	12	0.01
dc:creator	foaf:mbox_sha1sum	2	2	5977.51
dc:creator	http://iswc2004.semanticweb.org/ns#Author	2	347	0.07
dc:creator	http://prismstandard.org/namespaces/pcv/1.0/Descriptor	2	2	0.19
dc:creator	http://cain.nbii.gov/schemas/biodiv_resources#WeedManagementArea	2	2	0.26
dc:creator	http://xmlns.com/wordnet/1.6/Agent-3	1	7	93.44
dc:creator	http://www.kanzaki.com/ns/music#Conductor	1	1	0.01
dc:creator	foaf:Organisation	1	2	0.13
dc:creator	http://example.org/util#JoeGregorio	1	1	0.02
dc:creator	http://www.kanzaki.com/ns/music#Orchestra	1	1	0.01
dc:creator	http://web.mit.edu/simile/www/2004/02/display-config#PreferredTerm	1	2	0.00
dc:creator	foaf:name	1	1	5258.01
dc:creator	http://www.markwatson.com/index.rdf#Consultant	1	1	0.01
dc:creator	http://haystack.lcs.mit.edu/schemata/haystack#Person	1	3	0.01
dc:creator	http://redfoot.net/2005/redfoot#Kernel	1	98	0.03
dc:creator	http://www.aaronsw.com/2002/rdftools.n3#apps?Project	1	2	0.01
dc:creator	http://www.ontoweb.org/ontology/1#Person	1	1	0.34
dc:creator	http://annotation.semanticweb.org/iswc/iswc.daml#Associate_Professor	1	1	0.02
dc:creator	http://annotation.semanticweb.org/iswc/iswc.daml#PhDStudent	1	1	0.07
dc:creator	http://annotation.semanticweb.org/iswc/iswc.daml#Researcher	1	1	0.12
dc:creator	http://www.mindswap.org/2003/vegetarian.owl#Vegetarian	1	1	0.03
dc:creator.e-mail	rdfs:Literal	804	804	3372.21

dc:creator.email	rdfs:Literal	1	7	3372.21
dc:creator.name	rdfs:Literal	17	23	3372.21
dc:creator.personalname	rdfs:Literal	3	12	3372.21
dc:date	rdfs:Literal	264253	5231875	3372.21
dc:date	dct:W3CDTF	502	502	113.91
dc:date	rdf:Bag	213	213	1191.79
dc:date	dc:description	2	2	5682.88
dc:date	http://dublincore.org/2000/03/13/dcq#W3CDTF	1	2	0.17
dc:date	rdf:description	1	1	1.05
dc:date	dct:issued	1	1	2266.30
dc:date	dct:Created	1	1	0.00
dc:date	http://www.siderean.com/ISO8601#Date	1	110	0.01
dc:date-x-metadatalastmod	rdfs:Literal	3	6	3372.21
dc:date.x-unixtime	rdfs:Literal	10	90	3372.21
dc:dateSubmitted	rdfs:Literal	14	14	3372.21
dc:decription	rdfs:Literal	10	10	3372.21
dc:description	rdfs:Literal	400052	2484990	3372.21
dc:description	rdf:Alt	132	158	1309.40
dc:description	http://www.w3.org/1999/xhtmldiv	3	9	148.28
dc:description	rdf:Bag	2	2	1191.79
dc:description	http://tbc.sk/RDFS/person.rdf#Person	2	2	0.04
dc:description	foaf:br	2	10	0.00
dc:description	dc:InteractiveResource	1	2	0.01
dc:description	http://purl.org/dc/dcmitype/InteractiveResource	1	8	0.82
dc:description	foaf:Document	1	81	1915.82
dc:description.notes	rdf:Seq	2	2	2426.25
dc:email	rdfs:Literal	1	26	3372.21
dc:fecha	rdfs:Literal	1	15	3372.21
dc:format	rdfs:Literal	13183	287664	3372.21
dc:format	dct:IMT	12518	14498	63.22
dc:format	http://dublincore.org/2000/03/13/dcq#IMT	13	13	0.06
dc:format	http://www.iwi-iuk.org/material/RDF/Schema/Class/iwi#FormatObject	8	8	0.10
dc:format	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#FormatObject	8	8	0.04
dc:format	http://cain.nbii.gov/schemas/biodiv_resources#NBIIFormatTerm	7	170	0.40
dc:format	rdf:Bag	3	3	1191.79
dc:format	http://purl.org/dc/terms#IMT	1	50	0.02
dc:format	dct:extent	1	1	37.47
dc:hasPart	foaf:Group	2	4	189.52
dc:hasPart	foaf:Image	1	4	501.58
dc:hasPart	dc:Image	1	2	0.02
dc:hasPart	http://purl.org/dc/dcmitype/Image	1	1	32.08
dc:hasPart	http://www.ontoweb.org/ontology/1#Misc	1	1	0.04
dc:hasPartOf	rdfs:Literal	1	75	3372.21
dc:hasVersion	rdfs:Literal	1	1	3372.21
dc:html	rdfs:Literal	1	15	3372.21
dc:id	rdfs:Literal	4	134	3372.21
dc:identifier	rdfs:Literal	12936	148095	3372.21
dc:identifier	http://purl.org/dc/dcmitype/Text	11547	15123	55.84
dc:identifier	dct:URI	497	501	9.00
dc:identifier	rdf:Alt	353	353	1309.40
dc:identifier	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#InstitutionalHomePa	161	293	1.44
dc:identifier	http://www.iwi-iuk.org/material/RDF/Schema/Class/iwi#InstitutionalHomePage	107	112	0.85

dc:identifier	http://purl.org/rss/1.0/item	73	1412	5504.11
dc:identifier	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#HomePage	63	88	0.64
dc:identifier	http://purl.org/dc/dcmitype/StillImage	46	155	1.93
dc:identifier	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#MathNetPage	43	43	0.23
dc:identifier	http://www.iwi-iuk.org/material/RDF/Schema/Class/iwi#HomePage	30	32	0.37
dc:identifier	http://dublincore.org/2000/03/13/dcq#URI	12	12	0.09
dc:identifier	http://purl.org/rss/1.0/channel	7	7	1644.66
dc:identifier	http://www.iwi-iuk.org/material/RDF/Schema/Class/iwi#PersonalHomePage	6	10	0.81
dc:identifier	http://www.schemas-forum.org/registry/schemas/SCHEMAS/1.0/smes#NsSche	3	3	0.01
dc:identifier	rdf:Bag	3	3	1191.79
dc:identifier	http://139.91.183.30:9090/RDF/VRP/Examples/smes.rdf#NsSchema	2	2	0.01
dc:identifier	http://purl.org/dc/dcmitype/Collection	2	24	1.06
dc:identifier	http://139.91.183.30:9090/RDF/VRP/Examples/smes.rdf#ApSchema	2	2	0.01
dc:identifier	foaf:Demo	2	4	0.02
dc:identifier	rdfs:Resource	1	1	10899.26
dc:identifier	http://web.resource.org/cc/Work	1	10	2242.33
dc:identifier	http://purl.org/rss/1.0/modules/taxonomy/topic	1	3	0.06
dc:identifier	dc:Image	1	1	0.02
dc:indentifier	rdfs:Literal	6	462	3372.21
dc:institution	rdfs:Literal	2	2	3372.21
dc:isPartOf	http://sw.nokia.com/VOC-1/Vocabulary	3	74	0.17
dc:isPartOf	foaf:Project	2	6	328.12
dc:isPartOf	http://heima.olivant.fo/~styrheim/gallery/stadium/stadium.rdf#Rivalry	1	12	0.01
dc:isPartOf	http://xmlns.com/wordnet/1.6/document	1	1	0.00
dc:isPartOf	dc:InteractiveResource	1	2	0.01
dc:isPartOf	http://www.ontoweb.org/ontology/1#Book	1	2	0.05
dc:isPartOf	http://xmlns.com/wordnet/1.6/Document	1	1	139.88
dc:isReferencedBy	dc:Image	1	1	0.02
dc:isReferencedBy	dc:InteractiveResource	1	1	0.01
dc:isReferencedBy	http://xmlns.com/wordnet/1.6/Document	1	4	139.88
dc:isVersionOf	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#PersonalHomePage	17	17	0.20
dc:isVersionOf	rdfs:Literal	1	40	3372.21
dc:issued	rdfs:Literal	5	44	3372.21
dc:journal	rdfs:Literal	1	1	3372.21
dc:keyword	rdfs:Literal	12	154	3372.21
dc:keywords	rdfs:Literal	1	4	3372.21
dc:label	rdfs:Literal	2	70	3372.21
dc:lang	rdfs:Literal	15	55	3372.21
dc:language	rdfs:Literal	207382	341020	3372.21
dc:language	dct:RFC1766	12775	15395	62.23
dc:language	dct:RFC3066	498	498	6.57
dc:language	http://dublincore.org/2000/03/13/dcq#RFC1766	13	13	0.61
dc:language	dct:ISO639-2	11	11	4.20
dc:language	http://purl.org/dc/terms#RFC1766	1	47	0.02
dc:language	rdf:Bag	1	1	1191.79
dc:licence	http://web.resource.org/cc/Licence	6	4239	0.00
dc:link	rdfs:Literal	5	59	3372.21
dc:maintainer	foaf:Person	1	2	12377.71
dc:mediator	http://web.resource.org/cc/Agent	11	11	159.74
dc:mediator	rdfs:Literal	3	3	3372.21
dc:modified	rdfs:Literal	22	658	3372.21
dc:modified	dct:W3C-DTF	4	4	0.03

dc:name	rdfs:Literal	3	43	3372.21
dc:point	rdfs:Literal	1	50	3372.21
dc:product	rdfs:Literal	1	20	3372.21
dc:pubdate	rdfs:Literal	3	3	3372.21
dc:publisher	rdfs:Literal	43751	173864	3372.21
dc:publisher	http://www.hackcraft.net/bookrdf/vocab/0_1/Publisher	16369	16369	85.30
dc:publisher	foaf:Person	653	654	12377.71
dc:publisher	rdf:Seq	477	477	2426.25
dc:publisher	http://www.w3.org/2000/10/swap/pim/contact#Person	203	203	111.14
dc:publisher	http://cain.nbii.gov/schemas/biodiv_resources#Organization	68	158	3.72
dc:publisher	http://web.resource.org/cc/Agent	28	40	159.74
dc:publisher	foaf:Organization	17	21	673.94
dc:publisher	http://example.org/iemsr/terms/Agency	14	20	0.03
dc:publisher	http://purl.org/ws-mmi-dc/terms/Agency	9	21	0.04
dc:publisher	http://cain.nbii.gov/schemas/biodiv_resources#Person	4	4	2.02
dc:publisher	http://www.edutella.org/bibtex#Organization	3	10	0.01
dc:publisher	rdf:Bag	3	3	1191.79
dc:publisher	http://web.resource.org/cc/Work	3	6	2242.33
dc:publisher	rdf:Alt	3	3	1309.40
dc:publisher	http://purl.org/ws-mmi-dc/Agency	3	9	0.01
dc:publisher	http://cain.nbii.gov/schemas/biodiv_resources#WeedManagementArea	2	2	0.26
dc:publisher	http://purl.org/rslp/terms#Agent	2	2	0.01
dc:publisher	dc:title	2	2	23387.65
dc:publisher	http://example.org/dcap/Agency	1	1	0.00
dc:publisher	http://www.ukoln.ac.uk/metadata/education/regproj/reg/Agency	1	1	0.01
dc:publisher	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#Organization	1	1	2.12
dc:publisher	foaf:Organisation	1	1	0.13
dc:publisher	http://heima.olivant.fo/~styrheim/gallery/stadium/stadium.rdf#Team	1	8	0.01
dc:publisher.corporatenam	rdfs:Literal	3	12	3372.21
dc:publisher.corporatenam	rdfs:Literal	3	12	3372.21
dc:publishingDate	rdfs:Literal	2	2	3372.21
dc:rank	rdfs:Literal	1	50	3372.21
dc:references	http://heima.olivant.fo/~styrheim/gallery/stadium/stadium.rdf#Team	1	1	0.01
dc:references	http://purl.org/dc/dcmitype/InteractiveResource	1	1	0.82
dc:references	http://purl.org/dc/dcmitype/Image	1	1	32.08
dc:relation	rdfs:Literal	128	8388	3372.21
dc:relation	rdf:Bag	51	53	1191.79
dc:relation	foaf:chatEvent	48	131	78.62
dc:relation	foaf:Document	27	86	1915.82
dc:relation	http://www.rossettiarchive.org/schema#figure	15	400	0.11
dc:relation	rdfs:Class	7	198	12949.78
dc:relation	dc:description	2	4	5682.88
dc:relation	http://heima.olivant.fo/~styrheim/gallery/stadium/stadium.rdf#Rivalry	1	6	0.01
dc:relation	http://xmlns.com/wordnet/1.6/document	1	2	0.00
dc:relation	http://heima.olivant.fo/~styrheim/gallery/stadium/stadium.rdf#Team	1	13	0.01
dc:relation	http://purl.org/rss/1.0/item	1	16	5504.11
dc:relation	dc:Image	1	1	0.02
dc:relation	http://www.mathematik-21.de/03KRM1FR/math.shtml.rdf	1	10	0.04
dc:relation	http://www.mathematik-21.de/03GR7ZIB-7/math.shtml.rdf	1	2	0.03
dc:relation	http://www.mathematik-21.de/03GR7ZIB-7/apply.shtml.rdf	1	4	0.03
dc:relation	http://www.mathematik-21.de/03KRM1FR/apply.shtml.rdf	1	8	0.03
dc:relation	http://www.mathematik-21.de/03GRMIB4/science.shtml.rdf	1	2	0.02

UMBC Swoogle data 3/24/2006

dc:relation	http://purl.org/dc/dcmitype/Image	1	3	32.08
dc:relation	http://xmlns.com/wordnet/1.6/Document	1	1	139.88
dc:relation.IsChildOf	dc:description	2	2	5682.88
dc:relation.IsChildOf	rdf:Seq	2	2	2426.25
dc:relation.Requires	rdf:Bag	2	2	1191.79
dc:relation.requires	dc:description	2	2	5682.88
dc:relationship	dc:relationship.References	2	2	0.00
dc:relationship	rdf:Bag	2	2	1191.79
dc:relationship.references	rdf:Bag	2	2	1191.79
dc:relationship.requires	dc:description	2	14	5682.88
dc:relationship.requires	rdf:Bag	2	2	1191.79
dc:replaces	rdfs:Literal	1	1	3372.21
dc:resource	rdfs:Literal	332	1268	3372.21
dc:right	rdfs:Literal	1	8	3372.21
dc:rights	rdfs:Literal	45368	153451	3372.21
dc:rights	http://web.resource.org/cc/Agent	8652	17275	159.74
dc:rights	foaf:Person	5282	12691	12377.71
dc:rights	http://web.resource.org/cc/license	649	649	856.47
dc:rights	foaf:Agent	600	1200	1066.99
dc:rights	http://xmlns.com/wot/0.1/User	7	2140	0.34
dc:rights	rdf:Bag	6	7	1191.79
dc:rights	rdf:Alt	5	5	1309.40
dc:rights	foaf:Document	4	4	1915.82
dc:rights	http://purl.org/rss/1.0/Agent	3	3	0.02
dc:rights	http://web.resource.org/cc/Work	3	6	2242.33
dc:rights	http://web.resource.org/cc/License	2	20	855.42
dc:rights	foaf:person	1	1	1.97
dc:rights	http://purl.org/rss/1.0/channel	1	1	1644.66
dc:rights	http://purl.org/rss/1.0/item	1	1	5504.11
dc:rightsHolder	dc:InteractiveResource	1	1	0.01
dc:rightsHolder	foaf:Person	1	3	12377.71
dc:section	rdfs:Literal	1	3	3372.21
dc:size	rdfs:Literal	7	14	3372.21
dc:source	rdfs:Literal	3390	86949	3372.21
dc:source	http://web.resource.org/cc/Work	1925	1925	2242.33
dc:source	http://simile.mit.edu/2003/10/ontologies/artstor#Collection	18	1608	0.14
dc:source	http://www.w3.org/2002/07/owl#Thing	11	16	1287.58
dc:source	foaf:Document	6	167	1915.82
dc:source	http://www.w3.org/2001/02pd/rec54#WD	2	2	1.81
dc:source	rdf:Bag	2	2	1191.79
dc:source	http://purl.org/dc/dcmitype/Image	1	2	32.08
dc:source	http://xmlns.com/wordnet/1.6/Document	1	1	139.88
dc:source	http://www.w3.org/2002/07/owl#Ontology	1	1	1547.73
dc:source	http://purl.org/rss/1.0/channel	1	20	1644.66
dc:source	http://purl.org/rss/1.0/item	1	2	5504.11
dc:source	http://xmlns.com/wordnet/1.6/document	1	6	0.00
dc:source	dc:Image	1	1	0.02
dc:source	http://www.aktors.org/ontology/portal#Airport	1	1	0.04
dc:source	http://www.aktors.org/ontology/portal#Coordinate-Location	1	1	0.14
dc:spatial	http://www.cyc.com/2004/06/04/cyc#SpatialThing-Localized	59	1254	0.35
dc:spatial	http://www.w3.org/2003/01/geo/wgs84_pos#SpatialThing	2	2	324.14
dc:sqndsqgy	rdfs:Literal	1	1	3372.21

dc:subject	rdfs:Literal	173686	2941474	3372.21
dc:subject	rdf:Bag	544	4296	1191.79
dc:subject	http://cain.nbii.gov/schemas/biodiv_resources#ISMTterm	74	312	15.76
dc:subject	http://cain.nbii.gov/schemas/biodiv_resources#Taxon	35	202	1.22
dc:subject	http://www.w3.org/2004/02/skos/core#ConceptScheme	32	36	0.50
dc:subject	http://www.w3.org/2004/02/skos/core#Concept	31	5207	5.35
dc:subject	http://www.w3.org/2002/07/owl#Class	26	130	4250.11
dc:subject	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#MSC2000	25	99	0.83
dc:subject	rdf:Seq	22	38	2426.25
dc:subject	http://www.iwi-iuk.org/material/RDF/Schema/Class/iwi#Descriptor	20	75	0.16
dc:subject	rdfs:Class	20	83	12949.78
dc:subject	http://protege.stanford.edu/swbp/books2.owl#Lion	9	9	0.40
dc:subject	http://web.resource.org/cc/Work	9	9	2242.33
dc:subject	http://protege.stanford.edu/swbp/books1.owl#AnimalClass	9	18	0.44
dc:subject	http://protege.stanford.edu/swbp/books3.owl#Subject	9	18	0.44
dc:subject	http://protege.stanford.edu/swbp/books2.owl#AfricanLion	9	9	0.30
dc:subject	x-urn:flickr:tag	8	32	0.09
dc:subject	dct:LCSH	7	7	4.27
dc:subject	dct:LCC	6	6	4.30
dc:subject	dct:UDC	6	6	4.22
dc:subject	dct:MESH	6	6	4.34
dc:subject	dct:DDC	6	6	4.34
dc:subject	http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class/mn#Descriptor	5	9	0.13
dc:subject	foaf:Document	4	8	1915.82
dc:subject	http://protege.stanford.edu/books3#Subject	3	6	0.15
dc:subject	http://protege.stanford.edu/books2#Lion	3	3	0.13
dc:subject	http://protege.stanford.edu/books2#AfricanLion	3	3	0.10
dc:subject	http://protege.stanford.edu/books1#AnimalClass	3	6	0.17
dc:subject	http://a.com/ontology#Lion	2	2	0.22
dc:subject	http://a.com/ontology#AnimalClass	2	3	0.15
dc:subject	http://purl.org/rss/1.0/a	2	45	0.43
dc:subject	http://www.learndirect-advice.co.uk/provider/standardsandclassifications/classp	2	6	0.01
dc:subject	http://prismstandard.org/namespaces/pcv/1.0/Descriptor	2	2	0.19
dc:subject	http://xmlns.com/wordnet/1.6/Category	2	60	0.16
dc:subject	http://a.com/ontology#Subject	2	3	0.16
dc:subject	http://www.w3.org/TR/2004/WD-swbp-classes-as-values-20040721/books2.n3A	1	1	0.03
dc:subject	http://www.w3.org/TR/2005/NOTE-swbp-classes-as-values-20050405/books3.n	1	2	0.05
dc:subject	http://mirrors.webthing.com/view=Asis-Links/www.w3.org/TR/2004/WD-swbp-cl	1	2	0.05
dc:subject	http://www.mathematik-21.de/03BA7CO1-2/apply.shtml.rdf	1	2	0.02
dc:subject	foaf:Person	1	1	12377.71
dc:subject	http://mirrors.webthing.com/view=Asis-Links/www.w3.org/TR/2004/WD-swbp-cl	1	2	0.06
dc:subject	http://www.siderean.com/Z3919#Descriptor	1	286	0.01
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Jun/att-0031/books3.n3Sub	1	2	0.05
dc:subject	http://www.w3.org/TR/swbp-classes-as-values/books2.n3Lion	1	1	0.04
dc:subject	http://www.w3.org/2001/sw/BestPractices/OEP/ClassesAsValues-20040623/boo	1	1	0.04
dc:subject	dc:Image	1	1	0.02
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Jun/att-0031/books2.n3Afri	1	1	0.03
dc:subject	http://www.w3.org/TR/swbp-classes-as-values/books2.n3AfricanLion	1	1	0.03
dc:subject	http://www.w3.org/2001/sw/BestPractices/OEP/ClassesAsValues-20040623/boo	1	1	0.03
dc:subject	http://protege.stanford.edu/swbp/books1.n3AnimalClass	1	2	0.06
dc:subject	dc:InteractiveResource	1	1	0.01
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Jun/att-0031/books2.n3Lion	1	1	0.04

dc:subject	http://www.w3.org/2001/sw/BestPractices/OEP/ClassesAsValues-20040623/boo	1	2	0.06
dc:subject	http://www.w3.org/TR/2004/WD-swbp-classes-as-values-20040721/books1.n3A	1	2	0.06
dc:subject	http://protege.stanford.edu/swbp/books3.n3Subject	1	2	0.05
dc:subject	http://www.mathematik-21.de/03BA7CO1-2/science.shtml.rdf	1	1	0.02
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Jun/att-0031/books1.n3Ani	1	2	0.06
dc:subject	http://www.w3.org/TR/2004/WD-swbp-classes-as-values-20040721/books3.n3S	1	2	0.05
dc:subject	http://www.mathematik-21.de/03VEM1H1/math.shtml.rdf	1	4	0.03
dc:subject	http://smi-web.stanford.edu/people/noy/ClassesAsValues/books3.n3Subject	1	2	0.05
dc:subject	http://www.mathematik-21.de/03RONAVN/science.shtml.rdf	1	3	0.02
dc:subject	http://www.w3.org/2001/sw/BestPractices/OEP/ClassesAsValues-20040623/boo	1	2	0.05
dc:subject	http://purl.org/rss/1.0/marquee	1	3	0.01
dc:subject	http://www.mathematik-21.de/03STM5B4/math.shtml.rdf	1	1	0.01
dc:subject	http://smi-web.stanford.edu/people/noy/ClassesAsValues/books2.n3AfricanLion	1	1	0.03
dc:subject	http://www.mathematik-21.de/03MONJB1/math.shtml.rdf	1	2	0.02
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Jun/att-0007/books3.n3Sub	1	2	0.05
dc:subject	http://www.w3.org/2002/04/squish-qbe#Literal	1	1	0.06
dc:subject	http://www.mathematik-21.de/03STM5B4/apply.shtml.rdf	1	1	0.01
dc:subject	http://smi-web.stanford.edu/people/noy/ClassesAsValues/books2.n3Lion	1	1	0.04
dc:subject	http://www.mathematik-21.de/03MONJB1/apply.shtml.rdf	1	5	0.02
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Jun/att-0007/books2.n3Afri	1	1	0.03
dc:subject	http://heima.olivant.fo/~styrheim/gallery/stadium/stadium.rdf#Rivalry	1	6	0.01
dc:subject	http://www.mathematik-21.de/03MAM5DA/math.shtml.rdf	1	2	0.04
dc:subject	http://www.w3c.org/TR/2004/WD-swbp-classes-as-values-20040721/books3.n3	1	2	0.05
dc:subject	http://www.mathematik-21.de/03KRM1FR/math.shtml.rdf	1	6	0.04
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Jun/att-0007/books2.n3Lion	1	1	0.04
dc:subject	http://heima.olivant.fo/~styrheim/gallery/stadium/stadium.rdf#Team	1	21	0.01
dc:subject	rdf:Alt	1	18	1309.40
dc:subject	http://www.w3.org/2001/sw/BestPractices/OEP/ClassesAsValues-20040719/boo	1	2	0.05
dc:subject	http://www.w3c.org/TR/2004/WD-swbp-classes-as-values-20040721/books2.n3	1	1	0.03
dc:subject	http://www.mathematik-21.de/03GRMIB4/math.shtml.rdf	1	1	0.01
dc:subject	http://protege.stanford.edu/swbp/books2.n3AfricanLion	1	1	0.03
dc:subject	http://xmlns.com/wordnet/1.6/Document	1	1	139.88
dc:subject	http://www.w3c.org/TR/2004/WD-swbp-classes-as-values-20040721/books2.n3	1	1	0.04
dc:subject	http://www.mathematik-21.de/03GR7ZIB-7/apply.shtml.rdf	1	2	0.03
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Jun/att-0007/books1.n3Ani	1	2	0.06
dc:subject	http://protege.stanford.edu/swbp/books2.n3Lion	1	1	0.04
dc:subject	http://www.mathematik-21.de/03GRMIB4/science.shtml.rdf	1	2	0.02
dc:subject	http://purl.org/gem/qualifiers/NASA	1	4	0.01
dc:subject	http://www.mathematik-21.de/03KRM1FR/science.shtml.rdf	1	1	0.02
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Apr/att-0091/books3.n3Sub	1	1	0.09
dc:subject	http://purl.org/gem/qualifiers/ERIC	1	3	0.01
dc:subject	http://www.mathematik-21.de/03ZEM1M1/science.shtml.rdf	1	2	0.02
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Apr/att-0091/books2.n3Lion	1	1	0.06
dc:subject	http://www.w3c.org/TR/2004/WD-swbp-classes-as-values-20040721/books1.n3	1	2	0.06
dc:subject	http://www.mathematik-21.de/03KRM1FR/apply.shtml.rdf	1	3	0.03
dc:subject	http://lists.w3.org/Archives/Public/www-archive/2004Apr/att-0091/books1.n3Ani	1	1	0.10
dc:subject	http://www.w3.org/TR/2005/NOTE-swbp-classes-as-values-20050405/books2.n	1	1	0.04
dc:subject	http://www.w3.org/TR/2004/WD-swbp-classes-as-values-20040721/books2.n3L	1	1	0.04
dc:subject	http://www.w3.org/TR/2005/NOTE-swbp-classes-as-values-20050405/books2.n	1	1	0.03
dc:subject.keyword	rdf:Bag	2	2	1191.79
dc:thumbnail	foaf:Image	15	46	501.58
dc:thumbnail	http://xmlns.com/wordnet/1.6/Document	1	6	139.88

UMBC Swoogle data 3/24/2006

dc:title	rdfs:Literal	433091	843259	3372.21
dc:title	rdf:Seq	191	191	2426.25
dc:title	rdf:Alt	67	137	1309.40
dc:title	http://www.w3.org/2000/10/XMLSchema#string	1	93	3.83
dc:title.alternative	rdfs:Literal	3	12	3372.21
dc:title.main	rdfs:Literal	3	12	3372.21
dc:type	rdfs:Literal	3903	62182	3372.21
dc:type	dct:DCMIType	970	970	19.60
dc:type	dct:IMT	477	477	63.22
dc:type	http://purl.org/dc/dcmitype/DCMIType	214	214	1.44
dc:type	rdf:Bag	204	204	1191.79
dc:type	http://cain.nbii.gov/schemas/biodiv_resources#NBIITypeTerm	18	102	0.90
dc:type	http://dublincore.org/2000/03/13/dcq#Text	13	13	0.05
dc:type	http://purl.org/gem/qualifiers/GemType	5	9	0.03
dc:type	rdfs:Class	5	180	12949.78
dc:type	dct:Text	3	3	0.41
dc:type	http://daml.umbc.edu/ontologies/webofbelief/bibtex-ext#Category	2	38	0.00
dc:type	http://purl.oclc.org/gem/gemtype/EducatorsGuide	2	2	0.01
dc:type	http://purl.oclc.org/gem/gemtype/LessonPlan	2	2	0.01
dc:type	http://purl.oclc.org/gem/gemtype/PrimarySource	2	2	0.01
dc:type	http://purl.org/dc/dcmitype/Collection	1	1	1.06
dc:type	http://web.resource.org/cc/Agent	1	1	159.74
dc:type	http://purl.oclc.org/gem/gemtype/ImageSet	1	1	0.00
dc:type	http://purl.oclc.org/gem/gemtype/Activity	1	1	0.00
dc:type	http://purl.oclc.org/gem/gemtype/Curriculum	1	2	0.01
dc:type	http://purl.oclc.org/gem/gemtype/Reference	1	1	0.00
dc:type	http://purl.oclc.org/gem/gemtype/UnitOfInstruction	1	1	0.00
dc:type	http://purl.org/gem/qualifiers/DCType	1	1	0.00
dc:uri	rdfs:Literal	2	32	3372.21
dc:value	rdfs:Literal	4	62	3372.21
dc:version	rdfs:Literal	2	6	3372.21
abbreviations				
dc	http://purl.org/dc/elements/1.1/			
rdfs	http://www.w3.org/2000/01/rdf-schema#			
rdfs	http://www.w3.org/1999/02/22-rdf-syntax-ns#			
dct:	http://purl.org/dc/terms/			
foaf:	http://xmlns.com/foaf/0.1/			

Title: Finalizing DCMI Type Vocabulary after the comment period
Identifier: /admin/www/usage/meetings/2006/04/seattle/dcmitype/
Created: 2006-03-30

Shepherd: Stuart

Required reading
[1] /usage/public-comment/2005/12/type-vocabulary-changes/
[2] /usage/public-comment/2006/03/type-vocabulary-comments/

See also
[3] /usage/meetings/2006/04/seattle/domains-ranges/

A public comment period was held from 8 Dec to 31 Jan on changes to the DCMI Type Vocabulary [1]. The responses have been summarized in tabular form [2].

The finalization of the DCMI Type Vocabulary changes depends on resolution of the "stylistic" issue in the topic "DCMI property domains and ranges" [3].

Title: Public Comment on Changes to the DCMI Type Vocabulary
Identifier: http://dublincore.org/usage/public-comment/2005/12/type-vocabulary-changes/
Date: 2005-12-08

The DCMI Usage Board proposes the attached changes (see below) to definitions and comments of terms of the DCMI Type Vocabulary [1,2]. The changes, which were discussed at recent meetings of the Usage Board [3,4], are essentially editorial in nature:

1) Definitions were wordsmithed for grammatical correctness and stylistic consistency.

2) As they were originally written, the definitions of DCMI Type terms included information about expected uses of the term -- information which, for other types of DCMI terms, is typically provided in a separate Comment. For stylistic consistency, therefore, parts of the current Definitions were split off into Comments.

It is worth noting that the Usage Board did not undertake any more fundamental discussion of the typology offered in the DCMI Type Vocabulary. As clarified in a posting to DC-GENERAL in December 2003 [5], the Usage Board does not intend to expand or otherwise clarify the existing vocabulary; rather, it "prefers to focus on interoperability between the DCMI Type Vocabulary and other, more granular vocabularies created and maintained by communities of interest for specialized uses."

Public comment is therefore solicited primarily not on the concepts underlying the DCMI Types, but on the potential impact of the these editorial changes on existing interpretations and implementations.

Public Comment will be open from 12 December 2005 to 31 January 2006.

[1] http://dublincore.org/documents/dcmi-type-vocabulary/
[2] http://dublincore.org/documents/dcmi-terms/
[3] http://dublincore.org/usage/meetings/2005/05/washdc/topic-type-vocabulary/
[4] http://dublincore.org/usage/meetings/2005/09/madrid/type-vocabulary/
[5] http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0312&L=dc-general&T=0&F=&S=&X=3E9FF4152D1B3DB3DF&Y=tbaker%40tbaker.de&P=444

URI: http://purl.org/dc/dcmitype/Collection
Label: Collection

OLD> Definition: A collection is an aggregation of items. The term
OLD> collection means that the resource is
OLD> described as a group; its parts may be
OLD> separately described and navigated.

NEW> Definition: An aggregation of items.
NEW> Comment: The term Collection means that the resource
NEW> is described as a group; its parts may also
NEW> be separately described and navigated.

URI: http://purl.org/dc/dcmitype/Dataset
Label: Dataset

OLD> Definition: A dataset is information encoded in a
OLD> defined structure (for example, lists,
OLD> tables, and databases), intended to be
OLD> useful for direct machine processing.

NEW> Definition: Information encoded in a defined structure (for example,
NEW> lists, tables, and databases).
NEW> Comment: A Dataset is intended to be useful for direct machine
NEW> processing.

URI: http://purl.org/dc/dcmitype/Event
Label: Event

```
    OLD> Definition:     An event is a non-persistent, time-based
    OLD>                 occurrence.  Metadata for an event provides
    OLD>                 descriptive information that is the basis
    OLD>                 for discovery of the purpose, location,
    OLD>                 duration, responsible agents, and links to
    OLD>                 related events and resources. The resource
    OLD>                 of type event may not be retrievable if
    OLD>                 the described instantiation has expired
    OLD>                 or is yet to occur. Examples - exhibition,
    OLD>                 web-cast, conference, workshop, open-day,
    OLD>                 performance, battle, trial, wedding,
    OLD>                 tea-party, conflagration.

    NEW> Definition:     A non-persistent, time-based occurrence.
    NEW>
    NEW> Comment:        Metadata for an Event provides descriptive
    NEW>                 information that is the basis for discovery of the
    NEW>                 purpose, location, duration, and responsible
    NEW>                 agents associated with an event. Examples
    NEW>                 include an exhibition, webcast, conference, workshop,
    NEW>                 open day, performance, battle, trial, wedding, tea
    NEW>                 party, conflagration.

-------------------------------------------------------------------------
URI:            http://purl.org/dc/dcmitype/Image
Label:          Image

    OLD> Definition:     An image is a primarily symbolic visual
    OLD>                 representation other than text. For example -
    OLD>                 images and photographs of physical objects,
    OLD>                 paintings, prints, drawings, other images and
    OLD>                 graphics, animations and moving pictures,
    OLD>                 film, diagrams, maps, musical notation.
    OLD>                 Note that image may include both electronic
    OLD>                 and physical representations.

    NEW> Definition:     A primarily symbolic visual representation other than text.
    NEW> Comment:        Examples include images and photographs
    NEW>                 of physical objects, paintings, prints, drawings, other
    NEW>                 images and graphics, animations and moving
    NEW>                 pictures, film, diagrams, maps, musical
    NEW>                 notation.  Note that Image may include both
    NEW>                 electronic and physical representations.

-------------------------------------------------------------------------
URI:            http://purl.org/dc/dcmitype/InteractiveResource
Label:          Interactive Resource

    OLD> Definition:     An interactive resource is a resource
    OLD>                 which requires interaction from the user
    OLD>                 to be understood, executed, or experienced.
    OLD>                 For example - forms on web pages, applets,
    OLD>                 multimedia learning objects, chat services,
    OLD>                 virtual reality.

    NEW> Definition:     A resource which requires interaction from the user to
    NEW>                 be understood, executed, or experienced.
    NEW> Comment:        Examples include forms on Web pages, applets, multimedia
    NEW>                 learning objects, chat services, virtual reality
    NEW>                 environments.

-------------------------------------------------------------------------
URI:            http://purl.org/dc/dcmitype/MovingImage
Label:          Moving Image

    OLD> Definition:     A series of visual representations that,
    OLD>                 when shown in succession, impart an
    OLD>                 impression of motion. Examples of moving
    OLD>                 images are: animations, movies, television
    OLD>                 programs, videos, zoetropes, or visual
    OLD>                 output from a simulation.
    OLD> Comment:        Instances of the type "Moving Image" must
    OLD>                 also be describable as instances of the
    OLD>                 broader type "Image".

    NEW> Definition:     A series of visual representations that, when shown in
    NEW>                 succession, impart an impression of motion.
    NEW> Comment:        Examples include animations, movies, television programs,
    NEW>                 videos, zoetropes, or visual output from a simulation.
    NEW>                 Instances of the type Moving Image must also be describable
    NEW>                 as instances of the broader type Image.

-------------------------------------------------------------------------
URI:            http://purl.org/dc/dcmitype/PhysicalObject
Label:          Physical Object

    OLD> Definition:     An inanimate, three-dimensional object or
    OLD>                 substance. For example -- a computer,
    OLD>                 the great pyramid, a sculpture. Note that
    OLD>                 digital representations of, or surrogates
    OLD>                 for, these things should use Image, Text
    OLD>                 or one of the other types.

    NEW> Definition:     An inanimate, three-dimensional object or substance.
    NEW> Comment:        Note that digital representations of, or surrogates for,
    NEW>                 these objects should use Image, Text or one of the other
    NEW>                 types.
```

URI: http://purl.org/dc/dcmitype/Service
Label: Service

OLD> Definition: A service is a system that provides one or
OLD> more functions of value to the
OLD> end-user. Examples include: a photocopying
OLD> service, a banking service, an authentication
OLD> service, interlibrary loans, a Z39.50 or
OLD> Web server.

NEW> Definition: A system that provides one or more functions of value to
NEW> the end user.
NEW> Comment: Examples include a photocopying service, a banking service,
NEW> an authentication service, interlibrary loans, a Z39.50 or
NEW> Web server.

URI: http://purl.org/dc/dcmitype/Software
Label: Software

OLD> Definition: Software is a computer program in source
OLD> or compiled form which may be available for
OLD> installation non-transiently on another
OLD> machine. For software which exists only
OLD> to create an interactive environment,
OLD> use interactive instead.

NEW> Definition: A computer program in source or compiled form
NEW> which may be available for installation nontransiently
NEW> on another machine.
NEW> Comment: Examples include a C source file, MS-Windows
NEW> .exe executable, or Perl script.

URI: http://purl.org/dc/dcmitype/Sound
Label: Sound

OLD> Definition: A sound is a resource whose content is
OLD> primarily intended to be rendered as
OLD> audio. For example - a music playback file
OLD> format, an audio compact disc, and recorded
OLD> speech or sounds.

NEW> Definition: A resource whose content is primarily intended to be
NEW> rendered as audio.
NEW> Comment: Examples include a music playback file format, an audio
NEW> compact disc, and recorded speech or sounds.

URI: http://purl.org/dc/dcmitype/StillImage
Label: Still Image

OLD> Definition: A static visual representation. Examples
OLD> of still images are: paintings, drawings,
OLD> graphic designs, plans and maps.
OLD> Comment: Recommended best practice is to assign the
OLD> type "text" to images of textual
OLD> materials. Instances of the type "Still
OLD> Image" must also be describable as instances
OLD> of the broader type "Image". Narrower Than:
OLD> http://purl.org/dc/dcmitype/Image

NEW> Definition: A static visual representation.
NEW> Comment: Examples include paintings, drawings,
NEW> graphic designs, plans and maps. Recommended
NEW> best practice is to assign the type Text to
NEW> images of textual materials. Instances of the
NEW> type Still Image must also be describable as
NEW> instances of the broader type Image.

URI: http://purl.org/dc/dcmitype/Text
Label: Text

OLD> Definition: A text is a resource whose content is
OLD> primarily words for reading. For example
OLD> - books, letters, dissertations, poems,
OLD> newspapers, articles, archives of mailing
OLD> lists. Note that facsimiles or images of
OLD> texts are still of the genre text.

NEW> Definition: A resource whose content is primarily words for reading.
NEW> Comment: Examples include books, letters, dissertations,
NEW> poems, newspapers, articles, archives of mailing
NEW> lists. Note that facsimiles or images of
NEW> texts are still of the genre Text.

Title: Summary of public comments on editorial changes to the DCMI Type Vocabulary
Identifier: http://dublincore.org/usage/public-comment/2006/03/type-vocabulary-comments/
Description: This document summarizes the comments made in a public comment period on changes to the DCMI Type Vocabulary from 8 December 2005 to 31 January 2006 (see http://dublincore.org/usage/public-comment/2005/12/type-vocabulary-changes/).
Created: 2006-03-12

Term	Proposed Language	Public Comment
Collection	An aggregation of items. *Comment:* The term Collection means that the resource is described as a group; its parts may also be separately described and navigated.	*Gregory Renaud:* A resource which is [a]n aggregation of items. *Comment:* The term Collection means that the resource is described as a group; its parts may also be separately described and navigated. Properties of a Collection may not be identical to properties of its parts; properties of all parts may not be identical.
Dataset	Information encoded in a defined structure (for example,lists, tables, and databases).	*Gregory Renaud:* A resource in which data is encoded in a defined structure.
Event	A non-persistent, time-based occurrence.	*Gregory Renaud:* A resource which is a time-based activity.
Image	A primarily symbolic visual representation other than text.	*Gregory Renaud:* A resource which is a visual, non-textual representation.
Moving Image	A series of visual representations that, when shown in succession, impart an impression of motion.	*Gregory Renaud:* A resource which is a visual representation other than text that imparts motion or an impression of motion.
Physical Object	An inanimate, three-dimensional object or substance.	*Gregory Renaud:* A resource which has three dimensions and mass.
Service	A service is a system that provides one or more functions of value to the end-user.	*Ann Apps:* A service is a system that provides one or more functions. *Vladimir Makarov:* A service is a system that provides one or more functions of value to [an] actor. *Gregory Renaud:* A resource which is a system or activity that is needed or requested.
Software	A computer program in source or compiled form which may be available for installation nontransiently on another machine.	*Gregory Renaud:* A resource which is expressed as a computer program in source or compiled form; may be available for installation nontransiently on another machine. *Misha Wolf:* A resource which is expressed as a computer program in source or compiled form. *Gregory Renaud:* A resource which is encoded information used to store and process information as part of a computer system.

Sound	A resource whose content is primarily intended to be rendered as audio.	*Gregory Renaud:* A resource which is primarly intended to be heard.
Still Image	A static visual representation.	*Gregory Renaud:* A resource which is a static visual representation other than text.
Text	A resource whose content is primarily words for reading.	*Gregory Renaud:* A resource which is in primarily words intended to be read.

Summary of Individual Posts

Ann Apps (Fri, 9 Dec 2005 15:05:53 +0000)
DC-General: http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0512&L=dc-general&T=0&P=890

Term	Proposed Language	Public Comment
Service	A service is a system that provides one or more functions of value to the end-user.	A service is a system that provides one or more functions ~~of value to the end-user~~.

Vladimir Makarov (Fri, 9 Dec 2005 10:56:26 -0500)
DC-General: http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0512&L=dc-general&T=0&P=1004Response to Ann Apps (Fri, 9 Dec 2005 15:05:53 +0000)

Term	Proposed Language	Public Comment
Service	A service is a system that provides one or more functions of value to the end-user.	A service is a system that provides one or more functions of value to ~~the end-user~~ [an] actor.

Gregory Renaud (Fri, 16 Dec 2005 10:42:27 -0500)
DC-General: http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0512&L=dc-general&T=0&P=2031

Term	Proposed Language	Public Comment
Collection	An aggregation of items. *Comment:* The term Collection means that the resource is described as a group; its parts may also be separately described and navigated.	A resource which is [a]n aggregation of items. *Comment:* The term Collection means that the resource is described as a group; its parts may also be separately described and navigated. Properties of a Collection may not be identical to properties of its parts; properties of all parts may not be identical.
Dataset	Information encoded in a defined structure (for example,lists, tables, and databases).	~~Information encoded in a defined structure (for example,lists, tables, and databases).~~ A resource in which data is encoded in a defined structure.
Event	A non-persistent, time-based occurrence.	~~A non-persistent, time-based occurrence.~~ A resource which is a time-based activity.
Image	A primarily symbolic visual representation other than text.	~~A primarily symbolic visual representation other than text.~~ A resource which is a visual, non-textual representation.
Moving Image	A series of visual representations that, when shown in succession, impart an impression of motion.	~~A resource which requires interaction from the user to be understood, executed, or experienced.~~ A resource which is a visual representation other than text that imparts motion or an impression of motion.

Physical Object	An inanimate, three-dimensional object or substance.	An inanimate, three-dimensional object or substance. A resource which has three dimensions and mass.
Service	A system that provides one or more functions of value to the end user.	A system that provides one or more functions of value to the end user. A resource which is a system or activity that is needed or requested.
Software	A computer program in source or compiled form which may be available for installation nontransiently on another machine.	A computer program in source or compiled form which may be available for installation nontransiently on another machine. A resource which is expressed as a computer program in source or compiled form; may be available for installation nontransiently on another machine.
Sound	A resource whose content is primarily intended to be rendered as audio.	A resource whose content is primarily intended to be rendered as audio. A resource which is primarly intended to be heard.
Still Image	A static visual representation.	A static visual representation. A resource which is a static visual representation other than text.
Text	A resource whose content is primarily words for reading.	A resource whose content which is in primarily words for reading intended to be read.
Misha Wolf (Fri, 16 Dec 2005 16:25:47 +0000) **DC-General: http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0512&L=dc-general&T=0&P=2161Response to Gregory Renaud (Fri, 16 Dec 2005 10:42:27 -0500)**		
Software	A computer program in source or compiled form which may be available for installation nontransiently on another machine.	*Renaud's Revised Proposal:* A resource which is expressed as a computer program in source or compiled form.~~; may be available for installation nontransiently on another machine.~~
Gregory Renaud (Fri, 16 Dec 2005 11:58:18 -0500) **DC-General: http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0512&L=dc-general&T=0&P=2280**		
Software	A computer program in source or compiled form which may be available for installation nontransiently on another machine.	~~A computer program in source or compiled form which may be available for installation nontransiently on another machine.~~ A resource which is encoded information used to store and process information as part of a computer system.

Metadata associated with this resource: http://dublincore.org/usage/public-comment/2006/03/type-vocabulary-comments/index.shtml.rdf

Title: The NSDL Registry
Identifier: /usage/meetings/2006/04/seattle/nsdl-registry/index.shtml
Source: /usage/meetings/2006/04/seattle/nsdl-registry/index.txt

Shepherd: Diane, with Jon Phipps <jphipps@madcreek.com>

Reading:
[1] http://phoenix.ischool.washington.edu/wiki/index.php/Working_Notes_and_Documentation#NSDL_Registry_Use_Cases

See also:
[2] http://purl.org/nsdlregistry/

Diane proposes a demonstration of the metadata registry currently in development, which in its first stage will be focused on controlled vocabularies:

> In later stages, we intend to include metadata schemas and application profiles. I'd suggest that we schedule this prior to our discussion on UB review of APs, in hopes of bringing to the table some practical issues that sometimes get lost as we hover around 50,000 feet in our discussions.
>
> Jon Phipps, who is working with me and Stuart on this project, will conduct the demonstration, and will also be observing the UB meeting. I think about 45 minutes should cover it, including discussion.

We have some use case documentation available which includes Application Profile use cases [1].

Title: Review of Application Profiles - the pipeline
Identifier: /usage/meetings/2006/04/seattle/profile-pipeline/index.shtml

[1] http://homes.ukoln.ac.uk/~lispj/dc-cd/rep200602.html

-- Collection Description [1] - will not be ready for Seattle - planned for Mexico, see:
/usage/meetings/2005/09/madrid/profile-collection/
/groups/collections/
/usage/meetings/2006/04/seattle/profile-pipeline/2005-11-03.dc-collections-report.html

-- Agents: may be ready by Mexico
/groups/agents/

-- Libraries
/groups/libraries/

DCMI Collection Description

Report of Working Group, February 2006

1. Activities and results since Madrid conference

The principal item of the workplan agreed in Madrid [1] is the finalisation of the Application Profile for collection-level description [2]. Four principal outstanding issues were identified; work is in progress on two of these:

1.1 Revise DC CD AP in light of Usage Board comments

A revised draft of the DC CD AP has been circulated. This draft addresses two changes requested by the DCMI Usage Board after their preliminary review of the DC CD AP:

- minor presentational changes, particularly to expand abbreviations
- the introduction of some new properties to address the problem that in the previous version the constraint on the value-space of a property was not consistent with the value-space of a subproperty of that property

1.2 Representing Item Format

Not yet discussed.

1.3 Collection-Location/Collection-Service relationships

An extension [3] to the *Analytical Model of Collections and their Catalogues* has been developed by Michael Heaney of the University of Oxford, with the aim of clarifying some of the issues surrounding the differences between the Location of a Collection and the Service which makes a Collection available. The issue is currently under discussion via the mailing list, and opinion seems to be in favour of retaining both the isLocatedAt and isAccessedVia properties.

1.4 Encoding Scheme(s) for Open-Ended Date Ranges

Depends on work by DC Date WG.

2. Further planning and activities for 2006

Progress is now running slightly behind the schedule described in the Workplan.

A proposal to address issue 1.2 will be made during March by the WG chair.

3. Use of Wikis to support WG work

Although the DC CD WG does have a Wiki set up, relatively little use has been made of it. It has been used to capture some example instance metadata descriptions based on the DC CD AP.

The principal working documents of the DC CD WG (the draft application profile and subsidiary documents) had already been created as static HTML pages before the Wiki was installed, and work has continued on the HTML versions. If the work on editing these documents was starting now, authoring might have been carried out using the Wiki (though one of the documents does contain some quite complex tables).

Minor problems have been encountered with Wiki spam.

References

[1] Report of Working Group meeting, 11 October 2004, DC-2005, Madrid. http://www.jiscmail.ac.uk/cgi-bin/webadmin?

A2=ind0511&L=dc-collections&P=60

[2] DC Collection Description Application Profile. http://www.ukoln.ac.uk/metadata/dcmi/collection-application-profile/

[3] An Extension of the Analytical Model of Collections and their Catalogues into Usage and Transactions. http://www.ukoln.ac.uk/cd-focus/model-ext/

Title: Wikipedia article on Dublin Core
Identifier: /usage/meetings/2006/04/seattle/wikipedia/index.shtml

Discussant: Diane

Readings:
[1] http://en.wikipedia.org/wiki/Dublin_Core
[2] /usage/meetings/2006/04/seattle/wikipedia/2006-03-21.digest.txt

See also:
[3] http://en.wikipedia.org/wiki/Talk:Dublin_Core

In March, Diane made some initial edits on the Wikipedia article about Dublin Core [1], opening a discussion on dc-usage. Andy posted a strawman description putting DCAM at the center. For Seattle, everyone please read the Wikipedia article and dc-usage digest [2] for a discussion of how "the DCMI story" should be told.

Dublin Core

From Wikipedia, the free encyclopedia

The **Dublin Core** metadata standard is a simple yet effective element set for describing a wide range of networked resources. The Dublin Core standard includes two levels: Simple and Qualified. Simple Dublin Core comprises fifteen elements; Qualified Dublin Core includes three additional elements (Audience, Provenance and RightsHolder), as well as a group of element refinements (also called qualifiers) that refine the semantics of the elements in ways that may be useful in resource discovery. The semantics of Dublin Core have been established by an international, cross-disciplinary group of professionals from librarianship, computer science, text encoding, the museum community, and other related fields of scholarship and practice.

Contents

Simple Dublin Core

The Simple **Dublin Core Metadata Element Set (DCMES)** consists of 15 metadata elements:

1. Title
2. Creator
3. Subject
4. Description
5. Publisher
6. Contributor
7. Date
8. Type
9. Format
10. Identifier
11. Source
12. Language
13. Relation
14. Coverage
15. Rights

Each Dublin Core element is optional and may be repeated. The **Dublin Core Metadata Initiative (DCMI)** has established standard ways to refine elements and encourage the use of encoding and vocabulary schemes. There is no prescribed order in Dublin Core for presenting or using the elements.

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

Qualified Dublin Core

Subsequent to the specification of the original 15 elements, an ongoing process to develop exemplary terms extending or refining the Dublin Core Metadata Element Set (DCMES) was begun. The additional terms were identified, generally in working groups of the Dublin Core Metadata Initiative, and judged by the DCMI Usage Board to be in conformance with principles of good practice for the qualification of Dublin Core metadata elements.

Element refinements make the meaning of an element narrower or more specific. A refined element shares the meaning of the unqualified element, but with a more restricted scope. The guiding principle for the qualification of Dublin Core elements, colloquially known as the "Dumb-Down Principle," states that an application that does not understand a specific element refinement term should be able to ignore the qualifier and treat the metadata value as if it were an unqualified (broader) element. While this may result in some loss of specificity, the remaining element value (without the qualifier) should continue to be generally correct and useful for discovery.

In addition to element refinements, Qualified Dublin Core includes a set of recommended encoding schemes, designed to aid in the interpretation of an element value. These schemes include controlled vocabularies and formal notations or parsing rules. A value expressed using an encoding scheme will thus be a token selected from a controlled vocabulary (e.g., a term from a classification system or set of subject headings) or a string formatted in accordance with a formal notation (e.g., "2000-01-01" as the standard expression of a date). If an encoding scheme is not understood by an application, the value may still be useful to a human reader.

DCMI also maintains a small, general vocabulary recommended for use within the element Type. This vocabulary, consisting of 12 terms, can be found on the DCMI webpage or the Dublin Core Metadata Registry (see References).

Dublin Core Syntaxes

Syntax choices for DC metadata depend on a number of variables, and "one size fits all" prescriptions rarely apply. When considering an appropriate syntax, it is important to note that Dublin Core concepts and semantics are designed to be syntax independent, are equally applicable in a variety of contexts, as long as the metadata is in a form suitable for interpretation both by machines and by human beings.

The **Dublin Core Abstract Model** (see References) provides a reference model against which particular DC encoding guidelines can be compared, independent of any particular encoding syntax. Such a reference model allows implementors to gain a better understanding of the kinds of descriptions they are trying to encode and facilitates the development of better mappings and translations between different syntaxes.

Application Examples

One Document Type Definition based on Dublin Core is the Open Source Metadata Framework (http://www.ibiblio.org/osrt/omf/) (OMF) specification. OMF is in turn used by ScrollKeeper, which is used by the GNOME desktop and KDE help browsers and the ScrollServer documentation server. PBCore is also based on Dublin Core. The Zope CMF's Metadata products, used by the Plone and the Nuxeo CPS Content management systems, also implement Dublin Core.

DCMI also maintains a list of projects using Dublin Core on its website.

See also

- Controlled vocabulary
- Interoperability
- Open Archives Initiative
- Semantic Web

References

- Dublin Core Metadata Initiative (http://dublincore.org/)
- "Using Dublin Core" –the official usage guide (http://dublincore.org/documents/usageguide/)
- Dublin Core Metadata Registry (http://dublincore.org/dcregistry/)
- Dublin Core Abstract Model (http://dublincore.org/documents/abstract-model/)
- Dublin Core Metadata Initiative Publishes DCMI Abstract Model (http://xml.coverpages.org/ni2005-03-21-a.html) (*Cover Pages*, March 2005)

Retrieved from "http://en.wikipedia.org/wiki/Dublin_Core"

Categories: Knowledge representation | Library and information science | Metadata | Semantic web

---

Date: Tue, 21 Mar 2006 11:04:37 +0000
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Updated Wikipedia article
To: DC-USAGE@JISCMAIL.AC.UK

Diane I. Hillmann wrote:
> Folks:
>
> I went in and changed the page: http://en.wikipedia.org/wiki/Dublin_Core
> using stuff from "Using Dublin Core" primarily.
>
> You're welcome to see what you think and do some editing yourself, if
> you've a mind to do so. I won't take it personally, I promise!

I haven't edited the document on Wikipedia (yet), and I don't really want to do so unilaterally, but (and this really applies to other introductory resources like the "Using Dublin Core" document and like the conference tutorials) I think that the role of the DCMI Abstract Model as providing the conceptual framework for DC should be presented more "up front" e.g. there should be a short/simplified summary description of the DCAM in the introductory section, or as a separate section following the intro and before the discussion of Simple DC and Qualified DC. Mentioning the DCAM _only_ as a tool for comparing different syntaxes is only telling half the story (IMHO) - though I think the DCAM doc itself might be slightly more "bullish" in its own intro paragraph! ;-).

Now sure, I appreciate that that might look like a case of "historical revisionism" to the casual observer who knows "the 15 elements came first" (or indeed only knows the 15 elements), but I do think we need to shift firmly towards putting the DCAM at the centre of our explanations of "what DC is". (If people want a history of the evolution of DC, and how DCMI got from "the 15 elements" to "the qualifiers" to the grammatical principles to the DCAM, OK, that's fine, but that's a different document.)

I recognise this probably goes against the way we've tended to introduce DC, but I'd go as far as saying that it is confusing/unhelpful to start talking about "elements" without first describing the DCAM, at least in some way - maybe not every fine detail, but the fundamental points about making statements that assert relationships between resources and values. Without such "contextual" information, it just begs the question of what an "element" is. It makes a "leap of faith" that readers already share a common understanding of what an element is, but (as we've found out somewhat painfully over the last few years), that is not the case: the term "element" is used to refer to different things in different contexts and readers draw their own (different, incompatible) conclusions ("Ah, they're talking about XML elements", "Ah, so they're referring to things like LOM elements", "Ah, they mean attribute-value pairs" etc etc etc).

I think the account of "Simple Dublin Core" also blurs the distinction between the DCMES as a set of terms, each of which may be deployed in many different "DC application profiles" with different constraints on their usage in a description set, and "Simple Dublin Core" as one such DCAP with one particular set of constraints. And in the account of "Qualified DC" I'm not sure the word "value" is being used in the way it is used in the DCAM. I think phrases like "the value may still be useful to a human reader" suggest that the reference is to (what the DCAM calls) "value strings".

Date: Tue, 21 Mar 2006 11:55:30 -0000
From: Andy Powell <andy.powell@EDUSERV.ORG.UK>
Subject: Re: Updated Wikipedia article
To: DC-USAGE@JISCMAIL.AC.UK

Yes, I agree with this... though it's certainly not easy to come up with the one paragraph summary of what DC is. That's one of our problems! But talking about 'elements', 'simple' and 'qualified' doesn't help much!

Here's my stab at a summary...

--- cut ---

Dublin Core (DC) is a metadata standard for describing a wide range of digital, physical and conceptual resources (i.e. just about anything!).

A DC description is made up of a set of statements, each of which comprises a property/value pair. Typically, the described resource is identified using its URI and the value is either identified using its URI or represented using a simple string (the 'value string'). In many cases, multiple descriptions are combined in order to build up richer descriptions ('description sets') about related groups of resources. For example, in describing a digitised painting, it may also be appropriate to describe the original painting from which the digitisation was made and the original artist, thus creating a description set of three related descriptions. Statements may be refined by indicating the language of the value string, any data-type ('syntax encoding scheme') to which the value string conforms or the class ('vocabulary encoding scheme') from which the value is taken. Properties, syntax encoding schemes, vocabulary encoding schemes and concepts in controlled vocabularies are known as 'terms'. All terms in DC metadata are assigned URIs, and schema languages are used to indicate the relationships between them.

The features of the DC metadata standard are fully described in the Dublin Core Abstract Model (DCAM).

The Dublin Core Metadata Initiative (DCMI), the body that facilitates the community development of the DC metadata standard, provides a core set of about 80 properties, encoding schemes and controlled vocabularies from which descriptions can be constructed, but encourages communities to create additional terms as necessary, within the framework provided by the DCAM.

Finally, the DCMI community has defined three encoding syntaxes that can be used to encode DC metadata records for exchange between software systems and services using XHTML, XML and RDF.

Historically, DC refered to properties as 'elements' and is perhaps best known for the Dublin Core Metadata Element Set (DCMES) - a set of 15 elements, created originally to support the discovery of resources on the Web.

--- cut ---

This, rightly(!), relegates DCMES to a footnote and doesn't even mention simple and qualified! :-)

On the face of it, it may seem harder to grasp than the traditional 'element', DCMES, 'simple DC', 'qualified DC' approach - but I think that is largely to do with where we come from. Personally, I think it is much clearer - it emphasises what is important and ignores what can safely be forgotten.

Date: Tue, 21 Mar 2006 12:06:53 -0500
From: "Diane I. Hillmann" <dih1@CORNELL.EDU>
Subject: Re: Updated Wikipedia article
To: DC-USAGE@JISCMAIL.AC.UK

Tom, responding to Pete and Andy:

>I agree that the Abstract Model should be at the center of
>explanations, and I like most of the text. I also think we
>need to revisit the explanation for "simple" versus "qualified"
>Dublin Core.

I agree that the Abstract Model should be better integrated into the text, and have said more than once that I'd like to find a way to integrate the AM better into "Using Dublin Core," but could use some help with that. Because I used text from "Using Dublin Core" as the basis for what I added to the article, it suffers from some of the same problems.

>Two issues however:
>
>1) To start with statements, property-value pairs, and URIs is
> to jump in at the deep end. The notion of "core metadata

>   properties" has been and should continue to be a key part
>   of the message. Introducing "the Dublin Core" up-front also
>   helps explain the funny name.
>

Yes, in my experience, the DCAM certainly is the deep end for most
people. I think it makes more sense to use Wikipedia to provide a
"soft" approach, dealing first with things people might have already
heard about DC, and making that clearer before getting into deeper
water.

>   > The Dublin Core Metadata Initiative (DCMI), the body that facilitates
>   > the community development of the DC metadata standard, provides a core
>   > set of about 80 properties,
>
>   I suspect that the notion of "core metadata properties" really
>   is easier for most people to grasp than the notion of an
>   abstract model. I would not want DCMI to lose that focus --
>   80 properties is no longer really a "core".

Agreed. Despite our mere 10 years of "legacy," we are seen by many in
an increasingly complex world as more approachable and
understandable--metadata for the masses, perhaps. We can't afford to
lose that advantage.

>2) "DC" and "DCMI"
>
>   > The features of the DC metadata standard are fully described
>   > in the Dublin Core Abstract Model (DCAM).

Usage aside, I think the sentence above denies the fact that the
explanation provided in the DCAM, while certainly full, is fairly
technical from the point of view of the metadata novice.

>   This is not quite how I understand our current use of
>   "DC" and "DCMI". To take two important examples, we
>   currently have the "DCMI Abstract Model" and we talk
>   about "Dublin Core Application Profiles" (but not "DC
>   Application Profiles" or "DCMI Application Profiles").
>   In a reasonably consistent way, we currently:
>
>   -- use "Dublin Core" instead of "DC". The exception is when
>   we use the names of Jiscmail lists as handles for working
>   groups (e.g., "DC-Architecture"). In DCMI publications,
>   we have I think avoided using the free-standing acronym
>   "DC" to mean just "Dublin Core" for quite a long time.
>   If we were to revive it, would we do so with the intent
>   to refer to "the Dublin Core" (DC-15) or to designate
>   the abstract model?
>
>   -- use "DCMI" for things that are managed or maintained by
>   DCMI as an organization -- e.g., "DCMI Usage Board".
>   The DCMI Abstract Model is a model put forward and
>   maintained by DCMI for metadata that uses DCMI terms,
>   (whereas a Dublin Core Application Profile is in most
>   cases not put forward or maintained by DCMI).

I *think* I followed most of this pattern, but may have blundered a bit ... ;-)

Date: Wed, 22 Mar 2006 07:52:51 +0000
From: Pete Johnston <p.johnston@ukoln.ac.uk>
Cc: DC-USAGE@jiscmail.ac.uk
Subject: Re: Updated Wikipedia article

Quoting Thomas Baker <tbaker@tbaker.de>:

> Two issues however:
>
> 1) To start with statements, property-value pairs, and URIs is
>   to jump in at the deep end. The notion of "core metadata
>   properties" has been and should continue to be a key part
>   of the message. Introducing "the Dublin Core" up-front also
>   helps explain the funny name.
>
>   > The Dublin Core Metadata Initiative (DCMI), the body that facilitates
>   > the community development of the DC metadata standard, provides a core

> > set of about 80 properties,
>
> I suspect that the notion of "core metadata properties" really
> is easier for most people to grasp than the notion of an
> abstract model. I would not want DCMI to lose that focus --
> 80 properties is no longer really a "core".

Unless we explain what we mean by "property", and show how properties are used in DC metadata (to construct simple statements that assert relationships between resources), then we are left with the same problem with the word "property" that we have now with the word "element".

Agreed, "property" maybe comes with slightly less "baggage" than "element", but I've had experience of people picking up on my tendency to use "property" rather than "element" and altering their usage accordingly. But unless I explain why I'm saying property - what a property is - then they think I'm using it simply as a synonym, and they start using "property" for all the contexts in which they were using "element", including ways which are inconsistent with the DCAM.

And - perhaps even worse - they start to apply the term "property" to things defined in other meta-models - where those thinga are not properties at all, and the "native" (to that meta-model) term "element" is indeed the most appropriate term.

Whatever we call the fifteen things, we have to explain what they are and how they are used.

Date: Wed, 22 Mar 2006 13:34:49 -0000
From: Andy Powell <andy.powell@EDUSERV.ORG.UK>
Subject: Re: Updated Wikipedia article
To: DC-USAGE@JISCMAIL.AC.UK

> I suspect that the notion of "core metadata properties" really
> is easier for most people to grasp than the notion of an
> abstract model. I would not want DCMI to lose that focus --
> 80 properties is no longer really a "core".

OK, could drop the specific ('80') and leave the general ('core set of ...')?

But main problem is with emphasis and order. It's difficult to say what DCMI provides (essentially 'a core set of terms that can be used to create descriptions of resources') without describing the nature of what those terms are and how they fit together.

Perhaps in my suggested text,

The Dublin Core Metadata Initiative (DCMI), the body that facilitates the community development of the DC metadata standard, provides a core set of about 80 properties, encoding schemes and controlled vocabularies from which descriptions can be constructed, but encourages communities to create additional terms as necessary, within the framework provided by the DCAM.

needs to be moved up front and changed to

The Dublin Core Metadata Initiative (DCMI), the body that facilitates the community development of the Dublin Core metadata standard, provides a flexible and extensible metadata framework and a core set of terms from which descriptions can be constructed, but encourages communities to create additional terms as necessary to meet their own descriptive requirements.

Then go on to describe the framework (the DCAM) in more detail...

Does that help?

> 2) "DC" and "DCMI"
>
> > The features of the DC metadata standard are fully described
> > in the Dublin Core Abstract Model (DCAM).
>
> This is not quite how I understand our current use of
> "DC" and "DCMI".

Sorry. I was writing in a hurry so I didn't follow corporate
guidelines! :-)

Title: Review of Resource Description and Access (RDA, aka AACR3)
Identifier: /usage/meetings/2006/04/seattle/rda-review/index.shtml

Discussant: Diane

[1] /groups/libraries/rda/
[2] /usage/meetings/2006/04/seattle/rda-review/2006-02-06.digest.html
[3] /usage/meetings/2006/04/seattle/rda-review/2006-03-28.rda-discussion.html
[4] /usage/meetings/2006/04/seattle/rda-review/SCScommentsRDAPart1-excerpts.pdf
[5] /usage/meetings/2006/04/seattle/rda-review/RDA_for_who.htm

The DCMI Libraries Working Group has a Working Group for Reviewing Resource Description and Access (RDA) [1]. There has been discussion on various DCMI lists raising fundamental issues about "cataloging" in the Web environment [2,3]. Diane has included here a draft essay about the issues [5].

In particular, there have been some important comments on RDA part 1 from the the Program for Cooperative Cataloging of The Library of Congress, Standing Committee on Standards (SCS) [4], http://www.loc.gov/catdir/pcc/standards.html.

As a Usage Board we should be aware of these discussions and consider taking a stand or otherwise contribute to this discussion. Please review the readings included in the packet for discussion in Seattle.

Working Group for Reviewing Resource Description and Access (RDA)

Chair: Robina Clayphan
The British Library, UK
RDA Liaison: Matthew Beacom
Yale University
Members: Tom Baker - DCMI
Brad Eden - University of Nevada, USA
Corey Harper - University of Oregon, USA
Christine Frodl - Die Deutsche Bibliothek, Germany
Jane Greenberg - University of North Carolina, USA
Robin Wendler - Harvard University, USA
John Chapman - University of Minnesota, USA
Marty Kurth - Cornell University, USA
Dr. RSR Varalakshmi - Andhra University, India
Cassandra Armstrong - Library Associates, USA
Douglas King - University of South Carolina, USA
Leif Andresen - Danish National Library Authority, Denmark
Anchalee (Joy) Panigabutra-Roberts - St Cloud University, MN, USA
Susan Pyzynski - Harvard, USA
Description: This working group -- a task force of the DCMI Libraries Working Group -- has been formed to review drafts of the new content standard "Resource Description and Access" (RDA) and to offer input into the drafting process from a Dublin Core perspective.

Context

Rules for creating the content of metadata descriptions constitute an important aspect of interoperability. Content-level agreements exist within specific communities but interpretation of the content can become problematic when data is moved between domains.

"Resource Description and Access" (RDA) is a new content standard for resource description being developed on the foundations of library cataloguing practice established in the Anglo-American Cataloguing Rules (AACR).

The Joint Steering Committee intends that RDA will provide a comprehensive set of guidelines and instructions covering all types of content and media. To achieve this, consultations are being undertaken with other communities in an effort to attain an effective level of alignment between RDA and the metadata standards used in those other communities. DCMI has thus been approached and this working group has been formed to review drafts as they are produced and to offer any other relevant input from a DC perspective.

Timeline

- Comments on Draft of Part 1 due on 1 March 2006
- Committee meets 24-28 April to consider comments and launch Part 2
- Completion and Review of Part 2, May - September 2006
- Completion and Review of Part 3, October - April 2007
- Completion of Introduction, Appendices, Glossary May - September 2007
- Publication in 2008

It is planned to produce a Web version of the standard and a beta release may be ready for testing in mid-2007.

Mailing-list archive

An archive of the closed-membership DC-RDA mailing list may be found at http://www.jiscmail.ac.uk/lists/dc-rda.html .

Links

- Part 1 of RDA, draft
  http://www.collectionscanada.ca/jsc/rdadraftpt1.html
- Official overview of the whole proposed standard
  http://www.collectionscanada.ca/jsc/rdaprospectus.html
- Overview, with timeline
  http://www.collectionscanada.ca/jsc/rda.html
- Joint Steering Committee for Revision of Anglo-American Cataloguing Rules
  http://www.collectionscanada.ca/jsc/index.html

Developments (reverse-chronological order)

2006-02-06 Diane Hillmann report from ALA, San Antonio
-- http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0602&L=dc-general&P=56

2005-10-21 Comments from Die Deutsche Bibliothek
-- http://dublincore.org/groups/libraries/files/2005-11-14.Comments_Prospectus_DDB.pdf

2005-10-15 Invitation to join group
-- http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0510&L=dc-libraries&P=179

2005-09-15 Matthew Beacom presentation, DC-2005 in Madrid
-- http://www.britishlibrary.net/RDAandDCMI.ppt

2005-09-15 dc-libraries meeting report
-- http://www.jiscmail.ac.uk/files/DC-LIBRARIES/DC_Libraries_-_meeting_report_2.doc

2005-05-31 Reports from earlier AACR3 activities
-- http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0505&L=dc-libraries&P=53
-- http://dublincore.org/groups/libraries/files/aacr3-part1-DCMI-response.doc
-- http://dublincore.org/groups/libraries/files/aacr3-part1-DeutscheBibliothek.doc
-- Details of the JSC decisions can be found at
http://www.collectionscanada.ca/jsc/0504out.html

Metadata associated with this resource: http://dublincore.org/groups/libraries/rda/index.shtml.rdf

---
2006-02-06

From: General DCMI discussion list on behalf of Diane I. Hillmann
Sent: Mon 06/02/2006 13:48
To: DC-GENERAL@JISCMAIL.AC.UK
Cc:
Subject: Fwd: Issues in Applying RDA in Non-MARC Metadata Communities

Folks:

At the recent American Library Association meeting in San Antonio, I attended meetings concerned with the new Resource Description and Access (RDA) standard being developed as a replacement for the Anglo-American Cataloging Rules. The opinions represented below are mine only, intended for the purposes of inciting discussion about the issues.

Regards,
Diane

--- begin forwarded text

Date: Mon, 6 Feb 2006 08:45:39 -0500
To: RDA-L@INFOSERV.NLC-BNC.CA
From: "Diane I. Hillmann" <dih1@cornell.edu>
Subject: Issues in Applying RDA in Non-MARC Metadata Communities
Cc:
Bcc:
X-Attachments:

Issues in Applying RDA in Non-MARC Metadata Communities

Diane I. Hillmann
Jan. 31, 2006

Having come rather late to the RDA discussion, I recognize that the points I bring up in this document may not be the comments that CC:DA or JSC expects or desires at this stage, but I hope the concerns expressed may be useful nonetheless.

First, I'd like to describe the issues I see in applying the underlying assumptions made in RDA (most of which emanate from traditional cataloging practices) to the world of Non-MARC Metadata (NMM) Communities. Clearly, these issues are less apparent to traditional libraries whose preponderant exposure to digital resources are digital versions of published materials, but once outside that familiar boundary the environment is much less comfortable.

I believe the primary issues that concern me lie in the following areas:

Transcription as Identification

In the world of traditional cataloging and static published resources, the notion of consistent transcription as an important method to assure predictable access, from a variety of agencies handling exactly the same resources, made a great deal of sense.

However, digital resources carry no such assumption of stability-change is part of the package. In that environment, relying on use of consistently transcribed information as the primary method of identifying a resource makes much less sense. Resources in this environment are most often unique, and usually identified by a numeric or alpha-numeric string. In traditional cataloging, such identifiers are also used, of course: ISSNs and ISBNs are the most obvious examples, but they are generally not the primary identification of the resource.

As we all know, the current methodologies for identifying digital resources uniquely and unambiguously are still in flux and almost no one is satisfied with the current situation.

But whatever the ultimate answer, it will not rely on transcription, nor will decisions about what constitutes a "new" resource likely be susceptible to the rules defined for editions or versions. It should also be noted that the gold standard of infallible identification in a metadata description is not always necessary for digital information, where the resource itself can often be viewed easily and quickly.

Reliance on Notes

Oftentimes, the RDA (like traditional cataloging) herds catalogers to make decisions about what is "primary" or "secondary" and relegates the latter to the notes area. This is a significant problem for many NMM communities, who may either have no place to put this kind of descriptive "notes" or who rely on repetition of elements (with or without a notion of order) to capture information of the same kind within a single description, thus focusing more on access than descriptive integrity.

In most delivery systems for metadata (including OPACS, it must be noted), only the information in a small number of specified fields is actually displayed to the user (and we know few users actually look at full records). Additionally, because notes can contain so many different categories of information, they may not even be indexed (when they are, only as keywords). For systems using NMM, notes information is even less likely to be displayed, and may indeed be entirely ignored, since its "human-friendly" character makes it useless for machine processing and marginal for access.

Reproductions

I brought up the issue of reproductions on the RDA-L list and was dismayed to see how many catalogers were still trying to make the case for describing an original and a reproduction on the same record. If FRBR is truly underlying RDA, I believe this bullet must be bitten firmly and these practices explicitly marginalized within the context of the rules. In an environment where metadata of different formats created using different rules (or no rules) must be shareable, these residual practices keep us all from benefiting from our common enterprise.

Yes, it is certainly true that most vendor systems do not display multiply versioned resources acceptably, but we undercut the usefulness of our data by manipulating it to overcome system inadequacies; rather, we should address those problems with our vendors.

Source of Information

Specification of sources of information from which to record information grew logically from the reliance on transcription, the goal being consistency. Vital to this approach is the idea that resources have commonly identified and named parts that are similar within a specific category of materials, something that is not generally the case in the digital world.

Similarly, notions of whether information comes from the item itself or is supplied from somewhere else are often less important in NMM communities, even those who still deal primarily with physical, published items. In ONIX for example, information about the author (from the book jacket, reviews, or other marketing sources) is specifically tagged based on the function of the information, and it's often not explicitly descriptive in nature.

Future Considerations

As I mentioned in my comments at the Monday CC:DA meeting at ALA, we may increasingly be thinking less about the cataloging record as the lowest unit of description and more as the "statement" as the optimal unit. In that context, "Who says?" or "When said?" or "In what language?" is likely to be more important information to know in order to manage the information than where in a resource the information was

found, and the current RDA doesn't support these notions at all. I suspect we'll begin to see this change in thinking more as we discuss common authority files, where explicit specification of language and form of heading are critical to making appropriate choices for usage in different catalogs, in the context where the concept of an individual "statement" has already taken root.

Some of these attributes are easier to manage outside of MARC (XML, for instance, supports language specification at various levels), but it's really important that we start thinking along those lines sooner, rather than later.

The ideal of the current RDA still seems to be the anonymous cataloger acting objectively using a commonly understood set of rules, providing consistent records suitable for sharing. Clearly, the sharing and integration pieces are still critically important, but we may not be able to afford the levels of consistency and predictability that we've had in the past. Other mechanisms may be available to improve access in ways we don't understand fully at the moment, but we should probably at least explore some of the possibilities at this juncture.

I'm not entirely sure how to where to go from here, but it might be useful to examine some strategies whereby the most basic level of RDA instruction might be more generally useful outside the traditional library environment, given the dissonances noted above.

2006-02-06

On Mon, Feb 06, 2006 at 06:13:13PM -0000, Robina Clayphan wrote:
> This message is to ensure that you have all seen the message
> sent by Diane Hillmann to the DC-General and RDA lists
> (see below). Diane attended the CC:DA meeting at ALA with
> a DC hat on and raises some fundamental points about RDA and
> non-MARC- metadata creation that we could usefully consider.
>
> As a non-AACR person I have been struggling to review
> the current draft of RDA from a DC perspective. There is
> a mismatch that is not easy to express but which Diane
> has begun to articulate. What I felt was the need for
> a specification of DC requirements for content level
> rules in order to assess RDA against those requirements.
> This specification of requirements will need to be based on
> a statement of assumptions or principles which this group,
> possibly together with other DC groups, should undertake
> to create. Diane's paper points the way to such a statement.
> I would be interested to hear you thoughts on this.

Dear all,

I have spoken with Robina about the DCMI role in providing input to the revision of RDA. We agreed that it would be useful to have a two-page position paper with something like the following:

-- DCMI principles (e.g., one-to-one, "values" and "value representations" in the Abstract Model, etc).

-- Requirements: what do DC implementers look for from a standard like RDA?

We feel encouraged by the willingness (as I understand it) of the RDA maintainers to make it easier for users to extract selected sub-sets of the entire standard for particular applications. I'm wondering if we (i.e., DCMI) could focus in on a sub-set of rules of likely applicability to DC metadata implementers instead of trying to review RDA as a whole.

Ideally, we would have a draft two-pager on the table for the Usage Board meeting in Seattle for discussion and, possibly, some sort of official UB approval. Diane's posting to dc-general [1, and copied below] provides a start.

Does that seem like an achievable objective, and are there
any volunteers?

Tom

http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0602&L=dc-general&P=56

Date: Fri, 10 Feb 2006 22:39:58 -0500
From: "Diane I. Hillmann" <dih1@CORNELL.EDU>
Subject: Re: Input to RDA process
Comments: cc: "Clayphan, Robina" <Robina.Clayphan@bl.uk>
To: DC-USAGE@JISCMAIL.AC.UK

Tom:

I think your second point is the strongest need,
at this stage. One of the legacies of AACR is the
important distinction between "description" and
"access." I think that to some extent the DCAM
principles are not particularly relevant (except
for "one-to-one) to the cataloger notion of
"description", though there will likely be more
relevance when the discussion moves from the
"descriptive" rules to the "access" portion. To a
great extent the "one-to-one" rule is reflected
in RDA by the adherence (such as it is) to FRBR
principles.

Reviewing RDA as a whole, particularly for
someone not conversant with AACR2, is likely to
be a pretty frightening experience. There is an
expression of principles at the head of the
current Part I, which is only marginally helpful,
since it does not address some of the hidden
assumptions that are carried over from AACR2
(some of which I tried to address in my comment).

I suspect one of the most useful things that we
could do, as a group, is bring to the discussion
our experiences with interoperability issues and
the uses (and misuses) we find when applying
machine-based understanding to human created data.

We might also try and bring in Wayne Hodgins to
the discussion, who has also undertaken to
participate on behalf of LOM (I didn't get to
speak to him at any of the meetings but he was in
attendance at a few of them).

Date: Thu, 23 Feb 2006 11:32:57 -0500
Reply-To: "Diane I. Hillmann" <dih1@CORNELL.EDU>
Sender: General DCMI discussion list <DC-GENERAL@JISCMAIL.AC.UK>
From: "Diane I. Hillmann" <dih1@CORNELL.EDU>
Subject: Re: [RDA-L] NLM/MLA View on Critical Deficiencies of RDA
Comments: To: Resource Description and Access / Resource Description and Access
<RDA-L@INFOSERV.NLC-BNC.CA>
To: DC-GENERAL@JISCMAIL.AC.UK

Alice, et al.:

I'd like to comment briefly on the contents of your document, which I
agree with in principle and disagree with strongly in some specifics.

To quote the document:

"As stated previously, NLM remains dismayed that, as presently
constituted, RDA represents a largely a cosmetic revision, rather
than a high-level reexamination of content. In order to make the huge
investment that adopting RDA entails (e.g., changes to documentation,
training efforts, programming changes, etc.), we stress that libraries
need a document that addresses critical bibliographic issues/problems,
some of which have plagued libraries since AACR2 itself was published,
as well as broadens the bibliographic perspective to cover today's digital

environment. We concur with others, specifically Paul Weiss, who stated that it is appropriate and important to raise issues in the ALA response to the JSC that need to be addressed either from AACR2 or in the current RDA draft. In its strict constructionist view, we believe that the JSC is failing to meet its own stated goal that 'RDA =8Abe a new standard for resource description and access, designed for the digital world.'"

I think this point is absolutely correct, and concur with the concern that the rush to meet deadlines has to a great extent overpowered discussion of first principles. We know our world has changed and will continue to change dramatically, but you'd hardly know that from the highly specific discussions going on at present.

That said, I see a real disconnect between the general and the specific recommendations in the document, as it goes on to suggest:

o Incorporation of rules to allow use of a single record to describe multiple entities in certain instances

Since the beginning of the rule revision process, NLM has stressed the need to provide guidance on the critical cataloging question of multiple versions. We continue to view it as imperative to incorporate, in some fashion, the highly-used and pragmatic approach of using a single record for multiple entities in certain instances.

o Treatment of the description of a reproduction based on its original

We included this in our comments in confluence as a major problem and concur with the statement of others that this is a long-standing practice that likely will continue even if RDA cements its current position. We contend that, at a minimum, RDA should provide an option that would recognize what is now and what is anticipated to be U.S library practice in the future.

o Rules on Changes Requiring a New Description

=46irst, we believe that RDA should address what changes require a new description for all modes of issuance. The rules in the current draft relate solely to serials.

The first two points seem to me to contradict the general thrust of the earlier paragraph. We seem to be in deep denial that the "highly-used and pragmatic approach" is in fact a serious kludge, and one that will not allow us to either share data effectively or take advantage of some of the =46RBR-based research on improved display of manifestations and expressions. Consider for a moment some of the work OCLC has done with =46ictionFinder--all based on the notion that there is for the most part sufficient data in the records to associate them in this way and create a much more useful display than current library vendors now offer.

In my opinion, the more we cling to these practices, the more we hobble ourselves in the future. Instead, we should be examining the conclusions of projects such as FictionFinder and others to determine what should be included in records to allow the associations we need to create to make useful displays possible, and then =46LOG OUR VENDORS to provide that capability. Instead we insist on manipulating our essential data to overcome the deficiencies of our software, and thus mortgage our future options. This state of affairs is, in my opinion, a direct outcome of the lack of discussion on first principles that we've skated over in our rush to talk about specific rules, and it is exactly the reason why the goal of RDA as "... a new standard for resource description and access, designed for the digital world" is looking increasingly out of reach.

The third point about "what changes require a new description" is an important one, particularly in the digital world. What I'd suggest though, is that rather than proliferate the notion that there might be differences in the rules for what constitutes sufficient difference based on mode of issuance or genre, that we try to come up with something more general that can be applied across the board.

For those of you who don't believe that the powers-that-be in our libraries will over-rule any rearrangement of the deck chairs on the Titanic, I'd suggest you take a look at the new report from the University of California:
http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf
This report is being discussed at some level in most large academic

libraries in this country, and the implications for an RDA not based on a significantly broader view of the world are quite chilling.

Date: Thu, 30 Mar 2006 09:33:56 -0500
To: Thomas Baker <baker@SUB.UNI-GOETTINGEN.DE>
From: "Diane I. Hillmann" <dih1@cornell.edu>
Subject: Fwd: Re: [ALA-CCDA:5336] RDA: an analogy

Date: Tue, 28 Mar 2006 13:04:27 -0500
To: ala-ccda@ala.org, RDA-L@INFOSERV.NLC-BNC.CA
From: "Diane I. Hillmann" <dih1@cornell.edu>
Subject: Re: [ALA-CCDA:5336] RDA: an analogy
Cc:
Bcc:
X-Attachments:

Paul:

Thanks for sharing this, I think it expresses some of my concerns as well (though you're clearly more adept with a metaphor than I am!).

I'm also concerned that, if we defer dealing with these issues now (largely in our desire to meet a publication deadline) we may never have the opportunity to provide the leadership you discuss below. If the RDA proceeds along the current path, others will fill the void we refuse to address, and the RDA will indeed be AACR3 in everything but name.

I spent part of this weekend re-reading two reports: one is the recent one from California ("Rethinking How We Provide Bibliographic Services for the University of California," available at: http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf) and the other one (not yet publicly released) by Karen Calhoun for the Library of Congress. These two reports should be required reading for all those participating in RDA development--they are the leading edge of calls to drastically simplify what we do in library cataloging, and look towards other models of bibliographic service provision for cues to our future.

I take issue with a few of the specific recommendations, but support the general thrust away from our wagon-centric legacy. We have an enormous amount of experience and knowledge to offer the rest of the world, as they struggle with the challenges of metadata creation, maintenance and distribution, but we will be pushed to the side of the road if we can't broaden our view--now, not for RDA2.

Regards,
Diane

>At one point in the process of developing PCC SCS comments on part 1
>of RDA, I was having trouble communicating my
>thoughts/feelings/vision on the big picture aspects of RDA to other
>committee members. I developed the analogy below as a way for me to
>clarify my thoughts and to communicate them more clearly and at a
>deeper level. I'm sending this out more broadly in case it might be
>helpful for, possibly even resonate with, some of you.
>
>Let's view cataloging codes as technical/engineering plans. AACR1
>produced wagons that for the first time could really be
>mass-produced in the state of Libraria, and it allowed drivers to
>switch from one wagon to another without significant retraining.
>AACR2 was an improvement--it standardized even more pieces, so
>drivers switching wagons needed just a little retraining, but
>clearly showed where differences were appropriate and how to deal
>with them. AACR2 wagons were bright and shiny in their day, but
>their age is showing. AACR2 wagons don't work as well on newfangled
>paved roads, they look old-fashioned, people have gotten used to
>comfy furniture and don't like the hard seats, and some of our
>sister states have been developing their own modes of
>transportation. AACR3 required our manufacturing plants to vastly
>retool even though it would have produced wagons that looked and
>worked much the same as before, although in some cases less well.
>This draft of RDA is a more cohesive, thought-out plan, which would
>produce generally well-designed wagons, with all of their different
>parts working even better together, that would be easier to train

>new drivers on. RDA wagons would probably work better than AACR2
>wagons in the state of Libraria.
>
>However, the intention is that RDA is designed so that people in
>other states will use it to design and build wagons very similarly
>to ours, so that wherever people go in the country of
>Informationland, the wagons will be compatible (interchangeable
>parts, no retraining needed for drivers, etc.) and usable
>(comfortable, fast, going where people want to go). Libraria used to
>be seen as the Swiss watchmaker of wagons throughout Informationland
>and beyond, but no longer. Some of our newer sister states don't
>really know much about us at all, let alone about our wagons. Even
>the ones who are familiar with our wagons don't see them as relevant
>to their states. Many of our sister states have been busy. In the
>past several years they have been frantically designing, building,
>and selling many other modes of transportation. Wagons have become
>only one of many modes, they are slow and expensive to produce, and
>they take a long time to ship to the people who need transportation.
>Wagons are just not that efficient or effective any more, and are
>becoming increasingly unpopular as people migrate to other modes of
>transportation. Citizens of most countries throughout the Union of
>Users will ignore RDA and will come to view its wagons as cute, but
>archaic, and will rarely want to use one. Libraria's reputation will
>suffer, and other states and countries won't look to us for
>leadership, not just in transportation, but in all of our areas of
>expertise. Which is a shame, because Librarians really do know a lot
>that can help others achieve their Information goals.
>
>We need to change RDA to be a broader, integrated plan encompassing
>cars, airplanes, trains, ships, etc., and yes, even wagons. As we
>articulate and live up to our objectives and principles, and send
>emissaries to the other states, others will see how their modes of
>transportation can successfully integrate with other modes to
>achieve a cohesive and harmonious transportation system. Our
>neighboring states of Museumdom and Archivia will find the RDA model
>useful to them, and design their modes to interconnect with ours.
>There is a much better chance that states in the Digitalia region,
>such as DublinCore, the Republic of LOM, the GIS Kingdom, and new
>states that we may not even have heard of, will take notice of RDA
>and use it to meet some of their needs. We will likely produce fewer
>wagons, but we will trade modes with other states, and all of our
>states will flourish.
>
>As the various states in Informationland work together to build this
>wondrous system, citizens from other countries will look forward to
>spending time in our country, knowing that they can get from one
>state to another easily and quickly. Peoples of the world will come
>to know Libraria anew, and no longer will some of them think of
>Libraria as a historical landmark, but as a thriving, useful, fun
>hotspot.
>
>:-)
>
>Onward Libraria!
>Paul
>Chair, PCC SCS
>
>_____
>Paul J. Weiss
>Catalog Librarian and NACO Coordinator
>Metadata Services Department
>UCSD Libraries
>858-534-3537
>pweiss@ucsd.edu
>_____

PCC SCS comments on RDA part I

submitted February 2006, Paul J. Weiss, Chair, PCC SCS
Comments in red are the ones that I feel are of greatest importance. --Paul

Overall comments

We appreciate that JSC heard many of the comments that were made in the review of AACR3, and made significant changes that we see now in RDA. However, after much reading, thinking, and talking with colleagues, with very mixed emotions, Some of us feel that we need to again stop the current process and take an even more radical change of course. We don't say this lightly, as we greatly respect the people most closely involved with RDA development. Although this draft is a definite improvement, it has not gone far enough to become a modern, useful, usable metadata standard. A reorganization of rules with increased consistency is great, but is not worthwhile on its own to engender so much work. We strongly agree with Karen Coyle that "if we in the library field do not develop cataloging rules that can be used for this digital reality, we will find once again that non-librarians will take the lead in an area that we have assumed is ours." We more firmly believe now that it is far more important to get out a good standard, rather than one that can be published in a certain year. We doubt that we will likely never again such a golden opportunity for us to have a major positive impact in the broader information community.

As was suggested at Midwinter, some of us feel that we should restructure RDA as a data dictionary and application profiles. Such resources as MARC 21, EAD, VRA Core, and METS provide good models of a data dictionary structure. The data dictionary structure enforces consistency across element descriptions and makes it much quicker to find needed information in the standard. RDA should develop a method for registering application profiles. Each application profile would deal with such aspects as mandatory-ness, repeatability, transcription, data normalization capitalization, abbreviation, etc.), sources of information, level of granularity of a resource, level of specificity of elements (*title* vs. *title proper* + *parallel title* + *variant title* + ...), community-specific examples, etc. All RDA records would be explicitly identified as complying with one or more application profiles.

Currently in RDA there is some implicit hierarchy to data elements: the *title* element is composed of the *title proper, parallel title, variant title*, etc., subelements, for example. This is helpful, because it allows such sections as 2.3.0 to have scope over many particular subelements. There are some anomalies (the *edition* element is composed of the *edition statement, statement of responsibility relating to the edition,* etc., subelements, but the *edition statement* subelement seems to be composed of itself and the *parallel edition statement*), and these should be avoided. We should be more explicit with these hierarchies, because knowing about them will matter when different communities decide the structure/format/syntax for storing and exchanging RDA records. We also need to take more advantage of them, to help us with intrarecord interelement relationships, such as between a publisher, a place, and a date.

We strongly agree with Karen Coyle that RDA is "more about how cataloging was done in the past than how it might be done in the future. This is cataloging done by people sitting at desks with the item in hand, contemplating title pages, covers, prefaces, etc. ... There is also no recognition that many digital libraries expect that in the future much cataloging will have to be done by automated means in order to have any chance of creating metadata for digital resources. ... And we need to create cataloging rules that take into account the reality of machine-to-machine communication and the derivation of data elements by algorithms."

It is time for us to realize that we are not in the cataloging business. We are in the information services business. In the past, our traditional cataloging supported those services. However, the world is a vastly different place than it was when we first standardized cataloging. Metadata is far more generally available than it used to be, and comes in many more varieties of types, formats, sources, extents, and qualities. Administrators in libraries and their parent organizations continue to face limited budgets and increased calls for accountability. We need to ensure that our policies and practices are cost-effective. That is, we need to show that we deliver metadata that is worth more than the cost of producing it. The current path of RDA does not lead to such cost-effective instructions. As a profession, we need to face the new reality and evolve, or we will die. If we don't produce a 21st century standard, others with far less experience with metadata will. As Diane Hillmann has written, "There is no other alternative--we are already late to the table, and we come with a reputation for impatience with those who do not have library training."

One example of an issue for which changes in the world around us necessitate our relooking at the cost-effectiveness of what we are doing is duplicate detection. Some of us feel that duplicate records had a far more negative impact in the past than they do now. We need to consider the possibility of simplifying some rules so that metadata production will be faster and cheaper and able to be created by people other than fully trained catalogers, recognizing that we may generate some duplicate records along the way. It may be more cost-effective to generate 100 metadata records that may duplicate a couple of pre-existing records than only generate 70 records with no duplication.

We feel more strongly than ever that it is folly to believe that RDA can be effectively developed piecemeal. We need a tightly integrated standard, and each part informs the others. There are too many places in this draft that relate to parts 2-3 to be meaningfully evaluated before the other parts are available.

As UC's BSTF report states, libraries "offer a fragmented set of systems to search for published information (catalogs, A&I databases, full text journal sites, institutional repositories, etc) each with very different tools for identifying and obtaining materials. For the user, these distinctions are arbitrary." We need to recognize this as our current environment, and design RDA knowing that we directly control a smaller and smaller portion of the metadata pie.

As UC's BSTF report states, "It can be helpful to think of metadata provision as an ongoing process versus a one-time event." We need to design RDA for a world where metadata gets

created in one place but then gets copied, modified, abridged, expanded, translated, mapped into another system, etc. RDA needs both to produce records that are repurposable (including use by metasearch engines) and to provide for the repurposing of non-RDA data.

Add a new section on the concept of cataloger's judgement, and one on what to do when you are unsure about a situation (such as when you do not know the mode of issuance).

Most of the information in the Introduction to part I and in chapter 1 would seem to have parallels in the other parts, so it seems odd that this type of information is covered in part I. We would go so far as to say that maybe there shouldn't even be introductions for the parts, but rather a single Introduction for the whole document. Also, create a single chapter 0, covering language and script issues, transcription, and formulation of notes, that would apply to all three parts. We have problems currently in AACR2: some topics are covered in the Introduction to part I, but not in the Introduction to part II, and catalogers are left trying to figure out whether they should use the principles of part I in part II. LCRIs for 0.X rules are forced to include instructions for part II because there is nowhere else for those instructions to go. We point out at individual rules various instances of this needed generalization.

RDA needs to deal explicitly with relationships among metadata elements for a single resource, such as parallel titles and their statements of responsibility, and publishers and their places. Structural metadata is needed to indicate such relationships. (This is an example of the disadvantage of separating content from formatting.)

A new section is needed to discuss database management, as more and more of our time is spent modifying existing records rather than working on new ones. The rules need to support this activity as well as they support initial cataloging.

In addition to an alphabetical glossary, and repeating some definitions in the text, it would be quite helpful to also have a hierarchically arranged glossary. For example:

title ...
    title proper ...
    parallel title ...

Strategic Plan, goal 4, 4th bullet says we will repeat definitions from the Glossary in the text. Although we think it is a good idea for users of RDA to have terminology discussed where it is being used, in addition to a glossary, repeating text usually becomes a maintenance nightmare, with some places getting revised, but not others, etc. We hope that for the online version the text will only exist once, but displayed in two (or more) different places.

Some of us believe that continuing the ISBD convention of maintaining differences between information recorded in *statements* and *notes* is not helpful in our modern world. There are metadata standards, such as Dublin Core, that do not even have a concept of *note*. If the library community could itself get out of this mindset of strict distinguishing of notes from other data, it would make it much easier to implement the repeatable 260 field for example. If the concept

of *notes* is retained, create a general section on notes in chapter 1, and treat notes as subelements of each of the other data elements, rather than as separate data elements.

In the past (often for display real estate reasons), we have allowed only one instance of many data elements. This is no longer justified in our modern environment. Even *title* can be repeatable without disaster. One example of how changing our thinking on this will make database management easier is data about previous iterations of integrating resources. If rules said "Retain [information] ... on earlier iterations ..." rather than "Make notes on" would make maintaining this records simpler.

A *Conventions used* section should be added to the (General) Introduction describing such conventions as typefaces, special symbols (such as the diamond introduced in 0.1.7), and bulleting (if used in any special ways).

Objectives and principles document

We think the time spent on developing explicit objectives and principles has been quite fruitful. However, we agree with Diane Hillmann that "discussion has gotten quickly to the nitty-gritty, without spending sufficient time on the principles."

The intended difference between *objectives* and *principles* remains unclear.

1, Objectives, Adaptability: It is not clear what this objective means.

2, Objectives, Responsiveness to users needs: This objective needs to be weighed much more heavily than Representation or Attribution. We need to be user/service-centered first and foremost, and resource-centered a much more distant second.

2, Objectives, Responsiveness to users needs: If this is truly an objective that we want to achieve, then we need to allow for many more data elements (which of course can be optional). Examples include cover art for books, reviews, publisher blurbs, excerpts, color of the carrier, research methodology, disciplinary theory followed (such as relational grammar, lexical-functional grammar, government and binding theory, arc-pair grammar, etc., in linguistics), and source of funding.

2, Objectives, Responsiveness to users needs: As a basic requirement, FRBR has *find* not just for works and expressions, but also for manifestations. This objective should also include locating manifestations.

2, Objectives, Continuity: As we already have to deal with many non-AACR records (both pre-AACR2 as well as those based on other standards), we think objective should be relatively low on the totem pole.

2, Principles, Differentiation: We should strive for differentiation not just for entities in our local files, but also in broader national and international files.

2, Principles, Relationships: It is unclear how these two principles (for descriptive data and access point data) differ. Are they intended to be the same, except for type of data? If so, let's merge into a single sentence.

2, Principles, Representation: We need to become more user-focused, rather than resource-focused. Also, this purpose will likely be difficult for other metadata communities to understand.

2, Principles, Attribution: We do not serve our users well by knowingly using inaccurate data in the main part of the description. Accurate data should be used for title proper, etc., with inaccurate source data transcribed in notes, etc.

2, Principles, Common practice: It is unclear what this sentence means.

This is how we see the draft of part 1 relative the stated objectives and principles, on a scale of 1 to 5:

1. Objectives and Principles for the Design of RDA
    - Objectives
        - Comprehensiveness = 3
        - Consistency = 4
        - Clarity = 3
        - Rationality = 2
        - Currency = 3
        - Compatibility = 3
        - Adaptability = [not sure what this one means]
        - Ease and efficiency of use = 3
        - Format = [not yet clear]
    - Principles
        - Generalization = 4
        - Specificity = 3
        - Non-redundancy = 1
        - Terminology = 2
        - Reference structure = 1
2. Functionality of Records Produced Using RDA
    - Objectives
        - Responsiveness to user needs = 3
        - Cost efficiency = 2
        - Flexibility = 4
        - Continuity = 5
    - Principles
        - Differentiation = [can't say without parts 2-3]
        - Sufficiency = 4
        - Relationships = 3

Representation = 5
Accuracy = 5
Attribution = [parts 2-3 issue]
Common usage = 4 [but we haven't reviewed chapter 3 or the GMD/SMD report yet]
Common practice = [can't say without parts 2-3]
Uniformity = 5

Draft of part I
What is the title of part I?

We very strongly support the separation of metadata vs. its presentation, and focusing RDA on metadata proper, although we are concerned about how various data for one element will display. There are a number of rules that still deal with presentation topics, however, such as square brackets, capitalization (beyond normal English, etc., style), the hyphen in numbering, the question mark in publication (etc.) dates, copyright symbols, parentheses around standard number qualifiers, and punctuation in dissertation information.

This draft is written more tightly than what we saw last year. However, the document still suffers from textual bloat--it is much longer than it needs to be. The primary (web) version of RDA will not be improved by redundancy. Other better ways to provide context and help navigation were discussed at the ALA Editions focus groups. We point out several instances of superfluous text at individual rules.

This draft is till too centered on print, text, and resources in the West European/North American publishing tradition; we point out specific examples below. It is perhaps wise to even go beyond format and type of content neutrality, to putting more effort into types of resources that even digitally in full form are less easy for users to find. Full text can be indexed and searched relatively well compared with musical sound recordings, moving image materials, photographs, etc. Perhaps it is time to be focused more on non-textual resources.

We prefer the meaningful chapter numbering in the draft of AACR3 to what is in this draft. It also allows for easier addition of new chapters.

The cataloging community's use of *description*/*descriptive* does not match the use of those terms by other library folk, nor by the larger metadata community. For example, the NISO document *Understanding Metadata* defines *descriptive metadata* as "metadata that describes a work for purposes of discovery and identification, such as creator, title, and subject." Also, we currently use *description* in two different ways: a set of descriptive data for a resource, and the normal dictionary definition. Context does not always provide sufficient evidence to quickly and easily decide which meaning is intended. The seeming non-existence of substitute terminology is additional evidence that our distinction between subject and non-subject descriptive metadata is not useful (We have additional comments on this point at 0.1.0.). We propose that we observe the following usage:

descriptive: the sense described in the NISO definition above

description: normal English usage
record: the set of data that describes a resource

If rule numbering is deemed useful (before it always seemed to be, but now is not clear if it really is), number every level of hierarchy. For example, in 2.3.2, give the sections *Choosing the title proper*, *Recording the title proper*, and *Change in the title proper* numbers. Yes, this adds levels to rule numbers, but it provides for a more realistic structure. These "hidden" levels of hierarchy in AACR2 in Cataloger's Desktop cause problems--they are missing in TOC displays, links that apply to the whole section appear at the end of the previous rule, and therefore are very easy to miss, etc.

A style sheet would produce a more readable and consistently written document. I've pointed out many examples of lack of readability and consistently in our specific comments below. Here are a couple general examples:

Pick either "source of information" or "source" as the wording to use, and use that as much as possible.
Do this also with "numbering" and "enumeration or alphabetic designation".

The philosophy for the top-level structure of part I (other than chapter 1) is unclear. 0.1.4. says that "chapters 2-6 each cover a set of descriptive data elements that support a particular user task (e.g., identify or select)". It appears that the relationship between user tasks and part I chapters is:

chapter 2 find and identify
chapter 3 select
chapter 4 select
chapter 5 obtain
chapter 6 all, as they relate to items

What is the advantage of chapters 3 and 4 being separate, rather than being combined to form a single chapter for *select*. We realize that FRBR section 7 explicitly does not include user requirements related to items; nevertheless we would benefit from splitting this chapter up by user task, and moving those pieces to the relevant user-task chapter. If this is not done, the quotation above needs to be reworded to be accurate.

We're also concerned that RDA still requires us to base the description of a reproduction on the reproduction, rather than on the original. Since it is a long-standing practice that US libraries will likely continue to follow regardless of what RDA says, and for good reason, describing a reproduction based on its original needs to at least be an option. Evidence shows that users primarily want and need the data about the original publication, not the specifics on when it was digitized or microfilmed. Another option would be to follow FRBR more closely and either create descriptions for both and treat them as related resources, or in one description record data for both, with each data element labelled as to which it belongs to.

It is more helpful to organize and label instructions by the situation the cataloger has, rather than by the result of following the instruction. I've pointed out examples in our specific

comments below. One example is 2.3.0.4 and 2.3.0.5 are labelled “Names of persons and corporate bodies” and “Introductory words, etc.” respectively. These would be more helpful if merged, under the rubric “Determining the title”.

One of the major problems we are seeking a solution on from AACR2 is multitype resources. This draft does not yet achieve that (such as in 2.2.1); specifics are in comments below.

It is not clear why wording about options varies, such as:
- if it is considered to be important ...
- ... considered to be of particular importance
- if it is considered to be necessary
- optionally

Standardize the wording, and/or move all of this to 1.4.

Do not use the arrow bullet for general instructions. It reduces readability, loses meaning via ubiquity, is inconsistently utilized, and unnecessarily lengthens the document due to more white space.

Delete all instructions for prescribed punctuation. Rather, treat the pieces of data punctuated as separate data elements, with specified structural metadata as needed. If you retain any prescribed punctuation, centralize all such instructions in one section (probably also keeping such instructions in the elements).

Make clear that certain data elements only apply to particular kinds of resources, such as frequency only applying to resources issued in multiple parts.

Because the word *comprise* can mean either *compose* or *be composed of*, alternate terminology (perhaps *consist* of and *constitute)* would be clearer.

More modernization and dejargonization is needed. Some specific examples of this are described in comments below.

Without the glossary, we cannot fully evaluate part I. In particular, we need to see:
- *transcribe, record, supply, devise*
- *archival resource*
- *facsimile, reproduction* (define *reproduction* broadly enough to include facsimiles; then we can use just one word)

Introduction to part I

It is unclear why some general information is in the introduction, and why some is in chapter 1.

It would be much easier to digest and utilize information currently given in this Introduction if it were split between background information, and actual instructions. It would seem unnecessary, even counter-productive, to give background information rule numbers. The

content that is instructions (0.1.2, some of 0.1.6, and possibly 0.1.8 (intent here is not clear)) would be best placed in chapter 1.

One section that is missing is one on what kind of information to expect at each data element. This is an important part of any well-conceived standard containing elements. See these resources for examples of this:

CONSER Cataloging Manual, Introduction to Part 1, paragraph 5
CONSER Editing Guide, D1. Content, Organization, and Layout
Content Standard for Digital Geospatial Metadata, Organization of the Standard
DCMI Metadata Terms, Section 1. Introduction and Definitions
Encoded Archival Description Tag Library, Tag Library Conventions
Getty vocabularies (Getty Thesaurus of Geographic Names, Art & Architecture Thesaurus, or the Union List of Artist Names), About the [vocabulary], Information in the Record (Fields)
Library of Congress Subject Headings, Components of Entries
MARC 21 format (any), Introduction, Organization of This Document, Components of the Detailed Descriptions
Medical Subject Headings, XML MeSH Data Elements, Key to element information
Thesaurus of Graphic Materials, Introduction, I.C. Structure and Syntax
UDC MRF Database Development and Design, Database structure
UNIMARC Manual : Bibliographic Format, 2. Organization of the Manual
VRA Core Categories, Category attributes
many XML DTDs and schemas

The section headings in 0.1 are inconsistent: 0.1.1 and 0.1.2 include “resource description” while 0.1.0, 0.1.3-0.1.9 do not. Only the heading for 0.1.4 includes “Part I”.

If it is correct that there will be no other content between this Introduction and chapter 1, renumber all the sections of this Introduction from 0.1.X to just 0.X.

The paging of this Introduction does not match that of the rest of the draft.

0.1.0

As it stands now, part I of RDA most definitely does *not* “provides a comprehensive set of guidelines and instructions on recording data to describe a broad range of resources in both analog and digital formats.” As has been pointed out before, it does not contain any guidance on many areas of bibliographic descriptive metadata: not subjects, classification, or shelflisting, and not all the specialized metadata for particular media or types of content. Although this is likely not what was intended, the quotation above comes across as pompous and insulting. The word “comprehensive” needs to be removed if the scope does not change.

The traditional cataloging segmentation of bibliographic data into descriptive data (or descriptive and name/title authority control data) and subject data is an artificial one, that is not shared with the larger metadata community or, we would venture to say, our user communities.

Not covering subject headings, but including *nature and scope of the* content will not make sense to those uninitiated in traditional cataloging. We need a document that looks--at the highest level--at what a whole bibliographic metadata record should be: general philosophy, element list, references to other standards. RDA development has already made such important strides towards this, that it is a shame for JSC to not go that last step. If RDA doesn't provide this top-level perspective, it is likely that others (most likely computer folk, not librarians) will, and RDA will be relegated to a lesser role in metadata circles. This is a golden opportunity for our community, and we should not hesitate to seize it.

0.1.1
FRBR should be covered here.

If one of the reasons we are keeping successive entry cataloging for serials is to be compatible with ISSN cataloging practices, that should be covered here.

0.1.3
Delete the superfluous second and third sentences.

0.1.5
Either use a title such as "Intended use", which would more accurately describe the content of this section, or merge this into 0.1.4, since it is about structure.

Instructional manuals and reference tools serve very different purposes. The tradeoffs caused by trying to be both kinds of document make RDA less successful at being either. We should focus RDA to be a reference tool--a document that establishes a standard. Training material should be developed separately.

0.1.6
This section should probably be deleted; this information is also given at other, more logical places in the document.

If this rule is retained, given that parts 2-3 are not available, we cannot fully evaluate this rule.

The last paragraph is a very long sentence whose meaning is unclear.

0.1.7
It would be helpful to differentiate options that allow additional information to be recorded (such as at 2.3.0.5) from those that would result in *different* data being recorded (such as at 2.3.7.3). From a cooperative cataloging perspective, the latter are a much bigger concern than the former.

0.1.8
Extend this section to include script preferences.

To facilitate broader sharing of records, specify that the cataloger record the language of the description, and each use of data in the language of the catalog needs to be treated as a separate data element so that automatic translation can be applied as records are shared throughout various language communities

0.1.9
The first paragraph should be deleted. All examples should completely conform to RDA. We should not have any examples at one rule that contradict instructions given at another rule.

Delete the superfluous second paragraph. At any example we can give a note to help provide context; this is normal English usage.

Chapter 1. General guidelines on resource description
Move information that is not instructions (this would seem to be 1.0 and 1.1) to an unnumbered introductory section.

1.0
Delete this superfluous section; it is just a textual version of the TOC.

1.1
Change the title to *Types of resources* (moving 1.1.4 out, as described below).

Delete the superfluous first two sentence. At least delete the parenthetical, as it is only makes sense for documents that are primarily used in a linear fashion.

Move the third and fourth sentences to the General Introduction, as they have scope for the whole document.

1.1.1-1.1.4
We applaud the splitting apart of the actual definition from illustrative and exemplifying text for *resource*. Please do the same for the other definitions.

It is unclear why it was decided to present here these particular terms, but not other terms used in the chapter.

Most of the bulleting in these sections is inappropriate. The text at those bullets is not logically subordinate to the first sentence of each section; the bulleting should be removed. The second through fifth bullets in 1.1.1 are indeed subordinate to the current first bullet, and their bulleting is appropriate.

1.1.1
Delete the superfluous first sentence.

The first bullet implies that some terms might be used differently in different parts of RDA. Please do not do that; use each term consistently throughout the whole document. The bullet also is an example of the textual bloat we referred to above. It could be reduced to just the word and its definition without loss of meaning or readability: "**resource**: the entity that forms the center of focus for a resource description".

The definition of *resource* is not helpful; basically circular with the meaning of *description*. What is a resource? What a description describes. What is a description? Information about a resource. As the LOM and Dublin Core folks said at our CC:DA meetings in San Antonio, RDA needs to have a conceptual model of the resources it attempts to describe and give access to. It would seem easiest to use the FRBR group 1 entities. The document would be much tighter and more consistently understood if *resource* were replaced with *work, expression, manifestation,* and/or *item*, as appropriate. RDA needs to engender a common understanding on what it is that it describes. Can a person, for example, be described as a resource? What about other types of records that are currently in the MARC21 Community Information Format?

Resources exist whether or not there is a description of them. For example, the fact that we use a comprehensive description for a conference proceedings does not mean that the individual papers are not resources. This is also arises when discussing other related resources. Also, the phrase "center of focus" seems redundant. The definition could at least be reworded as "an entity that is or could reasonably be the focus of a descriptive metadata record".

Undescribed resources have the same variety of forms as those described; change "The resource described" to "A resource" in each the bullets.

The fourth bullet describes two subtypes of one of the types described in the previous bullet. Either combine into a single bullet, or indent this bullet under the third one.

The fifth bullet phrase "after the fact" begs the question of after what? The sentence is fine without that phrase. And technically someone could assemble a collection before its components are published.

1.1.2

We like that the draft deals explicitly with intangible resources, but we don't think "logical units" will be readily understood by the average RDA user.

Another opportunity for us to be more user-centered is to change the focus of these definitions. We believe that users perceive the content of resources via attributes of work and expression, not attributes of manifestation. We currently distinguish between serials and integrating resources based on how the manifestations are issued. Our users would be better served if instead we distinguished on how new content is intended to be used. It doesn't matter to them whether the manifestation of the *Encyclopedia of Associations* is successively issued or integrating; it is the fact that the content is kept up to date that matters.. Redefine as follows (changes underlined):

resource issued in successive parts

change from: a resource that is issued in a succession of discrete parts

to: a resource whose content is issued in a succession of discrete parts

integrating resource

change from: a resource that is added to or changed by means of updates that do not remain discrete and are integrated into the whole

to: a resource whose content is added to or changed by means of updates that do not remain discrete and are integrated into the whole

The first and third bullets include the word "issued" in their definitions. This seems to leave out assembled collections. We need either to define *issued* broadly, or change the wording to something like "issued or assembled".

There is a lack of parallelism in the terms and definitions; I've made specific suggestions below.

first bullet

The definition given in the first bullet would include a 1-issue serial. I'm guessing that was unintentional?

The term described is rather long; *unitary resource* would be shorter.

Different RDA users could reasonably interpret the definition to either include or exclude single-volume non-updating loose-leaf material. The individual leaves may be seen be some as separate physical units. And some of these resources ship with the pages shrink-wrapped separately from the binder, definitely separate physical units. How about changing "physical" to "cohesive" ("cohesive" could include the web "logical units")? Or changing "issued either as" to "issued in a form intended to be used either as"?

second and third bullets

There is still a need for an umbrella term for the types described in the second and third bullets. We suggest that we standardize on *multipart resource*.

The terms in the second and third bullets are unnecessarily non-parallel. Using *simultaneously issued multipart resource* and *successively issued multipart resource* would achieve worthwhile parallelism.

second bullet

Excluding the exemplary parenthetical, the definition reads "a resource comprising two or more physical units or, in the case of an intangible resource, two or more logical units." That would include resources issued in successive parts, which does not seem the intent, given the phrase "issued as a set" in one of the parenthetical examples, and the third bullet. The definition needs to include wording such as "issues as a set" or "together".

RDA for Who?

Diane I. Hillmann
Research Librarian

Cornell University Library

Mar. 12, 2006

When I think about the effort going on now with the evolving Resource Description and Access (RDA), the image that comes to my mind is some old film view of the Paris sewers, where sludge covered workers toil tirelessly and thanklessly to keep the excretions of the body politic moving. Perhaps not the best metaphor (or the best smelling one), but one important part to take from it is that this is all happening somewhat underground, far from the view of most librarians, and it concerns an important part of the infrastructure that we depend upon.

Someone will argue with me about the underground part, pointing out rightfully that the work is now going on far more transparently than in its first phase as AACR3 (when both process and product were roundly rejected by its constituency). As a former cataloger (and former MARBI member), who has now moved into the digital library world, I had resolved to keep my distance, with the hope that I'd argued my last about the minutiae of cataloging. But I kept running up against the stated goal of RDA to be relevant to the metadata world, as well as the traditional cataloging world, and sure as shootin', once I started making snarky comments, I got roped in: I'm now representing the Dublin Core community on CC:DA. Surely this is a far too onerous punishment for my sins!

I went into the process with the understanding that I was going to focus on the "big picture" not the details. This helps me avoid getting into the detailed discussions beloved of catalogers (for which I have not the time or inclination), and also has the benefit of allowing me to retain my reputation as an annoying irritant. As a former denizen of the traditional cataloging world who has emigrated permanently to another planet, I hope to bring a different perspective and worldview to the table. I'm fully reconciled to the notion that this will not increase my popularity ratings.

Due to the aforesaid reluctant involvement, I came to the discussions rather late. I'd heard, of course, some of the startling early criticism that there was not enough change from AACR2 being proposed in the early drafts, with the effort being characterized by some as "rearranging the deck chairs on the Titanic." Given the costs of change, said the critics, let's make sure the change is worth the investment.

The Joint Steering Committee (JSC, keepers of AACR) took notice, went back to the drawing board, and started over.

But still, as I pointed out to the RDA-L mailing list a few weeks ago, there are underlying (and largely unexamined) assumptions made in RDA, most of which emanate from traditional cataloging practices. These assumptions are very dissimilar to those inherent in the world of Non-MARC Metadata (NMM) communities (a term I made up, by the way). Because traditional librarians normally work with resources that are digital versions of published materials, they have difficulty seeing outside that familiar boundary to where the environment is much less comfortable.

I described my concerns with the current drafts as being in the following areas: transcription and specified sources of information, reliance on notes, and multiple versions.

Transcription as Identification

In the world of traditional cataloging and static published resources, consistent transcription is an important method to assure predictable metadata. However, digital resources carry no such assumption of stability—change is part of the package. In that environment, relying on use of consistently transcribed information as the primary method of identifying a resource makes much less sense. Resources in the digital environment are most often identified by a numeric or alpha-numeric string. In traditional cataloging, such identifiers are also used, of course: ISSNs and ISBNs are the most obvious examples, but they are generally not the primary identification of the resource.

As we all know, the current methodologies for identifying digital resources uniquely and unambiguously are still in flux and almost no one is satisfied with the current situation. But whatever the ultimate answer, it will not rely on transcription, nor will decisions about what constitutes a “new” or “different” resource likely be susceptible to the rules defined in AACR2 for editions or versions.

Part of the transcription tradition involves specification of sources of information from which to record. Vital to this approach is the idea that resources have commonly identified and named parts that are similar within a specific category of materials, something that is not generally the case in the digital world. Similarly, notions of whether information comes from the item itself or is supplied from somewhere else are often less important in NMM communities, even those who still deal primarily with physical, published items. In ONIX for example, information

about the author (from the book jacket, reviews, or other marketing sources) is specifically tagged based on the function of the information, and it's often not explicitly descriptive in nature.

Reliance on Notes

Oftentimes, the RDA (like traditional cataloging) herds catalogers to make decisions about what is "primary" or "secondary" and relegates the latter to the notes area. This is a significant problem for many NMM communities, who may either have no place to put these kinds of descriptive "notes" or instead rely on repetition of elements (with or without a notion of order) to capture information of the same kind within a single description, thus focusing more on access issues than descriptive integrity.

In most delivery systems for metadata (including OPACS, it must be noted), only the information in a small number of specified fields is actually displayed to the user (and we know few users actually look at full records). Additionally, because notes can contain so many different categories of information, they may not even be indexed (when they are, only as keywords). For systems using NMM, notes information is even less likely to be displayed, and may indeed be entirely ignored, since its "human-friendly" character makes it useless for machine processing and marginal for access.

Multiple Versions and Reproductions

Many traditional catalogers are unfortunately still trying to make the case for describing an original and a reproduction on the same record, for the sake of "practicality." If FRBR is truly the underlying model for the "new" RDA, it seems to me that this bullet must be bitten firmly and these practices explicitly marginalized within the context of the rules. In an environment where metadata of different formats, created using different rules (or no rules) must be shareable and interoperable, these residual practices keep us all from benefiting from our common enterprise.

The argument usually justifying multiple versions on one record is that most library vendor systems do not display resources with multiple editions and versions acceptably, but responding to this problem by undercutting the usefulness of our data by manipulating it to overcome specific system inadequacies is not the answer. We seem to be in deep denial that the approaches we've used in the past to overcome the inadequacies of our vendor systems are in fact serious kludges, that will not allow us to either share data effectively or take advantage of some of the

FRBR-based approaches to improved display of manifestations and expressions. A great example of what might be possible can be seen in the work OCLC has done with FictionFinder--all based on the notion that there is for the most part sufficient data in the (admittedly MARC) records to associate them more usefully for display than current library vendors now offer.

Interestingly, my post to the RDA-L list flushed out a few lurkers from the archives and museums community who said to me, in essence: “RDA doesn’t reflect the needs of our communities either, not any better than AACR2 did.” Most of these folks are not speaking out publicly, and I suspect that may well be a reflection of the fact that they have been cut out of the conversation so long they’ve given up. It’s interesting to note that in most cases they’ve gone off and created their own content standards, and in the days when such silos were a reasonable and acceptable solution, one could hardly blame them. Now, it’s a different matter—we all need to be thinking more broadly, as the traditional boundaries between what we’re doing dissolve. As Karen Coyle recently pointed out, we can
do plan for these changes ourselves, or someone else will do it for us.

I know that the organizers of the RDA effort are trying personfully to come up with a solution that is acceptable to catalogers and to those whose view is more to the forests. But, unfortunately, most of the conversation molding RDA is happening at the level of the workers under the streets, in hip boots navigating the sludge, who aren’t seeing the warning lights flashing above them: *WARNING, WARNING, the end of the world as we know it is upon us.* I’m afraid that the more we depend on catalogers to build the new RDA, the more likely
we are that it will be used only in traditional library settings—which I think most of us believe are becoming a smaller and more marginal part of what libraries do. If we want something else besides AACR3 by another name, we’d better start paying attention and participating, lest we waste yet another opportunity to move out ahead of change, instead of constantly cleaning up after the parade’s gone by.

Title: Alternative paths to semantic specificity
Identifier: /usage/meetings/2006/04/seattle/semantic-specificity/index.shtml
Source: /usage/meetings/2006/04/seattle/semantic-specificity/index.txt

Shepherd: Tom

[1] /usage/meetings/2006/04/seattle/semantic-specificity/2006-03-13.digest.txt

Is it better to use one broader property and specify its values with vocabulary encoding schemes? Or to use multiple, semantically more specific properties? Is it good practice to use a broad property but restrict the scope of its application by annotating the definition accordingly in an Application Profile?

This issue came up in particular with regard to the Accessibility profile. A digest of the discussion [1] is included in the packet. Let's please re-read this thread and consider, in Seattle, whether this is an area in which the Usage Board could provide further guidance.

2006-03-10

On 10/03/2006, at 6:23 PM, Andy Powell wrote:

I completely agree that the use of controlled vocabs is fine
and to be encouraged. But I'm suggesting that they are used
with a small number of properties (perhaps 4) rather than
with one single uber-property.

So instead of having one property (a4a:adaptability) with
one big controlled vocab (or 4 smaller controlled vocabs)
as I think you are currently suggesting, we should instead
have 4 properties (along the lines of a4a:perceptionMode,
a4a:controlMode, a4a:structuralFeatures, a4a:functionalFeatures
but note that I don't understand this space well enough to
know if these are correctly named), each with an associated
controlled vocabulary.

Is that clearer?

2006-03-10 From Liddy

Sent: 10 March 2006 23:56
To: DC-ACCESSIBILITY@JISCMAIL.AC.UK
Subject: Re: Liddy's comments in the wiki (long and techie comments)

Andy

I think it is a matter of style, usability, etc ....

If we look at what happens with subject, we find huge vocab
lists. In the case of adaptability, IMHO, there is work going
on in many places and some of the things that might need to
be included in metadata now will either be transmitted
automatically in other ways later, some new things will
arise, etc. I personally think they are all just adaptability
attributes. In putting the a4a before them, in a sense I
think you are saying the same thing.

You are saying, I think, that these should be part of an
application profile. I think how one understands the role and
value of application profiles might well be a matter for
debate. More and more elements being useful is not what I am
hearing where I work - people do not want to complete massive
long questionnaires to add a bit of metadata and if there are
4 or more additional elements, I suspect they will not be used.

As for processing - your original objection. If looking for
and finding values for attributes in one or four places makes
the difference, --- I cannot comment on that. I do know that
those who have already implemented this stuff are using a
single element with structured values so they musty be
processing the values somehow?

I would like whatever to be as simple as possible for those
being asked to add metadata, so long as that does not cause
problems for those trying to implement it. I am willing to be
guided on that balance but do want to take account of what I
hear from people who will be writing this metadata.

Re your choice of categories - we have worked with 3 dimensions:
control, display (presentation) and content choice. These are
the dimensions for adaptation for accessibility, as we see
it. So we'd have to think from the beginning again to come up
with the categories you suggest (very hypothetically). I did
group the attributes, as you know, so they would easily be
remembered etc - which is what, in fact, I think of as the
actual role for the groupings of DC elements for me.

Let's hope to hear from others - is it better to be choosing
from one list or four is the question??? Does it have any
implications for implementers that should be noted?

2006-03-11 From Andy Powell

Liddy,
I think you are mixing up the usability of metadata tools with the underlying structure of the metadata. (Actually, I think we all tend to do this at the moment because the quality of metadata user-interfaces tends to be rather poor in many tools). Just because we choose 3 properties in the underlying metadata doesn't mean that tools have to present 3 boxes to the end-user. Tools can choose to present a single list as part of the user-interface, but then partition the end-user selections into 3 metadata fields as necessary.

This cuts both ways of course. As I mentioned, dc:format covers at least three very distinct concepts (file format, physical media and dimensions). So a user-interface designer might choose to present 3 boxes to the end-user, but place all the resulting information into one metadata field.

As an aside, I would argue that dc:format is a good example of poor metadata design - i.e. its not a property that we want to copy!

So the question is *not*:

Is it better to be choosing from one list or four?

because that is a user-interface design question. We are intersted in the underlying structure of the metadata description. The question is more like:

Is it better to structure our metadata using a single very general property with 1 (or 3) vocabularies OR using 3 more specific properties each with a single vocabulary?

I agree that this is a design choice, and as such there are no clear-cut answers.

But I would argue that DCMI tends to lean towards the latter route (more specific properties). For example, DCMI has separated out spatial coverage ("it's about the 15th century") and temporal coverage ("it's about Mexico") from other kinds of topics ("it's about Chemistry") by creating several properties, rather than by simply using dc:subject with several controlled vocabularies (which would have been the alternative approach).

The justifaction for this approach is not easy to document - and as far as I know, DCMI has never tried to write down guidance on where to draw the line between using properties and vocabularies. Two points are worth noting though. Firstly, where applications choose not to use controlled vocabularies, it helps to have used more specific properties rather than very general ones (in order that some sense can be made of the resulting values by remote metadata systems). Secondly, where applications choose to define their own vocabularies, the relationship between any term in the vocabulary and the described resource is clearer (to remote metadata systems that don't know the vocabulary) if more specific properties have been used.

But, as I said above, it's a design choice, and there are arguments in both directions.

I still have a gut-feeling preference for something like

```
<meta name="a4a:controlMode"
      scheme="a4a:ControlCharacteristic"
      content="KeyboardOnlyControl" />
<meta name="a4a:displayMode"
      scheme="a4a:DisplayCharacteristic"
      content="Braille" />
```

rather than

```
<meta name="a4a:adaptability"
      scheme="a4a:AdaptabilityCharacteristic"
      content="KeyboardOnlyControl" />
<meta name="a4a:adaptability"
      scheme="a4a:AdaptabilityCharacteristic"
```

```
      content="Braille" />
```

which is what I think you are suggesting?? But as you can see from the above, I admit that I'm struggling to put that gut-feeling into a coherent argument! :-(

2006-03-12

From: Thomas Baker <tbaker@tbaker.de>
To: Andy Powell <andy.powell@EDUSERV.ORG.UK>
Cc: DCMI Usage Board <dc-usage@jiscmail.ac.uk>
Subject: [DC-USAGE] Semantic specificity options

> The justifaction for this approach is not easy to document - and as far
> as I know, DCMI has never tried to write down guidance on where to draw
> the line between using properties and vocabularies.

All,

Andy's posting to dc-accessibility about "where to draw the line between using properties and vocabularies" addresses a design issue we have acknowledged and discussed several times before, but I also do not recall that we ever wrote down or even ever articulated the problem more clearly than here.

Maybe it would help to give this issue a handle, like "Semantic specificity options". For completeness, those options would need to distinguish between term declarations and application profiles. The options might look something like the following:

1. Instead of using one broader property, use multiple, semantically more specific properties (i.e., declared in term declarations).

2. Use a broad property and specify its values with vocabulary encoding schemes (i.e., declared in term declarations).

3. Use a broad property but restrict its definition, domain, range, or use in an application profile.

Listing the options and discussing advantages and drawbacks of each -- discussing data design versus user-interface design, uncontrolled values versus use of controlled vocabularies -- would be a useful addition to the suite of documents DCMI offers to designers of application profiles.

Tom

Date: Sun, 12 Mar 2006 14:26:24 +0000
Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Semantic specificity options
To: DC-USAGE@JISCMAIL.AC.UK

Thomas Baker wrote:

> 3. Use a broad property but restrict its definition, domain,
>   range, or use in an application profile.

I understand what you mean, but I think we're going to have to be a bit careful with terminology here: in RDFS statements about the domain and range of a property are always global in scope. OWL introduces some constructs for doing restrictions on a per-class basis [1].

This is one of the areas where I think we need a properly worked-out model of a DCAP before we can nail down the guidelines.

Pete

[1] http://www.w3.org/TR/owl-guide/#PropertyRestrictions

Date: Sun, 12 Mar 2006 13:44:26 -0500

Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: "Diane I. Hillmann" <dih1@CORNELL.EDU>
Subject: Re: Semantic specificity options
To: DC-USAGE@JISCMAIL.AC.UK

All:

Well, yes, we did attempt to write this down. In the "old" version of the process document was a decision tree that asked the following of those considering a proposal for a new element:

1. Can the need be solved with a vocabulary encoding scheme for an existing DCMI Element or Element Refinement?=8A
2. Can the need be solved through an application profile that references an element or element refinement from an existing and recognized non-DCMI namespace?
3: Can the need be solved with a new refinement for an existing DCMI element ?
4: Create a new DCMI Element (and, if necessary, Element and Vocabulary Encoding Scheme) to meet the need

I have to admit that when I teach, I express this as the "Big Bucket" approach: generalized elements, specific vocabularies. It's an approach that is far more interoperable than the proliferation of more and more specific elements, and pushes implementors into more investment into controlled vocabularies that can be more responsive to changing needs without the overhead of change at the element/refinement level.

I have to admit that as I'm following the accessibililty discussions it is an approach that suggests itself as I hear the problems defined. =46or instance, if there are several categories of accessibility statements or criteria, one could envision a separate vocabulary for each, which could all be used within the same element. The fact that the vocabularies are separate gives you, to a great extent, the "structured" piece that many are looking for in DCSV, but again, without the overhead or potential for creating messes.

I'm all for writing this down in some way that's useful, but I'd go further and suggest that we recommend, and even STRONGLY recommend, the general buckets, specific vocabularies strategy.

Date: Sun, 12 Mar 2006 19:18:51 +0000
Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Semantic specificity options
To: DC-USAGE@JISCMAIL.AC.UK

Diane I. Hillmann wrote:

> I have to admit that when I teach, I express this as the "Big Bucket"
> approach: generalized elements, specific vocabularies. It's an approach
> that is far more interoperable than the proliferation of more and more
> specific elements, and pushes implementors into more investment into
> controlled vocabularies that can be more responsive to changing needs
> without the overhead of change at the element/refinement level.
>
> I have to admit that as I'm following the accessibililty discussions it
> is an approach that suggests itself as I hear the problems defined. For
> instance, if there are several categories of accessibility statements or
> criteria, one could envision a separate vocabulary for each, which could
> all be used within the same element. The fact that the vocabularies are
> separate gives you, to a great extent, the "structured" piece that many
> are looking for in DCSV, but again, without the overhead or potential
> for creating messes.
>
> I'm all for writing this down in some way that's useful, but I'd go

> further and suggest that we recommend, and even STRONGLY recommend, the
> general buckets, specific vocabularies strategy.

But within the DC model, elements are _not_ buckets: they are
properties, which express types of relationships betweeen resources.

I agree it's fine to use a single property to express the _same_ type of
relationship between one resource and two other resources (values) of
quite different types

e.g. if I want to talk about a concept being the subject of a book, and
a person being the subject of a book, then it is fine to use

book:B has-subject concept:C

book:B has-subject person:P

Yes, I agree we don't want to go coining new properties for
person-as-subject, physical-object-as-subject - well, assuming that the
definition of our has-subject property permits this range of values: for
some properties, the "range" of the property is indeed constrained by
the owner, and if our value is outside that range then in that case we
may need to define a new property.

But (it seems to me) the accessibility example shows a case where not
only are the values of widely different types, but the relationships
between the subject and those values are also of different types.

See Liddy's example here

http://dublincore.org/architecturewiki/MoreCarefullyThought

where the single proposed property a4a:adaptability is used with the
following different value strings (I think Liddy's first statement with
a4a:adaptability should be five separate ones)

visual
auditory
keyboardOnly
structuredPresentation
peerInteraction
textual
replacesVisual

As I said here

http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0603&L=dc-accessibility&P=2167

we have to think in terms of simple statements, and I can not envisage a
coherent definition for a single property which could be used to form
statements in which all of those values are used, apart from one which
says something like:

resource:r
is-related-in-some-undefined-manner-something-to-do-with-adaptability-to
value:v

And I can't understand what use that property is to anyone. If we want
to say something so vague, we could just use dc:relation ;-)

Date: Sun, 12 Mar 2006 21:19:29 +0000
Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Semantic specificity options
To: DC-USAGE@JISCMAIL.AC.UK

Pete Johnston wrote:

> But within the DC model, elements are _not_ buckets: they are
> properties, which express types of relationships betweeen resources.

Also the conversation a few months back about dc:coverage and the
confusion over whether it merged together "about-ness" and
"applicability" (at least on the "spatial" side) highlighted the
problems which arise if the broad bucket approach is applied loosely.

A statement using the dc:coverage property and having a place as value
is more or less useless because I don't know if it means the resource is
"about" the place or "applicable to" the place.

Date: Sun, 12 Mar 2006 17:09:53 -0500
Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: "Diane I. Hillmann" <dih1@CORNELL.EDU>
Subject: Re: Semantic specificity options
To: DC-USAGE@JISCMAIL.AC.UK

>Pete Johnston wrote:
>
>>But within the DC model, elements are _not_ buckets: they are
>>properties, which express types of relationships betweeen resources.

Yes, I understand that, but for most people the notion of
"properties" is not particularly helpful in a training context. In my
experience, it's important to deal with where folks are, not where
you wish they were, particularly when attempting to teach them to do
something new.

>Also the conversation a few months back about dc:coverage and the
>confusion over whether it merged together "about-ness" and
>"applicability" (at least on the "spatial" side) highlighted the
>problems which arise if the broad bucket approach is applied loosely.

But this implies that we are entirely in control about how people
understand and use metadata definitions, which we certainly aren't.
This is not to say that we shouldn't try to clarify some of the
ambiguity of the past (and avoid it in future), but we do need to be
realistic about what to expect from that exercise. Even if we had
gone the other way on that decision, I don't think it would have
materially changed what people did with the element.

>A statement using the dc:coverage property and having a place as
>value is more or less useless because I don't know if it means the
>resource is "about" the place or "applicable to" the place.

It is only useless if you require and expect the metadata you receive
from others to be entirely predictable and consistent. My experience
is that it's usually neither, so I've adjusted my expectations
accordingly. I'd be happy if all the providers I worked with managed
to put place names in Coverage (bonus if they spell them correctly!)
and used an identifier that could get me to the resource.

Your mileage may vary, of course ... ;-)

Date: Mon, 13 Mar 2006 07:36:41 +0000
Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Semantic specificity options
Comments: To: "Diane I. Hillmann" <dih1@cornell.edu>
To: DC-USAGE@JISCMAIL.AC.UK

Quoting "Diane I. Hillmann" <dih1@cornell.edu>:

> Yes, I understand that, but for most people the notion of
> "properties" is not particularly helpful in a training context. In my
> experience, it's important to deal with where folks are, not where
> you wish they were, particularly when attempting to teach them to do
> something new.
>
>> Also the conversation a few months back about dc:coverage and the
>> confusion over whether it merged together "about-ness" and
>> "applicability" (at least on the "spatial" side) highlighted the
>> problems which arise if the broad bucket approach is applied loosely.
>
> But this implies that we are entirely in control about how people
> understand and use metadata definitions, which we certainly aren't.

The issue here isn't - in the first instance - how implementers
(mis)use "our" terms: it's the principles "we" (the UB, WGs, other

designers of DCAPs) ourselves apply to constructing/creating/defining them. At this point we _do_ have control.

> This is not to say that we shouldn't try to clarify some of the
> ambiguity of the past (and avoid it in future), but we do need to be
> realistic about what to expect from that exercise. Even if we had
> gone the other way on that decision, I don't think it would have
> materially changed what people did with the element.

I can only guess! But if there had been two properties that separated out notions of "aboutness" and "applicability/validity", supported by clear documentation and examples, and appropriate machine-processable descriptions, it seems to me that there is a chance that they would have been used as intended. And - the bottom line - my consuming application would have been able to distinguish a course about Bristol from a course that is applicable/available only to residents of Bristol.

>> A statement using the dc:coverage property and having a place as
>> value is more or less useless because I don't know if it means the
>> resource is "about" the place or "applicable to" the place.
>
> It is only useless if you require and expect the metadata you receive
> from others to be entirely predictable and consistent. My experience
> is that it's usually neither, so I've adjusted my expectations
> accordingly. I'd be happy if all the providers I worked with managed
> to put place names in Coverage (bonus if they spell them correctly!)
> and used an identifier that could get me to the resource.

OK, but we have to start from the position of defining properties that we believe are

(a) formulated in a way which is consistent with the DCAM; and
(b) as fully, clearly and unambiguously defined as possible (both in terms of human-readable documentation and machine-processable assertions about relationships with other terms) (c) useful for the "functional requirements" they set out to address

I accept that in practice people will use DCMI terms in ways we hadn't predicted, and ways that result in nonsensical inferences. But I don't think we should treat that as a design consideration (except as something we have to do our best to avoid).

I don't understand an approach that places so much emphasis on not defining new properties because it's bad for interoperability. If an application needs to model a relationship type that is not covered adequately by an existing property, they need a new property. In terms of interoperability, that is a _better_ solution than "stretching" an existing property in ways that were never intended in the name of "reuse". It seems to me that it's the latter approach which is giving metadata consumers so many problems and damaging interoperability.

A consuming application encountering statements using an "unknown" property can then choose either to ignore statements or to make use of information about that property to infer other statements, which may result in "useful" data, referencing "known" properties. That is a better position for the consumer than having to deal only a small set of "known" properties, but having to grapple with the fact that they have been used in unpredictable and inconsistent ways.

> Your mileage may vary, of course ... ;-)

Going back to Andy's initial point, I do agree that the design choice between single propery/multiple vocab encoding schemes & multiple properties isn't clear cut at all.

However given the vocab encoding schemes currently proposed for the adaptability property, I can see no useful single property that can be defined to describe relationships between a resource and instances of _all_ those classes.

With a different set of vocabulary encoding schemes - derived from a different approach to modelling the problem space - that situation might be different. But the choice of which properties/classes required should be based on a modelling of the problem space which meets the functional requirements of the application, not on a principle of trying to "squeeeze everything in to" a single property.

Date: Sun, 12 Mar 2006 11:21:41 +0000
Reply-To: DCMI Accessibility Group <DC-ACCESSIBILITY@JISCMAIL.AC.UK>
Sender: DCMI Accessibility Group <DC-ACCESSIBILITY@JISCMAIL.AC.UK>
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Liddy's comments in the wiki (long and techie comments)
Comments: To: Andy Powell <andy.powell@eduserv.org.uk>
To: DC-ACCESSIBILITY@JISCMAIL.AC.UK

Quoting Andy Powell <andy.powell@eduserv.org.uk>:

> I think you are mixing up the usability of metadata tools with the
> underlying structure of the metadata. (Actually, I think we all tend to
> do this at the moment because the quality of metadata user-interfaces
> tends to be rather poor in many tools). Just because we choose 3
> properties in the underlying metadata doesn't mean that tools have to
> present 3 boxes to the end-user. Tools can choose to present a single
> list as part of the user-interface, but then partition the end-user
> selections into 3 metadata fields as necessary.

I strongly agree with this: designing a DCAP is not the same thing as
designing a user interface.

[snip]

> Is it better to structure our metadata using a single very general
> property with 1 (or 3) vocabularies OR using 3 more specific properties
> each with a single vocabulary?
>
> I agree that this is a design choice, and as such there are no clear-cut
> answers.

[snip]

> But, as I said above, it's a design choice, and there are arguments in
> both directions.
>
> I still have a gut-feeling preference for something like
>
> <meta name="a4a:controlMode"
> scheme="a4a:ControlCharacteristic"
> content="KeyboardOnlyControl" />
> <meta name="a4a:displayMode"
> scheme="a4a:DisplayCharacteristic"
> content="Braille" />
>
> rather than
>
> <meta name="a4a:adaptability"
> scheme="a4a:AdaptabilityCharacteristic"
> content="KeyboardOnlyControl" />
> <meta name="a4a:adaptability"
> scheme="a4a:AdaptabilityCharacteristic"
> content="Braille" />
>
> which is what I think you are suggesting?? But as you can see from the
> above, I admit that I'm struggling to put that gut-feeling into a
> coherent argument! :-(

I dunno if this helps or not, but I think we need to remember that the
reason we coin properties is to use those properties to make statements
about things "out there in the world". Each of those statements says

resource-1 is-related-in-some-specified-way-to resource-2

So in the examples above, we have three resources:

(a) the document, (b) the concept of control only by keyboard, (c) the
Braille format

We want to make statements relating (a) to (b) and relating (a) to (c).

Whar is the relationship between (a) and (b)? What statement do I want
to make with (a) as subject and (b) as object value? What is the
"verb"? Something like "can be controlled using"?

Similarly, what is the relationship between (a) and (c)? What statement

do I want to make with (a) as subject and (c) as object value? What is the "verb"? I don't think it is the same as for (a) and (b). I think it is something like "can be displayed using"?

Now sure, we can coin a property that captures both ("can be dsplayed or controlled using"), but the more different relationships we try to collapse into one property ("can be displayed or controlled using or has a supporting tool", "can be displayed or controlled using or has a supporting tool or (some other factor related to adaptability)"), the more general and imprecise the property becomes.

In any one particular use of the property I can't tell whether the statement is intended to convey a "can be displayed using" or "can be controlled using" relationship. I'm left to guess from the type of the value what the type of the relationship is. (Ah, the value is a class of tool, so the realtionship must be "has supporting tool", not "is controlled using" etc). (As Andy says, we have exactly this problem with dc:format, and probably also with dc:coverage)

As you collapse more different relationships into a single property, you eventually end up with something not far from dc:relation - which says only that there is some unspecified relationship - here something to do with adaptability - between (a) and (b). And (it seems to me) people very rarely want to say only that: they want to express a specific type of relationship.

Subject: RE: [DC-USAGE] Semantic specificity options
Date: Mon, 13 Mar 2006 15:50:13 -0000
From: "Andy Powell" <andy.powell@eduserv.org.uk>
To: "Thomas Baker" <tbaker@tbaker.de>
Cc: "DCMI Usage Board" <dc-usage@jiscmail.ac.uk>

> 3. Use a broad property but restrict its definition, domain,
> range, or use in an application profile.

3 is an area where we need to be careful. I don't think we should ever talk about restricting a property definition as part of an application profile. We can talk about annotating the wording of the definition to make its usage clearer in the context of an application - but we should not talk about changing the definition.

Essentially, we should think of an application profile as adding an additional application-specific comment.

Date: Mon, 13 Mar 2006 11:13:23 -0500
From: "Diane I. Hillmann" <dih1@CORNELL.EDU>
Subject: Re: Semantic specificity options
To: DC-USAGE@JISCMAIL.AC.UK

Andy:

I think you're right on this. I've noted that some people setting up APs have a tendency to change the wording in definitions from "resource" to "something-else-that-they're-interested-in" which has the effect of restricting the definition. I've tried to convince the ones I've worked with to leave the defs as is and add a comment that accomplishes essentially the same thing.

Date: Mon, 13 Mar 2006 16:18:12 -0000
From: Andy Powell <andy.powell@EDUSERV.ORG.UK>
Subject: Re: Semantic specificity options
To: DC-USAGE@JISCMAIL.AC.UK

> Well, yes, we did attempt to write this down. In the "old"
> version of the process document was a decision tree that
> asked the following of those considering a proposal for a new element:
>
> 1. Can the need be solved with a vocabulary encoding scheme
> for an existing DCMI Element or Element Refinement?=A9 2. Can
> the need be solved through an application profile that
> references an element or element refinement from an existing
> and recognized non-DCMI namespace?
> 3: Can the need be solved with a new refinement for an
> existing DCMI element?

> 4: Create a new DCMI Element (and, if necessary, Element and
> Vocabulary Encoding Scheme) to meet the need

Points 1 and 2 of these guidelines help in those cases where a WG is adding something that is closely related to stuff that already exists in DC. But they don't help much where WGs are moving into semantically new areas. In such cases one just falls thru to 4, which doesn't provide guidance on the 'property' vs. 'vocab' balance.

> I have to admit that when I teach, I express this as the "Big
> Bucket" approach: generalized elements, specific
> vocabularies.

Hmmm... interesting. That's pretty much the opposite of what I've been arguing on the dc-accessibility list :-(

On the other hand, I think we have a sliding scale here... what does "generalised elements" mean in practice. One logical conclusion of your argument is that we should only have one "general element", dc:description, and do everything else with "specific vocabularies". But we don't do this... and the reason we don't is because not everything can be expressed thru controlled vocabularies. If we just had

dc:description = text
dc:description = plumbing
dc:description = bath
dc:description = bristol

we've lost a significant part of the information carried by

dc:type = text
dc:subject = plumbing
dc:coverage = bath
dc:coverage = bristol

So, we do choose appropriately narrow elements - the question is, what is appropriately narrow!? :-) I think part of the answer lies in how far we can reasonably expect the values of any newly proposed property to come from a controlled concept space (a controlled vocab) and how far we expect the values to be less structured statements.

Very braod buckets, with unstructured statements as values do not lead to any interoperability.

There's a tension in the adaptability proposal because the new element is defined as "a statement ..." but all the examples provided tend to look like controlled vocabs. Part of the difficulty with the proposal as it stands is that the extent to which the expected values are statements is not very clear.

> I'm all for writing this down in some way that's useful, but
> I'd go further and suggest that we recommend, and even
> STRONGLY recommend, the general buckets, specific
> vocabularies strategy.

Yes, I think I probably agree with this. But only

- where controlled vocabs are a possibility

- where the norm is that values will be taken from vocabs

- where any proposal for a general property is accompanied by a reasonably full set of proposed vocabs

- where the relationship between the resource being described and *all* the suggested vocab terms can be expressed in a reasonable way

---

Date: Mon, 13 Mar 2006 17:19:09 +0000
Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Re: Semantic specificity options

Quoting "Diane I. Hillmann" <dih1@cornell.edu>:
> I think you're right on this. I've noted that some people setting up

> APs have a tendency to change the wording in definitions from
> "resource" to "something-else-that-they're-interested-in" which has
> the effect of restricting the definition. I've tried to convince the
> ones I've worked with to leave the defs as is and add a comment that
> accomplishes essentially the same thing.

I think it's fine for a DCAP to include a literal where "resource" is
replaced by "something-else-that-they're-interested-in". The fact that
it looks like what DCMI calls a "definition" rather than a "comment"
doesn't matter: it is still just an application-specific
comment/annotation, and doesn't change the "global" DCMI-provided
definition.

Date: Mon, 13 Mar 2006 17:17:24 +0100
From: Thomas Baker <tbaker@tbaker.de>
To: Andy Powell <andy.powell@eduserv.org.uk>
Cc: Thomas Baker <tbaker@tbaker.de>,
DCMI Usage Board <dc-usage@jiscmail.ac.uk>
Subject: Re: [DC-USAGE] Semantic specificity options

On Mon, Mar 13, 2006 at 03:50:13PM -0000, Andy Powell wrote:
> > 3. Use a broad property but restrict its definition, domain,
> > range, or use in an application profile.
>
> 3 is an area where we need to be careful. I don't think we should ever
> talk about restricting a property definition as part of an application
> profile. We can talk about annotating the wording of the definition to
> make its usage clearer in the context of an application - but we should
> not talk about changing the definition.
>
> Essentially, we should think of an application profile as adding an
> additional application-specific comment.

If the definitions are owned by DCMI, then arguably, nobody can
really "change the definition" (except DCMI). But "annotation"
is what I meant so I'm happy to refer to it only as annotation
if we think this will avoid confusion.

"Use a broad property but restrict the semantic scope of
its application in a particular context by annotating its
definition in an application profile." Is that better?

Date: Mon, 13 Mar 2006 17:03:19 +0000
From: Pete Johnston <p.johnston@ukoln.ac.uk>
To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@jiscmail.ac.uk>,
Thomas Baker <tbaker@tbaker.de>
Cc: DC-USAGE@jiscmail.ac.uk
Subject: Re: Semantic specificity options

Quoting Thomas Baker <tbaker@tbaker.de>:

> On Mon, Mar 13, 2006 at 03:50:13PM -0000, Andy Powell wrote:
>> > 3. Use a broad property but restrict its definition, domain,
>> > range, or use in an application profile.
>>
>> 3 is an area where we need to be careful. I don't think we should ever
>> talk about restricting a property definition as part of an application
>> profile. We can talk about annotating the wording of the definition to
>> make its usage clearer in the context of an application - but we should
>> not talk about changing the definition.
>>
>> Essentially, we should think of an application profile as adding an
>> additional application-specific comment.

That is the approach I argued for at some point on DC collections.

In the DC CD AP, we are currently using "DC CD AP Usage" (rather than
"DC CD AP Definition"). We do have both a "DC CD AP Usage" and a "DC CD
AP comment", though I regard both as supplementing/annotating the DCMI
definition.

> If the definitions are owned by DCMI, then arguably, nobody can
> really "change the definition" (except DCMI).

This is what I _want_ to be true, and I've tried to approach DCAPs as

if this is the case.... but I do seem to recall a conversation with Dan
and Charles on top of a skyscraper in Shanghai in which they argued
that "definitions" and "meaning" were _always_ "social" and shifting
and context-sensitive.

At that point I covered my ears and started singing loudly to myself
and clutched my security blanket tightly, least my carefully
constructed world fall apart around me.

> But "annotation"
> is what I meant so I'm happy to refer to it only as annotation
> if we think this will avoid confusion.
>
> "Use a broad property but restrict the semantic scope of
> its application in a particular context by annotating its
> definition in an application profile." Is that better?

Title: DCMI Documentation Roadmap
Identifier: e:/admin/admin/www/usage/meetings/2006/04/seattle/docn-roadmap/index.shtml
Source: e:/admin/admin/www/usage/meetings/2006/04/seattle/docn-roadmap/index.txt

Discussion leader: Tom

[1] /usage/meetings/2006/04/seattle/docn-roadmap/2006-03-31.roadmap.pdf

Tom originally prepared this Documentation Roadmap for
discussion with the DCMI Board of Trustees in Madrid.
It will be revised in light of ongoing progress for the
Trustee meeting in Mexico. It could not be discussed by the
Usage Board in Madrid because the UB met before the Trustees.

In Seattle, I would like for us all to take a critical look
at the Roadmap. Are there major items missing? Or have
some work items become unnecessary? Does it identify the
right priorities?

Roadmap for DCMI Documentation

Contents

Introduction

The Dublin Core Metadata Initiative (DCMI) has evolved over the past ten years as a community of developers. At ten years, DCMI finds itself with a stable set of metadata terms ("the Dublin Core" and related vocabularies) and a coherent, Semantic-Web-compatible data model ("DCMI Abstract Model"). Among DCMI's participants and stakeholders there is a widespread feeling that DCMI should now shift its focus away from the incremental expansion of the DCMI vocabularies and towards the practical and interpretive issues raised by the real-world deployment of those vocabularies in application profiles.

This report takes a fresh look at the documentation produced by DCMI from the standpoint of Dublin Core users and describes a path towards accomplishing such a shift. Part One considers DCMI documentation from the standpoint of users, describing materials needed by Dublin Core implementers and, more generally, by organizational decision-makers. Part Two looks behind the scenes of DCMI's public presentation and examines the

documentary tasks and workflows involved in the ongoing maintenance of DCMI vocabularies and specifications.

1 User-oriented documentation

1.1 High-level overviews

1.1.1 Dublin Core from 10,000 feet — the five-pager

There is a need for a document which explains the Dublin Core Model in simple terms for ordinary users, potential implementers, and decision-makers. This story would need to cover:

- What metadata is and what it is used for. Reliable information about a resource — who created it and when — provides a basis for managing information throughout its life-cycle, for finding the information with accuracy, and for sorting or filtering the information once found.
- Why "interoperability" of metadata is important and what it means in a practical sense. For example: that descriptive information collected from many sources can be merged into a common view in a portal.
- The Dublin Core philosophy: that application profiles use both DCMI and non-DCMI elements, as needed, to create implementation-specific metadata on the basis of an underlying model for interoperability. In particular:
  - The DCMI Abstract Model shares a grammar of properties and classes with other Semantic Web applications and is designed to facilitate the integration of data from multiple sources.
  - Building on core elements, implementors can use the model to build application profiles for specialized implementations, flexibly extending the implementation model in principled ways with additional modules as needed.
  - Featured application profiles [1.3.3] illustrate the use of Dublin Core in practice.

1.1.2 Dublin Core from 30,000 feet — the one-pager

There is a need for a well-written, carefully vetted one-page summary of the above for use in boardrooms.

1.1.3 Why metadata? — the boardroom presentation

There is a need for overhead presentations to introduce Dublin Core to a range of audiences, from Webmasters and decision-makers. This presentation could be featured on the DCMI home page for the benefit of people who prefer to get an overview from presentation slides.

1.2 The Dublin Core model

Since the invention of qualifiers in the late 1990s and the parallel rise of data-modeling activities in the World Wide Web Consortium ("Resource Description Framework"), the Dublin Core grammar has matured into a Semantic-Web-compatible Abstract Model and stands alongside the element set itself as a defining feature of Dublin Core.

1.2.1 The DCMI Abstract Model, extended

The DCMI Abstract Model (DAM) has withstood close scrutiny since its publication in March 2005 as a DCMI Recommendation, though some minor points have been identified for possible revision or extension. Specifically, the DAM does not yet formally define the notion of an "application profile." This remains to be done somewhere, whether as a separate document or as part of a revised DCMI Abstract Model.

1.2.2 Clarifications of the model

The use and application of the Abstract Model continues to raise issues of interpretation. These issues are being addressed in a growing category of "model-related" notes such as "Element refinement in Dublin Core metadata" [1] [2.1.4].

1.2.3 The Abstract Model made simple

Understanding the DCMI Abstract Model (DAM) as currently written presupposes that the reader knows something about data modeling. A document is needed which describes the abstract model in a more intuitive (if less rigorous) manner.

1.3 Featured application profiles

1.3.1 Simple Dublin Core

"Simple Dublin Core" has long been defined as the classic fifteen elements of ISO 15836, used exclusively with string values. In terms of the DCMI Abstract Model, the use of these restrictions means that Simple Dublin Core is, in effect, an "application profile," though it has hitherto never been defined and presented as such. Presenting it as such would help clarify, among other things, how the simple schema mandated by the Open Archives Initiative Protocol for Metadata Harvesting relates to other types of Dublin Core–based records.

1.3.2 Minimal Dublin Core

It has long been recognized that there is a need in many implementation scenarios for descriptive metadata even less detailed than the fifteen elements of the ISO 15836 Dublin Core. The emphasis in the DCMI community on the extensibility of Dublin Core has sometimes obscured the point that, in many cases, it could be enough to specify just (for example) Title, Contributor, Type, Date, and Description. A short profile for "Minimal Dublin Core" (a "Kernel") could make this point quite effectively.

1.3.3 Good examples of application profiles in practice

DCMI should provide good examples of application profiles — profiles which have been subject to qualified review and found to be "conformant" with the Dublin Core model [2.1.2]. Featured application profiles could be presented on a DCMI Web page along with a description of their implementation contexts and, when applicable, with review texts commenting on interesting or unusual features of the profiles [2.3.1]. Application profiles to be reviewed could be drawn from research projects, institutional or corporate implementations, or from DCMI working groups.

1.3.4 Record templates for featured application profiles

Ideally, a featured application profile should be linked to templates for metadata produced in accordance with that profile — for example, the XML schema used to validate metadata records. Such templates would constitute good-practice models directly usable by systems vendors in developing applications.

1.4 Practical guidance for implementers

1.4.1 How to implement an application profile

A document is needed for describing the DCMI notion of an application profile — its motivation (interoperability on the basis of a model), its role (to declare restrictions not present in the standard itself), and its form (documentation of how Dublin Core and other vocabularies are being used to describe information in a specific context). Specifically, it should describe:

- How to build an application profile: the process by which an implementor meets common descriptive requirements in DAM-conformant ways.
- How to present an application profile: what information should be documented in a profile, along with a simple tabular format for its presentation [2.3.1].
- How to evaluate an application profile in terms of conformance to the DCMI Abstract Model, with clearly explained checklists [2.1.4].

1.4.2 How to declare a metadata vocabulary

By design, a Dublin Core Application Profile should only use terms which have already been declared somewhere, whether by a vocabulary maintenance organization such as DCMI or by a metadata-using project or organization. To meet this requirement, implementers need to know how to first declare any new terms they have defined in order then to be able to cite them in their own application profiles. There are several aspects to this:

- How to assign identifiers to new (local) properties [2.1.4]. A draft "Guidelines for assigning identifiers to metadata terms" [17] addresses a key aspect of the problem and should be reviewed for DCMI status.
- How to publish terms — how to publish locally declared terms as a Web page, an RDF schema, or both, with advice about maintaining such documents [2.2.4].
- How to identify a controlled vocabulary as a DCMI encoding scheme. (Note: DCMI has in the past considered various ways of opening DCMI namespaces to the "registration" of controlled vocabularies with DCMI-maintained URIs for use in metadata records as encoding schemes but abandoned this approach in favor of encouraging implementers to coin and maintain URIs for their own vocabularies [2.1.4].)

- How to maintain a controlled vocabulary. The document on publishing terms could point readers to related work on publishing thesauri and other controlled vocabularies using, for example, the RDF vocabulary "SKOS Core" [19].

1.4.3 How to encode Dublin Core metadata in XML

The DCMI Recommendation for embedding Dublin Core metadata in HTML and XHTML using the META and LINK elements is currently regarded as stable [6], though recent developments such as the draft W3C specifications for GRDDL [18] and XHTML 2.0 [20] may require that this recommendation be revised in the medium term.

The recommended guidelines on implementing Dublin Core in XML [2], however, need to be revised in order to clarify their support for specific features of the Abstract Model.

1.4.4 How to encode Dublin Core metadata in RDF/XML

DCMI currently has a Recommendation for expressing "Simple Dublin Core" in RDF/XML" [3] and a Proposed Recommendation on expressing "Qualified Dublin Core" in RDF/XML [7]. For historical reasons, these two specifications take different approaches to modeling Dublin Core descriptions [15] — i.e., limits values to strings, while the latter considers values as resources in themselves (i.e., as "nodes") — with subtle implications for interoperability.

The evaluation of both specifications in light of the DCMI Abstract Model has recently been undertaken by a task force of the DCMI Architecture Working Group, and the emerging consensus seems to be that both of the above recommendations be superseded by one consolidated specification. The resolution of this issue would directly affect how DCMI expresses its own metadata term vocabularies in RDF [2.2.3].

1.4.5 Frequently Asked Questions

The FAQ currently posted on the DCMI Web site no longer accurately reflects the questions which are in fact frequently asked about Dublin Core [8]. Questions continually get asked on DCMI mailing lists or in the AskDCMI service, and in most cases the questions do get answered — often in tightly-argued, FAQ-sized paragraphs. However, there currently exists no mechanism for capturing these answers, incorporating them into a FAQ, and assigning each answer a URI for the purposes of citation. As answers often involve fundamental issues of interpretation, they should ideally be subject to periodic

review and revision. For the FAQ to remain fresh, each answer should ideally be reviewed by qualified experts at least once per year.

1.4.6 Introductory texts and tutorial materials

Tutorials are offered every year at the annual DCMI conference and should continue to be a primary source of presentation materials about Dublin Core. These, along with other Powerpoint presentations, training materials, and project manuals, are currently made available on the Web page "Metadata training resources" [9].

1.5 Distribution channels

1.5.1 User-oriented Web interface

The DCMI home page is currently very well designed for use by active participants of the DCMI community — working group members and metadata experts. However, there is a recognized need for an additional Web page designed primarily for interested non-experts ranging from potential implementers to boardroom decision-makers. The design of such an interface may be covered in a future marketing and communication plan for DCMI.

1.5.2 DCMI textbook ("Dublin Core Primer")

As the success of O'Reilly Publisher demonstrates, there is a continuing demand, even in the Web age, for printed documentation. All of the materials described above (except for Powerpoint slides) could be re-purposed and packaged in a Dublin Core Primer to be published, perhaps for profit, by DCMI.

2 DCMI behind the scenes

Part One discussed documentation intended for use by the general public — by decision-makers, vendors, system implementers, and end-users. Part Two discusses the internal maintenance processes underpinning those public results. Documents produced in support of these processes are intended primarily for use by members of the DCMI community, though they will of course be available to interested members of the community on the public Web.

2.1 Principled review

2.1.1 Approval and maintenance of metadata terms

The review of metadata terms for conformance with the Dublin Core model has and should continue to be a key activity of DCMI. In this report, the DCMI vocabularies themselves fall under the heading "behind the scenes" as a consequence of the shift in emphasis away from term sets per se towards specific application profiles that use and package those terms for specific implementation purposes.

2.1.2 Conformance review of application profiles

Reviews of application profiles would aim primarily at verifying the conformance of the profile as a whole (and of its constituent terms) to the DCMI Abstract Model. The result would typically be the assignment of a DCMI status to an application profile (e.g., "conforming"), after which a profile could be presented on the DCMI Web site as a Featured Application Profile [1.3.3]. Reviews would be undertaken with regard to profiles of strategic importance for DCMI; other profiles could potentially be reviewed on a consulting basis. Policies and procedures for review of non-DCMI profiles would need to be formulated as part of a broader DCMI approach to the development of review and certification services.

2.1.3 Endorsement of terms in non-DCMI namespaces

In cooperation with the US Library of Congress, DCMI is currently finalizing an agreement whereby the Library of Congress formally asserts a set of MARC Relator terms (such as Translator) to be sub-properties of the Dublin Core element Contributor, while DCMI, in turn, formally endorses those assertions. The recognition of MARC Relator terms allow implementers to use an existing, well-maintained term set, thereby avoiding the needless creation

and maintenance of a new vocabulary for the same purpose by DCMI. If such a need were to present itself with regard to other vocabularies, this assertion-endorsement form could be followed for other such agreements.

2.1.4 Criteria for review

One very important outcome of DCMI's experiences with reviewing both individual terms and application profiles as wholes for conformance with the Abstract Model will be the formulation of criteria by which terms and profiles can be evaluated in general. First drafts of a "DCMI mixing and matching FAQ" [14] and a draft "DCMI-compliant 'term' decision tree" [16] are important steps in this direction and will undoubtedly be refined as a result of experience. Documents such as these will provide crucial input to guidelines for creating application profiles [1.4.1] and metadata vocabularies [1.4.2].

2.1.5 Review-related processes and policies

In the interests both of transparency and of providing good-practice models for others to follow, DCMI should ensure that the processes and underlying policies for the review of metadata terms and and application profiles are properly documented. One key policy underlying all of these activities is the DCMI Namespace Policy [4], which must be amended to support DCMI Extension Namespaces. Process documents to be maintained also include Usage Board administrative processes [11].

2.2 Multiple-format publication of namespaces

DCMI's model for the formal documentation of its namespaces in the Semantic Web environment is one of the best-developed maintenance models in the Semantic Web environment. To the extent that Semantic Web applications become deployed in the mainstream, the model and work-flows used by DCMI to document and publish its namespaces should provide an example of good practice for maintainers of other such vocabularies.

2.2.1 An application profile for describing metadata terms

DCMI metadata terms are currently described with a set of attributes such as Label, Definition, and Status. This descriptive model is, in effect, a Dublin-Core-like application profile, though it has not yet been defined and presented as such.

There is currently a good opportunity for DCMI to cooperate with the maintainers of the FOAF [21] and SKOS [19] vocabularies on harmonizing the respective term-description models. Publishing DCMI's application profile of that model and assigning it a formal status would increase its value as a model of good practice for maintainers of other such metadata vocabularies.

2.2.2 A namespace for metadata term properties

Not all of the properties used to describe DCMI metadata terms have yet been formally declared (e.g., Status). These properties could be declared in a DCMI-maintained namespace, but it could also be desirable for DCMI to cooperate with other communities in the development of a shared namespace, whether that namespace is maintained by DCMI or by another organization.

2.2.3 The RDF model for namespace declarations

As one of the earliest adopters of the technology, DCMI has since 1997 maintained machine-processable representations of its vocabularies in RDF. In the context of a broader review of the technical specifications for encoding of Dublin Core metadata in RDF/XML [1.4.4], the RDF model for declaring DCMI vocabularies is also under review. Once this has been done, DCMI should formally clarify the status of these machine-processable representations and how they relate to natural-language definitions of the metadata terms.

(The question is: To what extent are DCMI terms "defined" by their English-language definitions and to what extent by formal assertions such as semantic relations to other terms? In what sense are alternative representations of DCMI vocabularies — Web documents as opposed to machine-processable schemas — to be regarded as "canonical" in this regard? How labels and definitions translated into Japanese or Spanish should exactly be modeled as independently-maintained RDF annotations also remains an open issue, with implications for the scalability of maintaining a namespace translated into many languages.)

2.2.4 Publishing namespaces in both XHTML and RDF

Whenever there is any change in, or addition to, the DCMI term sets, DCMI simultaneously issues new versions of its Web documents and new versions of the RDF schemas. The key Web document is "DCMI Metadata Terms" [5], a one-stop source of up-to-date information on all terms maintained by DCMI. Historical descriptions of DCMI terms are maintained for the interpretation of legacy metadata.

Each new or updated description of a metadata term is associated with a numbered decision [10], which in turn is itself associated with a decision text, meeting notes, and archival copies of supporting materials such as proposals. Every element of this chain — term description, decision, and related materials — is identified with a persistent URI.

For citation purposes, each new Web document and RDF schema is assigned a persistent URI. (These "historical" URIs for published sets have no affect on the persistent URIs used to identify the metadata terms documented therein.) At present, these Web documents and RDF schemas are generated from an single XML-tagged data source using XSLT scripts.

This "manual" approach to document management has proven effective enough given the relative infrequency of updates. If DCMI is to present itself as a leader in the deployment of semantic technologies, however, it would be desirable to solve this problem in a less ad-hoc manner using tools applicable, in principle, to any other vocabulary of metadata terms.

DCMI should therefore remain open to replacing the current term management system with a more sophisticated Semantic Web editing environment. The need for a more sophisticated solution may indeed become more urgent if the manual methods prove not to scale well to the demands of maintaining a growing number of DCMI Extension Namespaces, which are likely to hold a large number of controlled vocabularies and may need to be edited in collaboration with outside organizations. In the medium term, one possible solution might be to express the vocabularies primarily in RDF schemas, using these as the source for generating Web pages, as is currently done with the "SKOS Core" vocabulary [19] [2.2.1].

It is worth noting here that there has long been a request from some implementers that DCMI make its vocabularies available in their entirety as XML schemas. A shift in emphasis by DCMI from the term set to the application profile now provides an opportunity to properly re-define the task from that of maintaining an XML schema of a raw vocabulary to that of providing XML schemas for metadata records based on specific application profiles [1.3.4].

2.3 Multiple-format publication of application profiles

2.3.1 Specification of application profiles as documents

Guidelines for describing and formatting an application profile as a Web document have been formulated in the Dublin Core working group of the European Committee for Standardization (CEN) [13]. By prior agreement with CEN, these guidelines will be further developed by DCMI.

2.3.2 RDF model for application profiles

A draft model for expressing an application profile in RDF was also developed in the CEN working group [22], though the work should at this point still be considered research. While not an immediate priority, this work could be revisited as a basis for indexing multiple application profiles in portal environments and as an aid for constructing queries.

2.4 DCMI as a standards maintenance agency

The NISO and ISO standards for Dublin Core periodically come up for review. As the official maintenance agency for these standards, DCMI has the opportunity at such times to bring the formal standards editorially into line with the latest updates to the vocabularies. These procedures require that changes be documented and justified.

At present, issues arising with regard to Dublin Core metadata terms, best-practice usage, or DCMI as an organization are tracked primarily in the form of reports on meetings and teleconferences of the Directorate and Usage Board. In the longer term, more sophisticated approaches involving the identification of issues with URIs and formal procedures for resolution, such as those used by the World Wide Web Consortium, could be considered.

2.5 DCMI as publisher

2.5.1 The DCMI Publication Policy

DCMI operates as a publisher with respect to all of the resources described in this report. Changes in the nature of DCMI documentation should be reflected in a revised DCMI Publication Policy [12] [2.5.1].

2.5.2 Documentation of Website procedures

The work-flows and routines related to the publication processes described above need themselves to be documented in sufficient detail to be usable, in principle, by future successors to the current DCMI Webmasters and directors. For example, the scripts and procedures used by the Usage Board chair and by the Web Team for the simultaneous publication in XHTML and in RDF of DCMI terms documentation [2.2.4] should be captured in maintenance documentation for the DCMI Website.

3 Towards a work plan

3.1 High-level user-oriented overviews

In the near term, there is a need for high-level presentations of Dublin Core metadata in various formats — as a five-page introduction [1.1.1], a one-page summary [1.1.2], and as a Powerpoint presentation [1.1.3].

In the medium term, these materials should be featured in a new, user-oriented Web interface [1.5.1] to be designed in accordance with a future DCMI marketing and communication plan.

When	Where	What	Type
2005	1.1.1	Dublin Core from 10,000 feet — five-pager	general
2005	1.1.2	Dublin Core from 30,000 feet — one-pager	marketing
2005	1.1.3	Why metadata? — boardroom presentation	marketing
2006	1.5.1	User-oriented Web interface	marketing

3.2 Practical guidance for implementers

In the near term, existing drafts and specifications should be repurposed or revised, as necessary, in order to put into place a full set of user-oriented guidelines on creating application profiles [1.4.1] and declaring metadata vocabularies [1.4.2]. Taking into account ongoing clarifications of the DCMI Abstract Model [1.2.2], an effort should be made in the medium term to describe the Abstract Model is less–technical terms for ordinary users [1.2.3].

Also in the near term, work on revising the guidelines for Dublin Core in XML [1.4.3] and RDF/XML [1.4.4] in light of the Abstract Model should proceed as quickly as possible, bearing in mind that significant changes in existing recommendations could require a concerted communication effort by DCMI.

In the near-to-medium term, an editorial structure should be put into place for the maintenance of citable answers to Frequently Asked Questions [1.4.5]. In the longer term, DCMI should consider re-packaging its documents and user guidelines for a periodically revised printed textbook [1.5.2].

When	Where	What	Type
2005	1.4.1	How to implement an AP	general
2005	1.4.2	How to declare a metadata vocabulary	general
2005	1.4.3	How to encode DC metadata in XML	technical
2006	1.2.2	Clarifications of the Abstract Model	technical
2006	1.2.3	The Abstract Model made simple	general
2006	1.4.5	Frequently Asked Questions	general
2006	1.4.6	Introductory texts and tutorial materials	marketing
2007	1.5.2	DCMI textbook (“Dublin Core Primer”)	general

3.3 Review of application profiles

In the near term, the first experiences of the DCMI Usage Board in reviewing application profiles [2.1.2] should be used as a basis for progressively refining the criteria used for assessing the conformance of application profiles to the Abstract Model [2.1.4] as well as for testing the specifications for expressing application profiles as documents [2.3.1].

These criteria should be used to assign status to "conforming" application profiles, which could in turn be presented on the DCMI Web site as Featured Application Profiles [1.3.3]. In this context, "Simple Dublin Core" could be re-packaged as a Featured Application Profile [1.3.1]. In the medium term, an even simpler "Minimal Dublin Core" could be developed [1.3.2].

When available, record templates should be made available alongside the application profiles on which they are based [1.3.4].

These first experiences will need to proceed in parallel to the extension of the DCMI Abstract Model to include a formal specification for Application Profiles [1.2.1]. In the longer term, DCMI may want to resume work on a draft RDF model for application profiles [2.3.2].

These first experiences in the review of application profiles should also, in the medium term, flow back into the revision of the process and policy documents underlying DCMI's review processes [2.1.5] and, eventually, into a more fundamental discussion of the model by which DCMI and its affiliates might offer conformance testing and certification services.

When	Where	What	Type
2005	2.1.2	Conformance review of APs	technical
2005	2.1.4	Criteria for review	technical
2006	1.3.3	Featured application profiles	general
2006	1.2.1	The DCMI Abstract Model, extended	technical
2006	1.3.1	Simple Dublin Core	technical
2006	1.3.2	Minimal Dublin Core	research
2006	2.1.5	Review-related processes and policies	technical
2006	2.3.1	Specification of APs as documents	research
2007	1.3.4	Record templates for featured APs	technical
2007	2.3.2	RDF model for application profiles	research

3.4 Maintenance of DCMI vocabularies

In the near term, and on an ongoing basis, new terms will be added and existing terms maintained both in "central" DCMI namespaces and, increasingly, in new DCMI extension namespaces. This will involve the review of proposals for new terms, for editorial changes to existing terms, and for application profiles eligible to use extension namespaces [2.1.1]. As an aspect of these activities, DCMI should consider building on the precedent of MARC Relator terms to endorse the use of non-DCMI vocabularies with Dublin Core [2.1.3]. Any changes made in the vocabularies, along with issues arising in their regard, should be tracked in a form usable to support a periodic review of the NISO and ISO standards for Dublin Core [2.4].

In the near-to-medium term, DCMI should liaise with other vocabulary maintenance communities in the formalization of an application profile for describing DCMI metadata terms [2.2.1], which will involve the declaration of DCMI term properties in a namespace maintained either by DCMI or by another organization [2.2.2]. Once discrepancies between DCMI's RDF guidelines and the DCMI Abstract Model [1.4.4] are resolved, the RDF Model for DCMI namespace declarations should be reviewed and given formal status [2.2.3].

When	Where	What	Type
2005	2.1.1	Approval and maintenance of metadata terms	general
2005	1.4.4	How to encode DC metadata in RDF/XML	technical
2006	2.1.3	Endorsement of non-DCMI terms	technical
2006	2.2.1	An AP for describing metadata terms	technical
2006	2.2.2	A namespace for metadata term properties	technical
2006	2.2.3	The RDF model for namespace declarations	research
2007	2.4	DCMI as a standards maintenance agency	general

3.5 DCMI as publisher

DCMI operates as a publisher with respect to all of the resources described in this report. Changes in the nature of DCMI documentation as a result of this report will need to be reflected in a revised DCMI Publication Policy [2.5.1]. Work-flows and routines related to the publication processes described above need to be documented in sufficient detail to be usable, in principle, by any future successors to the current DCMI Webmasters and directors [2.5.2].

DCMI should be in the forefront of adopting new semantic technologies for the maintenance of its own specifications and vocabularies. Sophisticated editing environments currently under development in the marketplace could potentially be deployed with good effect for the simultaneous publication of DCMI vocabularies in XHTML and in RDF [2.2.4].

When	Where	What	Type
2005	2.5.1	The DCMI Publication Policy	general
2006	2.5.2	Documentation of Website procedures	general
2007	2.2.4	Namespaces in XHTML and RDF	technical

Thomas Baker/2005-08-28, revised 2006-03-31

References

[1] http://dublincore.org/documents/dc-elem-refine/

[2] http://dublincore.org/documents/dc-xml-guidelines/

[3] http://dublincore.org/documents/dcmes-xml/

[4] http://dublincore.org/documents/dcmi-namespace/

[5] http://dublincore.org/documents/dcmi-terms/

[6] http://dublincore.org/documents/dcq-html/

[7] http://dublincore.org/documents/dcq-rdf-xml/

[8] http://dublincore.org/resources/faq/

[9] http://dublincore.org/resources/training/

[10] http://dublincore.org/usage/decisions/

[11] http://dublincore.org/usage/documents/process/

[12] http://dublincore.org/usage/documents/publications/

[13] http://www.cenorm.be/isss/cwa14855/

[14] http://www.ukoln.ac.uk/metadata/dcmi/mixing-matching-faq/

[15] http://www.ukoln.ac.uk/metadata/dcmi/rdf-value-strings/

[16] http://www.ukoln.ac.uk/metadata/dcmi/term-decision-tree/

[17] http://www.ukoln.ac.uk/metadata/dcmi/term-identifier-guidelines/

[18] http://www.w3.org/TR/grddl/

[19] http://www.w3.org/TR/swbp-skos-core-spec/

[20] http://www.w3.org/TR/xhtml2/

[21] http://xmlns.com/foaf/0.1/

[22] ftp://ftp.cenorm.be/public/ws-mmi-dc/mmidc144.pdf

Title: DCMI Usage Board administrative process
Identifier: /usage/meetings/2006/04/seattle/ub-process/index.shtml

Discussants: Stuart and Diane

[1] /usage/documents/process/
[2] /usage/meetings/2006/04/seattle/ub-process/2006-02-23.process-ub-voting.txt

A revised process document was published on 2006-02-13 [1]. Subsequently, there was some discussion on the rules for voting -- whether people who did not participate in a face-to-face discussion should nonetheless have a vote [2] -- after which we agreed on a telecon to leave the wording vague for now.

On 2006-03-30, Stuart had an action item to write up a paragraph regarding UB process for revising the UB process document -- a statement regarding how that process document itself is revised so that there is something to point to should questions of process arise. Attached is a strawman section 10 for the _DCMI Usage Board Administrative Processes_ document [1]:

> 10. Revisions to Usage Board processes
>
> 10.1. Purpose of process revisions. Usage Board processes are revised and new versions of the _DCMI Usage Board Administrative Processes_ document published from time-to-time: (a) to clarify and correct process statements in the _DCMI Usage Board Administrative Processes_ document; (b) as the result of evolving knowledge and practice; (c) to conform to the evolving mission of DCMI; and (d) on request of the Board of Trustees or the DCMI Directorate.
>
> 10.2. Managing process revisions. The processes of revising Usage Board processes and the _DCMI Usage Board Administrative Processes_ document are informal, ongoing and managed by the Chair of the Usage Board.
>
> 10.3. Documenting Usage Board processes. All revisions to Usage Board processes are reflected here in the _DCMI Usage Board Administartive Processes_ document.
>
> 10.4. Communicating process revisions. Substantive revisions to the _DCMI Usage Board Administrative Processes_ document are announced on DC-General and other appropriate DCMI discussion lists.

DCMI Usage Board Administrative Processes

Creator: Diane I. Hillmann
Creator: Stuart A. Sutton
Contributor: Thomas Baker
Date Issued: 2006-02-13
Identifier: http://dublincore.org/usage/documents/2006/02/13/process/
Replaces: http://dublincore.org/usage/documents/2003/02/07/process/
Is Replaced By: Not Applicable
Latest version: http://www.dublincore.org/usage/documents/process/
Description of document: This document describes the process by which the DCMI Usage Board reaches and documents decisions. The Usage Board (UB) acts in accordance with its charter under the DCMI Bylaws, Article II, section D. The descriptions of process in this document are intended to guide the UB in executing its responsibilities for ensuring "an orderly evolution of the metadata terms maintained by the Dublin Core Metadata Initiative". The process statements are amended from time to time to reflect the evolving role of the Usage Board. In case of discrepancies, the DCMI Bylaws control.

Index

1. Usage Board Membership [from DCMI Bylaws].

1.1. Membership. The Usage Board will consist of at least seven and no more than eleven people (nine is ideal) appointed by the DCMI Directorate.

1.2. Responsibilities. The mission of the Usage Board is to ensure an orderly evolution of the metadata terms maintained by the Dublin Core Metadata Initiative. The Usage Board evaluates proposals for new terms (or changes to existing terms) in light of grammatical principle, semantic clarity, usefulness, and overlap with existing terms. To proposals that are accepted, it assigns a specific status. The Usage Board also evaluates constructs that use DCMI terms, such as Application Profiles.

1.3. Selection and Appointment Process. Members are selected based on the following criteria: knowledgeable concerning the development history and purpose of the DC element set and its relationship to the metadata world at large; related to a metadata community relevant to DCMI; willing and able to commit time and energy to the functions of the UB; able to communicate verbally and in writing in English well enough to prepare documents and discuss complex issues in a group setting; geographic and domain distribution of members is relevant but will not override other criteria. The DCMI Directorate will appoint the UB Chair from one of the membership. The DCMI Directorate can propose the appointment of non-voting Liaison members to the Usage Board. Liaison members may represent DCMI Affiliates or Sponsors, or other organizations that have a stake in the development of the Dublin Core semantic specifications.

1.4. Terms. Usage Board members are appointed for two-year, renewable terms. They may step down at their own discretion at any time.

1.5. Decision process. The Usage Board strives for consensus, justifying its decisions and interpretations in terms both of principle and of empirical practice. To be approved, a proposal needs more than 50% of assigned votes in favor and less than 25% of assigned votes against. Important decisions will be assigned a number for citation purposes and documented on the DCMI Web site. Decisions of the UB are forwarded to the DCMI Directorate for endorsement and approval.

1.6. Communication and documentation. For internal communication, the UB uses a closed mailing list. The messages are archived and are made publicly available. Meetings are held at least once a year. This meeting is scheduled during the annual DC general workshop/conference. Further meetings can be scheduled, preferably close to other conferences, so as to make attendance convenient for as many members as possible. Scheduling is done far enough in advance so that as many members as possible may be present.

1.7. Reporting. The chair of the Usage Board is responsible for the preparation of a report of meetings and conference calls and submission to the DCMI Directorate. Based on this report and after endorsement of Usage Board decisions, the Managing Director communicates the decisions to the DCMI community. Decisions on semantics are included in the reference documentation on the DCMI Web site.

2. Meetings

2.1. The calendar of UB meetings will be announced prominently on the UB DCMI homepage and announced on the DC-General mailing list.

2.2. Funding for regular UB members attendance at meetings should be supported as much as possible by DCMI. Funding for the attendance of Liaisons at UB meetings should be provided by their institutions.

2.3. The UB Chair maintains the agenda, which cites links to relevant supporting documentation, including JISCMAIL postings. All materials considered by means of the agenda are consolidated in a PDF briefing book and distributed electronically to UB members and Liaisons. At the conclusion of the meeting, the PDF briefing book becomes the official record of matters considered at the meeting. Important decisions will be assigned a number for citation purposes and documented on the DCMI website.

2.4. Additional meetings of the UB through conference calls may be scheduled at the discretion of the UB Chair.

2.5. The UB chair is responsible for assigning shepherds to proposals. Agenda items shall include the name of the UB member responsible for shepherding the proposal through the UB process. Agendas shall be available at the Web page DCMI Usage Board Meetings several weeks before the meeting.

2.6. Members must attend at least one meeting in a given year to maintain membership in good standing. Members who miss two meetings in succession may be replaced by the DC Directorate.

2.7. Attendance at any UB meeting by other than the UB members and liaisons is by invitation. People interested in attending should request an invitation via the UB Chair or the Managing Director. Participation in discussion of proposals by any interested parties is encouraged.

3. Status assigned by Usage Board decisions

Conforming [URI http://dublincore.org/usage/documents/process/#conforming]. Elements, Element Refinements, and Application Profiles may be assigned a status of conforming. Elements and Element Refinements assigned a status of conforming are those for which an implementation community has a demonstrated need and which conform to the DCMI Abstract Model.

Recommended [URI http://dublincore.org/usage/documents/process/#recommended]. Elements, Element Refinements, and DCMI-maintained Vocabulary Terms (e.g., member terms of the DCMI Type Vocabulary) that conform to the DCMI Abstract Model and do not semantically overlap with other terms in DCMI namespaces (i.e., http://purl.org/dc/elements/1.1/ and http://purl.org/dc/terms/) may be assigned the status of Recommended.

Obsolete [URI http://dublincore.org/usage/documents/process/#obsolete]. For Elements and Element Refinements that have been superseded, deprecated, or rendered obsolete. Such terms will remain in the registry for use in interpreting legacy metadata.

Registered [URI http://dublincore.org/usage/documents/process/#registered]. Used for Vocabulary Encoding Schemes and language translations for which the DCMI provides information but not necessarily a specific recommendation.

Endorsed [URI http://dublincore.org/usage/documents/process/#endorsed]. A non-DCMI assertion may be assigned the status of Endorsed where: (1) the term is managed by a registration authority other than DCMI and the assertion is that the term is conforming to the DCMI Abstract Model; or (2) the term is managed by a registration authority other than DCMI and the assertion is that the term bears either a property or subproperty relationship to a DCMI term.

5. Process for moving proposals

5.1. Assignment of shepherd Each impending proposal and application profile shall be assigned a shepherd by the UB chair from among the UB membership at the earliest opportunity. Shepherds should have knowledge of the proposal issues or application profile domain.

5.2. Shepherd responsibilities include:

- Monitoring discussion on relevant lists (shepherds should be members of the relevant DC WG list during the time of consideration of a proposal and are encouraged to join in the discussion to ensure that all relevant issues are exposed during the discussion period).
- Summarizing the comment period discussion and points of contention of the proposal for the UB, either verbally at the meeting or in writing prior to the meeting (preferred). Serving as liaison to the relevant WG or community during the time the proposal is under discussion and after a decision has been made. Preparing a draft of UB official decision on the proposal for review and approval by the UB.
- In general providing advice and expertise to the Working Group or domain on good practices, the Abstract Model, and other issues affecting the process of developing a proposal or application profile. The shepherd should bring issues of concern to the attention of the UB when appropriate.

5.3. Preparing for public comment periods. Completed proposals are forwarded to DCMI Managing Director or UB Chair. Proposals are given preliminary review for completeness by the DCMI Managing Director and UB Chair. If complete and no revisions are needed, proposals are circulated to UB members and announced for public comment by the Managing Director. A period of two weeks will be allowed between the date of the decision on completeness and the public announcement of the proposal to provide time for preparation of the supporting materials for public dissemination. If incomplete or revisions are needed, proposals are returned to the originator, with a request for revision or additional information.

5.4. Announcing the public comment period. Before announcement of the public comment period, proposals must be moved to the DCMI Web site, given DCMI page headers and a status of 'Proposed'. Announcement of the public comment period shall be made on the DC-General mailing list by the head of UB. Announcements should include links to the proposal; links to other relevant information; deadlines for comments; email addresses to be used for submitting comments. (In general, comments regarding a proposal may be addressed to the relevant WG mailing list, the DC-General mailing list or privately to the shepherd.) Announcements may also include information about the UB meeting at which the proposal will be discussed, including place, time, and how to request an invitation to participate; name and contact information for the assigned shepherd. The announcement should ask specifically for communications supporting the proposal in order to gauge the level of community support.

5.5. Managing public comment periods. The comment period for proposals should be managed on the DC-General list. Comment periods must be at least one month in length and commence at least six weeks before the UB meeting at which action is to be taken. Public discussions of UB related issues during public comment periods should take place on DC-General or other working group mailing lists as specified in the announcement. The public discussion must start at least six weeks before the UB meeting at which the issues will be discussed.

5.6. Communication responsibility

Communication Responsibility Table

What	Where	Who	Comment
Proposal draft posted	WG list, DC-General	WG Chair	
Proposal added to UB agenda	UB Website, UB list	UB Chair	
Proposal announced for public comment	DC-General	DCMI Directorate	
Usage Board decision announced	DC-General	DCMI Directorate	

5.7. Voting on proposals. Voting shall be limited to scheduled meetings and publicly announced conference calls. Voting shall be limited to UB members present at the meeting or conference call and able to participate in the discussion. UB members who cannot be present may present their arguments for or against a proposal in writing prior to a meeting (this shall not constitute a vote). UB members who cannot be present may explore other options with the chair, if they cannot be present for an important vote. In all cases, a vote may not be cast by a member who is not present, either physically or virtually, for the relevant discussion.

5.9. Documenting Usage Board decisions. A document explaining the UB decision regarding a proposal will be written in a timely fashion by the shepherd and approved by the UB. The decision will include brief statements of recommendations being issued and detailed explanations of UB decisions not to issue recommendations. UB decisions will be in a form determined by the UB and numbered consecutively for the purpose of citation. UB decisions must be sufficiently documented so that the rationale for the decision is clear and useful in guiding the development of future proposals. This is particularly true where the decision rejects a proposal or recommends further action.

5.10. Announcing and publishing Usage Board decisions. Decisions are published on the Web page DCMI Usage Board decisions. New terms will be added to the official DCMI documentation by the UB Chair.

6. Proposals for registration of application profiles

6.1. Application Profiles subject to review. Application profiles emanating from DCMI Strategic Activities may be reviewed by the Usage Board. Metadata implementers (established projects, communities or research groups) may also request review, subject to approval by the UB Chair. *Point to information regarding DCMI Strategic Activities when available.*

6.2. Documentation of Application Profiles. Application profiles must provide, for each term, an identifier of the element set where it is defined, ideally in the form of URIs for individual terms. If the terms in an application profile describe anything other than generic "resources" (the typical domain of Dublin Core), the application profile must make this clear. This is particularly important if an application profile is based on a data model that describes multiple classes of resources, such as agents or collections. It is recommended that application profiles be prepared using the guidelines .

6.4. Contextual information about Application Profiles. The documentation for each Application Profile must provide -- or point to a short text that describes -- the context and purposes in which the application profile is used or is likely to be used; the organizations or individuals involved in its development and a capsule history thereof; and any arrangements, policies, or intentions regarding the future development and maintenance of the application profile.

6.5. Evaluation of terms in Application Profiles. The use of terms related to Dublin Core (such as refinements of Dublin Core elements, or Dublin Core elements that have been constrained for particular contexts) will be evaluated from the standpoint of semantic conformance, grammatical principle (eg, "dumb-down"), clarity, and good practice. *Note: revisit this.*

6.6. Assignment of status "conforming". Application profiles which pass review will be assigned the status of conforming The status of conforming indicates a Usage Board assessment of the application profile as of the date of its submission for review. Changes to already conforming application profiles require further Usage Board review of the application profile in whole or in part according to the processes and criteria outlined in sections 6.1 through 6.3.

6.7. Publication of Usage Board reviews of Application Profiles. For application profiles that "pass" review, the Usage Board will publish a Review on a Web page for application profiles. Each Review will include, at a minimum: any comments from the Usage Board on the application profile; pointers to locally archived copies of the application profile as originally submitted and (if necessary) as subsequently amended in light of Usage Board comments; a pointer (with appropriate disclaimers) to the "latest version" of an application profile held by its maintainers.

6.8. Persistent identifiers for reviewed Application Profiles. Review represents a form of recognition,

and its URL will be persistent for purposes of citation.

7. New terms proposed with an Application Profile

7.1. Evaluation of new terms. New terms appearing in application profile submissions must be evaluated for compliance with the DCMI Abstract Model prior to evaluation of the Application Profile itself.

7.2. Assignment of DCMI term URIs and status. New terms deemed in compliance with the DCMI Abstract Model may be given URIs in DCMI namespaces and assigned a status of conforming.

7.3. Conformance criteria. Decisions as to whether a proposed term is in compliance with the DCMI Abstract Model will be made using the DCMI-Compliant Term Decision Tree.

8. Proposals for endorsement of terms in other namespaces for use within Application Profiles

8.1. Existing terms housed in other namespaces to be used within Application Profiles seeking review must be evaluated for compliance with the DCMI Abstract Model prior to evaluation of the Application Profile itself.

9.

Revisions of existing DCMI terms

9.1. Proposals for revisions. Requests to change terms in DCMI namespaces may originate within the Usage Board or externally. A Usage Board member will be assigned to draft a proposal for a change. Changes provisionally approved by the Usage Board will be circulated for general comment on the DC-General discussion list for one month before final approval. Final approval for term changes without significant opposition may be approved by email or teleconference vote.

9.2. Changes for formally standardized terms. Terms from namespace http://purl.org/elements/1.1/ require changes to the relevant standards: ISO Standard 15836-2003 (February 2003) and NISO Standard Z39.85-2001 (September 2001).

9.3. Terms from DCMI-hosted namespaces [to be added]

9.4. Application profile terms residing on DCMI hosted namespaces will be subject to the same change processes as other DCMI terms, but managed by the entities responsible for the terms. Application profile terms residing on non-DCMI namespaces will be subject to term policies of the host entity.

9.5. Changes to already 'conforming' application profiles require further Usage Board review of the application profile in whole or in part according to the processes and criteria outlined in previous sections. Changes to DCMI-registered "conforming" application profiles will be versioned according to DCMI namespace policies.

Metadata associated with this resource: http://dublincore.org/usage/documents/process/index.shtml.rdf

---

Date: Wed, 22 Feb 2006 13:16:44 -0500
From: "Diane I. Hillmann" <dih1@CORNELL.EDU>
Subject: Fwd: RE: Proposed rewording in Process doc
To: DC-USAGE@JISCMAIL.AC.UK

All:

Stuart and I concur with the following changes:

The wording as it stands:

5.7. Voting on proposals. Voting shall be limited to scheduled meetings and publicly announced conference calls. Voting shall be limited to UB members present at the meeting or conference call and able to participate in the discussion. UB members who cannot be present may present their arguments for or against a proposal in writing prior to a meeting (this shall not constitute a vote). UB members who cannot be present may explore other options with the chair, if they cannot be present for an important vote. In all cases, a vote may not be cast by a member who is not present, either physically or virtually, for the relevant discussion.

We suggest the following substitution:

5.7. Voting on proposals. Voting shall be limited to scheduled meetings and publicly announced conference calls. Except for special circumstances approved in advance, voting shall be limited to UB members present at the meeting or conference call and able to participate in the discussion. UB members who cannot be present are encouraged to present their arguments for or against a proposal in writing prior to a meeting (this shall not constitute a vote). UB members who cannot be present may explore other options with the chair, if they cannot be present for an important vote. [last sentence eliminated]

I think this wording bends over backwards in accommodating the bylaws, without opening us to votes by those who haven't heard the arguments.

Date: Thu, 23 Feb 2006 08:55:53 -0000
From: Andrew Wilson <andrew.c.wilson@AHDS.AC.UK>
Subject: Re: Proposed rewording in Process doc
To: DC-USAGE@JISCMAIL.AC.UK

Fwd: RE: Proposed rewording in Process docThe wording as it stands seems to allow a chair to publicly announce a conference call for the following day. Should the wording include something about the period of notice for both scheduled meetinsg and conference calls?

Perversely, I'm not sure I like imposing these limitations on the ability of the UB to have ad hoc meetings/calls at very short notice when something important suddenly needs to be acted on.

Andrew

Date: Thu, 23 Feb 2006 09:32:46 -0000
Reply-To: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
Sender: A mailing list for the Dublin Core Metadata Initiative's Usage Board <DC-USAGE@JISCMAIL.AC.UK>
From: Andy Powell <andy.powell@EDUSERV.ORG.UK>
Subject: Re: Proposed rewording in Process doc
To: DC-USAGE@JISCMAIL.AC.UK

OK, that makes sense, but the fact you had to explain it means either that I'm thick (a likely possibility) or that it isn't clear enough (or both of course! :-)).

I wonder if it would be helpful to more clearly distinguish between 'the discussion' and 'the vote', along the following lines:

5.7. Voting on proposals. Voting shall be limited to
scheduled meetings and publicly announced conference calls.
Except for special circumstances approved in advance, voting

shall be limited to UB members present at the meeting or
conference call and able to participate in the discussion. UB
members who cannot be present [for the discussion] are encouraged to
present their
arguments for or against a proposal in writing prior to a
meeting (this shall not constitute a vote). UB members who
cannot be present [for a vote but who were present for the discussion]
may explore other options with the chair.

Title: Back burner
Identifier: /usage/meetings/2006/04/seattle/backburner/index.shtml

-- Issues related to dc:date

[1] http://people.opera.com/charlesm/2006/shortdate/
[2] http://www.ukoln.ac.uk/metadata/dcmi/date-dccd-odrf/
[3] /usage/meetings/2006/04/seattle/backburner/2005-08-10.rebecca-comments.txt
[4] /usage/meetings/2006/04/seattle/backburner/2005-08-13.YearMonthDate-profile.txt
[5] /usage/meetings/2006/04/seattle/backburner/2005-08-22.douglas-campbell-long.txt

-- Accessibility application profile

-- Wiki versus Subversion for UB document management

-- Profile for the description of DCMI terms

[6] /usage/meetings/2006/04/seattle/backburner/2005-05-10.usageboard-profile.txt

Date and Time Short Form

Working Draft - 15 January 2005

This version:
http://people.opera.com/charlesm/2006/shortdate/060115.html
Newest version:
http://people.opera.com/charlesm/2006/shortdate/
Previous version:
http://people.opera.com/charlesm/2006/shortdate/051230.html
Author:
Charles McCathieNevile <chaals@opera.com>, Opera Software

Status of this document

This document is being prepared as a submission to W3C, to assist the work of the Dublin Core Metadata Initiative's Dates working group. It has been reviewed informally by several people, and this draft is offered to the DCMI Date group and other interested parties for review. This is the second published draft, incorporating changes based on feedback received to date (15 January 2006) on the first draft, which it obsoletes.

This draft may be updated or obsoleted at any time. The author expects to propose it as a member submission to W3C early in 2006.

Feedback on this document is actively sought until the end of January 2006, or until the publication of a new draft, whichever is later. Feedback should be sent to the author <chaals@opera.com>, and may be copied to the DCMI Date working group (subscription required) if a public record of response is desired.

Abstract

This document defines a profile of ISO 8601, the International Standard for the representation of dates and times. ISO 8601 describes a large number of date/time formats. To reduce the scope for error and the complexity of software, it is useful to restrict the supported formats to a small number. This profile specifies some date/time formats in the compact form YYYY[MM[DD[ThhmmTZID]]], in order to satisfy some additional requirements that were not met by the W3C Date Time Format [DTF] Note written by Misha Wolf and Charles Wicksteed or the IETF RFC by Graham Klyne and Chris Newman [r3339] describing roughly the same profile.

Contents

Introduction

This section is informative.

Purpose of this document

The "International Standard" for the representation of dates and times is ISO 8601. Its full reference number is ISO 8601 : 1988 (E), and its title is "Data elements and interchange formats - Information interchange - Representation of dates and times". A useful discussion of ISO 8601 [Kuhn] has been written by Markus Kuhn.

ISO 8601 describes a large number of date/time formats. A previous document [DTF] by Misha Wolf and Charles Wicksteed specified the "basic format" of that standard as a profile for use on the World Wide Web, clarifying the format

> to simplify the use of ISO 8601 in World Wide Web-related standards, and to avoid the need for the developers and users of these standards to obtain copies of ISO 8601 itself

An IETF Standards Track Request for Comments [r3339] describes a similar profile.

This specification defines another restricted range of formats, all of which are valid ISO 8601 dates and times, using the so-called compact syntax, which eliminates punctuation. The aim of this second specification is to simplify the process of using data which has been recorded using the compact format defined in ISO 8601 by being explicit about the form of the data.

Note that this profile may be used as a convenience for mapping pre-existing data which is already stored in the syntax defined.

However, recommended best practice is to use the clearer Date Time Format [DTF] wherever possible, in particular when choosing a format for any new work.

Dating issues

For dates between 1 January 1 CE and 4 October 1582 CE, this specification uses the Julian calendar, which was the most standardised calendar in the Western world over that period. This means that dates given at the time (such as 24 August 79 CE, the date of Vesuvius' eruption which destroyed Pompeii and Herculaneum) can generally be directly recorded, rather than being recalculated to conform to the proleptic Gregorian Calendar.

This specification defines dates in terms of the Gregorian Calendar since it came into use in the Gregorian year 1582. However it was not adopted in many countries until hundreds of years later - for example the UK (and her then American colonies which later became the USA) began to use it in 1752, Albania in 1917, Yugoslavia in 1919 and China in 1929. It is also not the only calendar in use today.

This means that it is important to know which calendar was in use even where the "common christian calendar" was used for dates after 4 October 1582 and before the Gregorian Calendar was adopted locally. For example, Shakespeare and Cervantes were both recorded as having died on 23 April 1616. But Cervantes died 10 days earlier in Spain, which was using the Gregorian Calendar. According to that calendar Shakespeare died on 3 May - the equivalent of the Julian date 23 April in use in England at the time.

This problem can be compounded - the Bolshevik revolution in Russia was known as the October revolution, since according to the Julian calendar it began on 25 October. The Soviet Union adopted the Gregorian Calendar, and commemorated the date of its "creation" as 7th November, the Gregorian equivalent.

This specification cannot be used for dates before the year 1 CE, nor for years after 9999 CE. Hopefully by that time it will have been superseded by a more effective date specification *in toto*.

The dates 5th October 1582 to 14th October 1582 (inclusive) do not exist in the Gregorian calendar - they were removed, in order to bring the calendar into synnchronisation with the seasons (the purpose of adopting the calendar in the first place).

Conversion to and from Hejri (Islamic) calendar dates is problematic, since the first and last days of the months can vary depending on regional differences or religious orientation. It is therefore approximate, with a possible degree of error of a couple of days.

Syntax

Different information may have been recorded using different levels of granularity for the date and time, so this profile defines four levels. Comparisons between dates are only valid to the precision of the least precise date in the comparison.

The formats are as follows. Exactly the components shown here must be present, with exactly this punctuation. Note that the "T" appears literally in the string, to indicate the beginning of the time element, as specified in ISO 8601, and that all dates are defined as being calculated according to the Gregorian Calendar for dates on or after 4 October 1582, and according to the Julian Calendar for dates on or before 4 October 1582.

A Year
: *YYYY* (for example `1997`).

Year and month
: *YYYYMM* (for example `200511` for November 2005)

A date
: *YYYYMMDD* (for example `19700101` for 1 January 1970)

A date/time (with hours and minutes)
: *YYYYMMDDThhmmTZD* (for example `19991231T2359+1200` for 1 minute to midnight, 31 December 1999, in the timezone GMT + 12 hours)

where:

YYYY
: is a four digit year. Valid values are in the range 0001 to 9999.

MM
: is a two digit month. Valid values are 01 to 12. Numbers in the range 01 to 12 correspond to months in the range January to December.

DD
: is a two digit day. All values are in the range 01 to 28 are valid, except values for days which would fall in the Date range [15821005,15821014].
For valid month values other than 02 (February) the values 29 and 30 are valid.
For month values 01, 03, 05, 07, 08, [10,12] the value 31 is valid.
For the month value 02, the value 29 is valid in leap years - before 1582 all years divisible by 4 and not divisible 100, and after 1582 all years divisible by 4 and not by 100 plus all years divisible by 400

hhmm
: is a four digit time (hours and minutes) expressed in the 24-hour clock. Valid values are in the range 0000 to 2359
Editorial Note: This should be clamped more precisely. If seconds are required, note that leap seconds will be necessary. Times may be removed if there is no existing use of them - and in hoping for that to be true I haven't yet cleaned up the range for them.

TZD
: is a time zone identifier. Valid values are Z (correpsonding to UTC), or an offset from Z expressed as a signed two-digit number of hours followed by a two-digit number of minutes. Valid values are in the range -2359 to +2359. -0000 and +0000 are equivalent.

Examples

This section is informative

`19941105T0815-0500` corresponds to November 5, 1994, 8:15:30 am, US Eastern Standard Time.

`19941105T1315Z` corresponds to the same minute.

`19700603` corresponds to the 3rd of June, 1970.

`00790824` corresponds to the 24th of August, 79 CE as calculated according to the Julian calendar.

`15821010` is not a valid date according to this specification, since the 10th of October 1582 does not exist in the Gregorian Calendar

`20050230` is not a valid date - February never has more than 29 days.

Sorting dates

One of the benefits of this format is that it allows for relatively simple sorting to be performed for chronological order, using only simple arithmmetic

operators. Sorting of dates should proceed as follows:

1. Where not all dates are given to the same precision, add trailing zeros or nines to normalise the precision. (This permits imprecise dates to be sorted
2. The dates should be sorted in numerical order
3. Remove trailing zeros or nines (note that zero and 99 are invalid values for each component of a date).

For sorting of Date/times the following steps are required:

1. Times should be normalised, by adding or subtracting the timezone indicator from the time. Any date/times lacking sufficient precision should be marked with a sortable token.
2. Sort by date as above.
3. For each set of values within the same date, sort the time values numerically.
4. Times may be relocalised to timezones if desired.

Namespaces and datatypes

Best practice for implementation of this specification depends on the usage. This specification defines datatypes for use in RDF and in XML.

The four datatypes use the same URI in XML and RDF

Year
: [namespacePrefix]Year

Year and month
: [namespacePrefix]YMon

Date
: [namespacePrefix]Date

Date / Time
: [namespacePrefix]DTime

In HTML or other usage there are a number of ways that data can reference information about the formats used.

Editorial Note: Should there be references to how this is done? Dublin Core in HTML, Microformat, GRDDL, ??? There probably should be RDF / XML examples at least.

Common algorithms for determining valid dates are relatively simple. The exception may be the calculation of leap years, given that this specification incorporates two different calendars. The following code excerpt will work as is in many common languages:

```
isLeap = year % 4 && ( year < 1582 ? !(year % 100) : !(year % 100) || year % 400) )
```

Conformance

In this section, the keyword *must* is used as defined in RFC 2119 [r2119].

There are several different types of conformance to this specification. Data can, itself be conformant. Generating and consuming agents can also be conformant.

Two types of conformance are defined - Strict and Loose conformance. Anything which is Strictly conformant is also Loosely conformant, but meets some additional requirements.

Loosely conformant...

Consuming Agents
: Must accept data conformant to this specification (Loosely conformant and Strictly conformant Data).
: Must either
  - - apply the Error Handling rules defined in this specification to non-conformant data, OR
  - - reject non-conformant data

Data

Must conform to the requirements of this specification
Must be valid Julian or Gregorian dates

Generating Agents
Must produce dates which satisfy the requirements of this specification.

Strictly conformant...

Consuming Agents
Must accept data conformant to this specification (Loosely conformant and Strictly conformant Data).
Must reject non-conformant data.

Data
Must conform to the requirements of this specification.
Must be valid Julian or Gregorian dates.
Must specify a datatype. For RDF or XML data formats, the datatype must be one of those defined in this specification, or one defined to inherit all characteristics required by those datatypes. See also XML Schema part 2 - Datatypes [XML-types], and datatypes in the RDF/XML Syntax specification [RDF] for more information in specifying a datatype.

Generating Agents
Must produce Strictly Conformant Data according to this specification.

A strictly conforming agent which therefore both consumes and generates data, such as a calendar application, must accept both kinds of data but convert loosely conformant data to strictly conforming.

Error handling

Loosely conformant Consuming agents may accept non-conformant data, if they process it according to the following rules.

1. If there are insufficient digits to match any version of the syntax, and this can be corrected by adding leading zeros to match the syntax requirements, the minimal number of leading zeros sufficient to match the syntax shall be added.
2. If there are excessive digits, they should be interpreted as exending the year beyond the conformant range, with the last four digits of the date specifying month and day.
3. Any value of zero for a year, month, or day is clamped to a value of 01 or 0001 as appropriate.
4. Any value greater than 12 for a month is clamped to 12
5. Any day value greater than the maximum acceptable for the given month is clamped to the maximum acceptable day value for that month.
6. Any time value outside the acceptable range is clamped to 0000+0000 (midnight GMT at the beginning of the relevant date).

Editorial Note: If there is significant use of times, this should be adjusted to allow for hours and minutes to be clamped seperately. If there is none, times will be dropped from this specification.

Example error handling

This section is informative

`911` would be normalised according to rule 1 to `0911`, and thus specify the year 911 CE, as loosely conformant data.

`790824` would be normalised according to rule 4 to `790812` - The month of December in the year 7908. `0790824` would be normalised according to rule 1 to `00790824` - 24 August in the year 79 CE. Editorial Note: Is "-" used in practice? This could better disambiguate dates like this one although it could mean dealing with the "current century", which is likely to be a nightmare in practice.

Further Reading

This section is informative

There are many discussions available about dating issues in general, and the ISO 8601 standard (in its various revisions) in particular. In addition to documents already explicitly referenced in this specification, those which were useful in preparing this draft and are recommended to others include:

- ISO 8601:1988 Date/Time Representations by Gary Houston [Houst]
- Usage of Dates and Times on the Web by Jon Hanna [WebDate]
- Wikipedia, particularly the entry on the Gregorian Calendar [WP]

Issues unresolved in this draft

This section is informative, and may be removed from a final version.

1. Are times necessary? The purpose of the draft is to enable data recorde using this profile to state conformance to this specification. Best practice is to use a more flexible method of recording dates, such as Date Time Format [DTF], so unless there are **existing uses** of the profile to record date/times they are not necessary. Disallowing Time would mean no need for timezones, which makes things simpler.
2. Normative grammars. An XML Schema should be generated. It is not yet clear how to define an RDF datatype in a machine readable way, so until the Semantic Web Best Practices Working Group or some similar group produces an appropriate recommendation that will not be resolved. If the draft is otherwise considered complete it will be submmitted, and resolution of the definition of RDF datatypes will be left for an update.
3. Namespace. This specification should define a namespace. If it is accepted as a W3C submission then it will presumably get one.
4. This specification uses the Julian calendar for dates prior to 1582, rather than a proleptic Gregorian Calendar. This means that dates recorded at the time are more likely to be transferable directly into this format, rather than calculating them. Is this the best approach, or should a proleptic Gregorian calendar be adopted?
5. The use of "-" to identify a missing part isn't included in this specification. Should it be? The advantage would be flexibility, the drawback would be in sorting information. Proposal: Do not do this, unless there is significant feedback that the feature is used in existing data.

ABNF Grammar

This section is informative

The following is a partial grammar for this format in augmented Backus-Naur format [ABNF]. While all dates conforming to this specification will conform to one of the four constructions

- full-year
- year-month
- full-date
- date-time

Some constructions which conform to this grammar are not dates conformant to this specification.

```
date-month      = 2DIGIT  ; 01-12
date-mday       = 2DIGIT  ; 01-31
time-hour       = 2DIGIT  ; 00-23
time-minute     = 2DIGIT  ; 00-59
time-numoffset  = ("+" / "-") time-hour time-minute
time-offset     = "Z" / time-numoffset
full-time       = time-hour time-minute time-offset

fullyear        = 4DIGIT
year-month      = fullyear date-month
full-date       = date-yearmonth date-mday
date-time       = full-date "T" full-time
```

Acknowledgments

This section is informative

This document is based on the "Date and Time Formats" Note [DTF] submitted to W3C in 1997 by Misha Wolf and Charles Wicksteed and would not have been possible without the efforts of those authors. A lot was also learned in the RDF Calendar work based around W3C [rCal] and the author is grateful to all those who have participated in that group for their ideas and explanations. It is motivated by the work of the Dublin Core Date Working Group. In particular the author would like to thank the following individuals for their invaluable assistance with this document:

- Karima Boudaoud
- Daniel Bratell
- Douglas Campbell
- Eric Childress
- Dan Connolly
- Jon Hanna
- Kjetil Kjernsmo

- Graham Klyne
- Lars Tomas Hansen
- Anne van Kesteren
- Pete Johnson
- Andy Powell
- Eddy Welbourne
- Misha Wolf

References

[ABNF] RFC 2234: Augmented BNF for Syntax Specifications: ABNF
D. Crocker and P. Overell, 1997. http://www.ietf.org/rfc/rfc2234.txt

[D8601] A discussion of ISO 8601
Markus Kuhn, http://www.cl.cam.ac.uk/~mgk25/iso-time.html

[DTF] Date and Time Formats
Misha Wolf, Charles Wickstreet, 1998. http://www.w3.org/TR/NOTE-datetime

[Houst] ISO 8601:1988 Date/Time Representations
Gary Houston, 1993. http://hydracen.com/dx/iso8601.htm

[r2119] RFC 2119: Key words for use in RFCs to Indicate Requirement Levels
S. Bradner, 1997. http://www.ietf.org/rfc/rfc2119.txt

[r3339] RFC 3339: Date and Time on the Internet: Timestamps
G. Klyne, C. Newman, 2002. http://www.ietf.org/rfc/rfc3339.txt

[rCal] RDF Calendar Workspace
D. Connolly et al. eds. Updated occasionally with community contribution. http://www.w3.org/2002/12/cal/

[RDF] RDF/XML Syntax Specification (revised)
Dave Beckett ed. W3C Recommendation 2004. http://www.w3.org/TR/rdf-syntax-grammar

[WebDates] Usage of dates and times on the Web
Jon Hanna, undated http://www.hackcraft.net/web/datetime/

[WP] WikiPedia: Gregorian Calendar
Community authored, continually updated. http://en.wikipedia.org/wiki/Gregorian_calendar

[XML-types] XML Schema part 2 - Datatypes Second Edition
Paul Biron, Ashok Malhotra, eds 2004 http://www.w3.org/TR/xmlschema-2/

Title:

Dublin Core Collection Description : Open Date Range Format (DCCD ODRF)

Creator:	Dublin Core Collection Description Working Group
Date Issued:	2005-08-13
Identifier:	http://www.ukoln.ac.uk/metadata/dcmi/date-dccd-odrf/2005-08-13/
Replaces:	Not applicable
Is Replaced By:	Not applicable
Latest Version:	http://www.ukoln.ac.uk/metadata/dcmi/date-dccd-odrf/
Status of Document:	This is a working document which has no status within DCMI.
Description of Document:	This document describes a format for encoding date ranges as text. It is based on ISO8601 (2000) and W3CDTF but is a profile of neither of those standards/specifications.

Introduction

This document describes a text format for representing:

- a **closed date range**: a time-interval expressed as a start date and an end date; and
- an **open date range**: a time-interval expressed using either a start date alone (to indicate the time-interval since that date up to the present) or an end date alone (to indicate the time-interval up to that date)

This format is based on the ISO8601 (2000) standard [ISO8601] and the W3CDTF "profile" of ISO8601 [W3CDTF]. However the format described by this document is *not* a profile of ISO8601 since ISO8601 does not support the representation of time-intervals based on "open" ranges; and it is not a profile of W3CDTF since W3CDTF supports the representation of dates and times, but not time-intervals.

The Format

This format supports the representation of **time-intervals** expressed as a **start date** and/or an **end-date**.

The start date and the end date must be represented using one of the following formats defined by W3CDTF [W3CDTF]:

```
Year:
   YYYY (eg 1997)
Year and month:
   YYYY-MM (eg 1997-07)
Complete date:
   YYYY-MM-DD (eg 1997-07-16)
```

where:

```
YYYY = four-digit year
MM   = two-digit month (01=January, etc.)
DD   = two-digit day of month (01 through 31)
```

Only this subset of the formats defined by W3CDTF are permitted. The representation of time of day is not supported by this specification.

A solidus ("forward slash") ["/"] is used to separate the representation of the start date from the representation of the end date.

Closed date range

A **closed date range** is represented by a start date, followed by a "/", followed by an end date:

```
Closed date-range:
   YYYY-MM-DD/YYYY-MM-DD (eg 1997-07-16/1998-09-17)
```

Any of the date formats above (YYYY-MM-DD, YYYY-MM or YYYY) may be used for both start date and end date.

This representation of a **closed date range** is compatible with the representation of a time-interval defined by ISO8601:2000 (section 5.5.4.1, extended format).

Open date range

An **open date range** is represented by:

- a start date, followed by a "/", or
- a "/" followed by an end date

```
Open date-range:
   YYYY-MM-DD/ (eg 1997-07-16/)
   /YYYY-MM-DD (eg /1997-07-16)
```

Any of the date formats above (YYYY-MM-DD, YYYY-MM or YYYY) may be used for both start date and end date.

This representation of an **open date range** is *not* compatible with the representation of a time-interval defined by ISO8601:2000.

Examples

`1997-07-16/1997-07-17` corresponds to a time-period starting on July 16, 1997, and ending on July 17, 1997. From the representation provided, it is not possible to say whether this is a period of a few seconds (from late on July 16 until early on July 17) or of 48 hours (from 0.00 on July 16 until 24.00 on July 17).

`1997-07/1997-08` corresponds to a time-period starting during the month of July, 1997, and ending during the month of August, 1997. From the representation provided, it is not possible to say whether this is a period of a few seconds (from late on July 31 until early on August 1) or of 62 days (from 0.00 on July 1 until 24.00 on August 31).

`1997/1998` corresponds to a time-period starting during the year 1997, and ending during the year 1998. From the representation provided, it is not possible to say whether this is a period of a few seconds (from late on Decemeber 31, 1997, until early on January 1, 1998) or of 730 days (from 0.00 on January 1, 1997, until 24.00 on December 31, 1998).

`1997-07-16/` corresponds to a time-period starting on July 16, 1997.

`1997-07/` corresponds to a time-period starting during the month of July, 1997.

`1997/` corresponds to a time-period starting during the year 1997.

`/1997-07-16` corresponds to a time-period ending on July 16, 1997.

`/1997-07` corresponds to a time-period ending during the month of July, 1997.

`/1997` corresponds to a time-period ending during the year 1997.

References

[ISO8601] *ISO8601:2000(E) Data elements and interchange formats - Information interchange - Representation of dates and times.* Second edition

[W3CDTF] Misha Wolf & Charles Wicksteed, *Date and Time Formats.* W3C Note. Submitted to W3C 15 September 1997.

http://www.w3.org/TR/NOTE-datetime

Changes made in this version

- Initial version.

Date: Wed, 10 Aug 2005 17:09:29 -0400
Reply-To: "Rebecca S. Guenther" <rgue@LOC.GOV>
Sender: DCMI Date Working Group <DC-DATE@JISCMAIL.AC.UK>
From: "Rebecca S. Guenther" <rgue@LOC.GOV>
Subject: Re: Adding ISO 8601 as a scheme -- discussion
To: DC-DATE@JISCMAIL.AC.UK

There are 2 things I want to be able to do:

1. Have a way to refer to an encoding scheme that uses what in ISO 8601 is called the "basic format" for date and time, i.e. YYYYMMDD etc. without hyphens. W3CDTF uses the "extended format", i.e. YYYY-MM-DD etc. with hyphens.

2. Have a mechanism to accommodate the other types of dates we talked about, most of which are not in ISO 8601. These include:
approximate dates
questionable dates
open ended date ranges
BCE dates
Non-Gregorian dates
and maybe also:
Quarter
Season
(the latter 2 have been expressed as needed in Z39.50 several years ago).

To get my need #1, a profile of ISO8601 could be done to include that and leave out what in ISO8601 we don't want. But I couldn't get those constructs not now included in 8601 in a profile because they weren't there to begin with.

I am trying to get a copy of the new version from NISO-- as we know ISO doesn't make these freely available. To my knowledge those expressions of date still aren't accommodated.

So there would need to be an extension or revision to 8601 rather than a profile to get #2.

If DCMI were to register all of ISO8601 then I guess I would have #1 available, although it still wouldn't be explicit that what I'm using is the basic notation in ISO8601, which was my point all along. And I'm told that registering ISO8601 basic was rejected so I'll have to live with it. I don't have the energy to fight this one any more.

I'm thinking about doing the extensions within NISO to attempt to get ISO 8601 revised to include these. Emphasize the word "thinking".

Rebecca

> Date: Fri, 5 Aug 2005 10:08:37 -0400
> From: "Childress,Eric" <childree@OCLC.ORG>
> Subject: Re: Adding ISO 8601 as a scheme -- discussion
>
> This is a multi-part message in MIME format.
>
> OK, we've heard opening comments from me, Andy, Charles, Misha, Pete
> about registering ISO 8601. How about some additional voices?
> Thanks,
>
> Eric

Date: Sat, 13 Aug 2005 16:49:28 +0200
Reply-To: Charles McCathieNevile <chaals@OPERA.COM>
Subject: Re: Adding ISO 8601 as a scheme -- discussion
Comments: To: Misha Wolf <Misha.Wolf@reuters.com>
To: DC-DATE@JISCMAIL.AC.UK

On Fri, 12 Aug 2005 23:48:12 +0200, Misha Wolf <Misha.Wolf@REUTERS.COM> wrote:
> Re an ISO 8601 Basic profile, I'm not sure that the verb "develop"
> applies. I think it would take anyone with a copy of the standard about
> 5 minutes. Then we'd spend 5 weeks discussing how to name it

If we can agree on a name, I am prepared to "write" a Note describing a format that is YYYYMMDD, which is my understanding of one of the particular formats that is being sought (are there others that people wanted?). I suggest we call it YearMonthDay format.

I think Misha's assessment of how much work is required is pretty accurate - armed with a copy of what you want plus Misha's Date Time Format note as a template...

Date: Sun, 14 Aug 2005 13:32:07 +0100
Reply-To: Pete Johnston <p.johnston@UKOLN.AC.UK>
Sender: DCMI Date Working Group <DC-DATE@JISCMAIL.AC.UK>
From: Pete Johnston <p.johnston@UKOLN.AC.UK>
Subject: Draft open date range format (Re: Adding ISO 8601 as a scheme -- discussion)
Comments: To: "Childress,Eric" <childree@oclc.org>
To: DC-DATE@JISCMAIL.AC.UK

Quoting "Childress,Eric" <childree@oclc.org>:

> DC does not "need" another body to develop encoding schemes, but
> I've been given to believe that DCMI prefers to outsource such activity
> if possible. That seems very reasonable to me -- with a primarily
> volunteer organization, development and maintenance bandwidth is always
> limited. Concentrating on DC's core standards is appropriate.
>
> If it is our opinion that this is a case where DC Date needs to
> develop and maintain a scheme then that too can be an appropriate
> activity. I ask only that we consider various alternatives and fix on
> such a course for cause. If we're convinced DC Date is the best hope
> for development of a suitable scheme, we should be able to persuade the
> UB as well.

I must admit that I had thought that it _was_ part of this groups work to "develop" some suggestions for date formats - even if ultimately the DCMI Usage Board recommends that we (or they) need to find a non-DCMI "owner"/maintenance agency for whatever we come up with.

I wasn't at the DC Date WG meeting at DC-2004 but

http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0411&L=dc-date&P=53

says:

===
After lively back-and-forth discussion it was agreed that the Date WG would explore modifying W3CDTF and/or developing a suitable DCDTF application profile to meet the requirements of DC Libraries.
===

>From a purely selfish viewpoint ;-) my own _very_ pressing concern is that this week (gulp!), I must submit a first draft of a specification based on a Dublin Core Application Profile. And in that specification I need to provide a URI for an encoding scheme to support the representation of "open ended" date ranges. It doesn't have to be a DCMI-owned URI, but I need a description of a date format, and a means of referring to it within the framework of the DCMI Abstract Model (as a class or as a datatype).

So, given that this is beyond what is in W3CDTF (for which DCMI provides the class with the URI http://purl.org/dc/terms/W3CDTF) and ISO8601 (and DCMI currently provides no URI for the class of things represented by ISO8601 anyway), it seems to me I/we (DC CD WG) have little option but to create our own specification for that format.

So I've had a go:

http://www.ukoln.ac.uk/metadata/dcmi/date-dccd-odrf/

I've just hacked this out this morning (pillaging from the W3CDTF doc) and I'm sure it needs more work, but, well, it's a start. And as I say, I need something now, even if it gets refined/tidied up later.

Comments welcome!

Note this is _not_ a profile of ISO8601 because ISO8601 doesn't support "open date-ranges". (And on that basis I probably shouldn't use phrasing like "an open date range: a time-interval expressed.... " because I don't think what I am calling an "open date-range" falls within what ISO8601 calls a "period of time" or "time-interval". It's a different concept, I think. But I'm not sure!)

Anyway, as Misha and Charles have suggested, I don't think it would be _that_ difficult to identify the formats required and to produce similar documents to address at least some of the other requirements.

And I think it may be a worthwhile exercise for us to try to do that - and deal with the social/political/administrative issues (who owns the specs, who assigns URIs for the encoding schemes etc etc) separately.

Date: Mon, 22 Aug 2005 16:49:09 +1200
Reply-To: Douglas Campbell <Douglas.Campbell@NATLIB.GOVT.NZ>
Sender: DCMI Date Working Group <DC-DATE@JISCMAIL.AC.UK>
From: Douglas Campbell <Douglas.Campbell@NATLIB.GOVT.NZ>
Subject: Re: Registering ISO 8601
To: DC-DATE@JISCMAIL.AC.UK

Eric, et. al.,

My apologies for only contributing to this list in fits-and-starts, but here's another "fit"... ;-)

As you probably know, I do not support the registration of the whole of ISO 8601 as an encoding scheme, so [luckily] I'm not about to relitigate the email sent to the UB. To add further weight to that email, I would point out that DCMI has a goal of "simplicity of creation and maintenance" (the only reference I could find to this was in section 1.2 of the Usage Guide [1]) - I think we need to find an easier-to-use solution than registering the whole of ISO 8601.

So, looking forward, it appears there are three main contenders:
1. ISO 8601 profile(s)
2. XML Schema dates
3. Um, keep looking...?

I'd like to try attacking the issue from a different angle: If we adopted 8601 profiles or XML schema, what, if any, requirements could be solved?

Warning, long post ahead - I decided to set aside some time for analysis of ISO 8601 and XML Schema dates. I apologise that it's in an email (so harder to read) but I didn't realise I would end up doing all this when I started. But go ahead, jump straight down to the conclusions part at the end of this post now [I know you were going to anyway]... :-)

To revise our requirements from the workplan:
a) B.C.E. dates [DC-Lib; DC CLD WG]
b) Questionable dates [DC-Lib]
c) Approximate dates [DC-Lib]
d) Open-ended date ranges [DC CLD WG]
e) Non-Gregorian dates
f) Large dates (e.g., geologic periods, astronomical time)
g) Soft termini (i.e. the outer bounds for one or more termini is known or can be associated with a known period, but one or both of the exact boundaries of the event referenced are not known)
h) Elapsed time less than date range interval (i.e. the duration is less than the complete interval between two termini, as in an intermittent activity)

And other requirements I can remember:
i) short format without hyphens, eg. YYYYMMDD [DC-Lib]
j) filename-friendly eg. no slash [Douglas Campbell]
k) accuracy to fractions of a second [I just thought of it]
l) date ranges (normal, not open-ended) [DC-General?] - this isn't explicitly in our list but should be as it's one of the reasons the DC-Date WG restarted - I have always contended that these are valid with "W3CDTF" encoded dates anyway, but others disagree

I now have a copy of ISO 8601:2004. It would appear the following would be possible using ISO 8601:
1. Continuation of the W3CDTF format [obviously]
2. Compact format (requirement i) - using 8601's "basic" format
3. Date ranges (requirement l) - we could now explictly cater for them
4. File-friendly (requirement j) - section 4.4.2 notes that "In certain application areas a double hyphen is used as a separator instead of a solidus [slash]" - I'm not sure if that means we can use it, but I quite like the idea of "2001-01-03--2005-06-10", many people find the slash confusing and don't intuitively think it means a range, this reminds me of the Liberty Date Recommendation [2] which decided to replace the T with an underscore just because it was "found visually obscurring and distracting by many users"! eg "2000-02-14_18:42"
5. Large dates (requirement f) - by extending the number of digits allowed in the year component, including positive/negative - though this is only usable where dates can be converted to a specific year, also the number of digits in the year component needs to be predetermined so we may need a

profile for 4 digits (normal), 6 digits, and 8 digits...?
6. BCE dates (requirement a) - by using the positive/negative component of point (5), though some care may be needed in encoding BCE dates (I seem to remember a discussion that 8601 "-0008" may not be 8 BCE?), also it makes a "+" mandatory for "ordinary" CE dates
7. Dates prior to 1583 - years 0000 to 1582 are only allowed to be used in 8601 with mutual agreement
8. Accuracy to fractions of a second (requirement k) - we would need to specify how many decimal places, all places must be included when a fraction is used, eg. "...T03:59:59,030" (comma is 8601's preferred seperator)
9. Some approximate dates (requirement c) - these are only possible to the extent an 8601 date can officially be truncated, eg "1999-06" means some time in June 1999, "1999" means sometime in 1999, "19" means sometime in the 1900s - it's not possible to use "195" meaning the 1950s, or to encode outside the scope of a single unit (century/year/month/day) in a single date expression eg. "sometime in 1920-1930"
10. Maybe open-ended dates (requirement d) - um, this would be seriously bending 8601, but what if we said a date plus a zero-length duration meant open-ended, eg. "1995/P0Y"???? This would clash with valid representations of "0 years from 1995", but a statement of zero years seems to mean very little to me
11. Other 8601 extensions deemed useful - eg. Week number (1995-W02-5), ordinal day in the year (1995-102), duration (1995/P2Y - a period of 2 years beginning from 1995), or abstract duration as suggested by Pete for dcterms:extent (P0Y2M - 2 months)

Of our main 8 requirements (a-h), 8601 profiles can really only satisfy 2 (or so). I should point out that I don't think that is a reason to abandon it, just that it can only be part of the solution. 8601 profiles can be developed fairly quickly (it is just configuring from existing specifications rather than dreaming up the specifications) - aside from the naming part ;-) I think we should make them (as others have noted) as we have been unsuccessful in finding existing ones. Who knows, they may become widely adopted since there doesn't seem to be much competition out there!

I'll move on to the XML Schema soon, but first, a cut at the new profiles..
.
NB: I'm assuming we'll incorporate as many "by mutual agreement" options as possible/necessary as we are defining a mutual agreement.

1. DCDATE (DCDTF?) - the basic date format most people will use, based on W3CDTF with some 8601 extensions.

Format sample: +YYYY-MM-DDThh:mm:ss,sss+hh:mm--+YYYY-MM-DDThh:mm:ss,sss+hh:mm

Notes
* Allow years 0000 to 9999
* Plus or minus prefix so can represent BCE dates (NB: the plus sign is mandatory for CE dates)
* Can add fractions of a second (I chose 3 digits arbitrarily)
* Can drop lower order components (as per W3CDTF) for reduced accuracy (eg. +1999-06), additionally can reduce year to "+YY" indicating century
* "--" for ranges separator (instead of forward slash), though could be confusing (for human readers only) if the second date begins with a minus?

2. DCDATE-COMPACT - same as DCDATE with separators removed.

Format sample: +YYYYMMDDThhmmssssss+hhmm--+YYYYMMDDThhmmssssss+hhmm

3. DCDATE-LARGE - same as DCDATE with extra year places (total 6 digits).

Format sample: +YYYYYY-MM-DDThh:mm:ss,sss+hh:mm--+YYYYYY-MM-DDThh:mm:ss,sss+hh:mm

Notes
* I have arbitrarily chosen 6 digits which may cover a lot of common large dates?
* We may need yet another profile called "very large" with, say, 10 digits???
* Would we also need DCDATE-LARGE-COMPACT for 6 digits but without hyphens?
* I left the full accuracy down to fractions of seconds so it is possible to encode all your dates within this single scheme

For each of these above we would need to consider if we wanted to add in alternatives like week number, ordinal day, or duration (including the zero duration as a hack for open-endedness).

We could also look at a profile for duration, but this is a lower priority.

These three or so profiles would give us a breather by solving some of the more basic date encoding problems. I realise it could result in a whole lot of date data being generated that could be inconsistent with our "final final solution", but we have no date in sight for that - in the meantime at least dates will be being created that comply with a major international Standard.

On the XML Schema front, I have to admit I haven't really used these so it's just from reading the specifications [3]...

xsd:duration
* basically an ISO 8601 duration eg "P8Y" (8 years) or "PT3H" (3 hours)
* appears to extend it by allowing negative duration

xsd:dateTime
* says "inspired" by 8601
* basically looks like an 8601 date (with hyphens and plus/minus prefix) but the number of year digits is completely variable (minimum of 4) and the + prefix for CE dates is not needed (actually it's prohibited)
* all components must be present down to the second (fractions of seconds are optional)
* the timezone (which is optional) must include the minutes

xsd:time
* basically the part of an 8601 date after the "T", including optional fractions of seconds and optional timezone specifier

xsd:date
* basically the part of an 8601 date before the "T" including optional minus prefix
* adds an optional timezone specifier at the end eg "1991-12-05+02:00"

xsd:gYearMonth
* basically an 8601 year and month with 4+ year digits and optional minus prefix
* adds an optional timezone specifier at the end

xsd:gYear
* basically an 8601 year with 4+ year digits and optional minus prefix
* adds an optional timezone specifier at the end

xsd:gMonthDay
* described as "a set of one-day long, annually periodic instances", eg. "3rd of May"
* no minus prefix allowed as the missing year is represented using a hyphen, eg. "--MM-D"
* adds an optional timezone specifier at the end

xsd:gDay
* described as "a set of one-day long, monthly periodic instances", eg. "the 15th of the month"
* no minus prefix allowed as the missing year/month is represented using hyphens, eg. "---D"
* adds an optional timezone specifier at the end

xsd:gMonth
* described as "a set of one-month long, yearly periodic instances", eg. Halloween is in "the month of November"
* no minus prefix allowed as the missing year/month is represented using hyphens, eg. "---D"
* adds an optional timezone specifier at the end

So, XML Schema dates offer a similar array of options to ISO 8601 profiles.

The nice things they offer are
* any number of digits for the year (4+)
* use minus for BCE dates, but don't need plus sign for CE dates

The limitations are basically around the reduced accuracy options:

* no 2 digit century option
* offers year/month/day level accuracy options, but a date plus time must be complete down to the seconds level (ie. can't just go to the hour or minute)

So if we to use them, I would anticipate we'd use:
* xsd:dateTime (YYYY-MM-DDThh:mm:ss)
* xsd:date (YYYY-MM-DD)
* xsd:gYearMonth (YYYY-MM)
* xsd:gYear (YYYY)
Maybe we'd use xsd:duration for dcterms:extent?

Since XML Schema dates are closely related to ISO 8601 dates, they only solve around the same number of our requirements:
a) BCE dates
c) Approximate dates (only to the extent of a single component (year/month/day) in a single date
f) Large dates
k) accuracy to fractions of a second

However, they do not, by themselves, offer date ranges or the short format (no hyphens).

CONCLUSIONS

ISO 8601 profiles:
* need to be defined ...and named ;-)
* probably a standard general-purpose profile plus specialist profiles for particular/advanced usages [samples defined above]
* catering for BCE and large dates requires placeholders/symbols to be added to "ordinary" dates (eg. leading zeros and plus sign prefix)
* offers (closed) date ranges
* offers reduced accuracy levels from seconds up to centuries
* offers a short format option (eg. YYYYMMDD)
* may be able to encode open-ended dates by bending it a little (eg. YYYY-MM-DD/P0Y)
* may be able to offer file-friendly format using double-hyphens instead of the slash in ranges

XML Schema dates:
* are already defined so could be deployed quickly
* offer an elegant method for BCE and large dates (just add digits or add a minus prefix as necessary)
* do not offer date ranges
* do not offer century, hour, or minute levels of reduced accuracy
* do not offer a short format (eg. YYYYMMDD)
* personally I'm not sure about the appropriateness of using a format designed for a particular markup language over an international Standard

Both:
* can only solve 2-3 of our 8 main requirements

Next steps:
* Are the suggested ISO 8601 profiles practical? usable?
* Start multiple threads about naming these profiles ;-]
* Given the low number of requirements 8601/XML schema solve, are either really worth pursuing?
* Should we consider following the XML Schema and Pete's "Open Date Range Format" [4] route of using ISO 8601 only as an "inspiration"?? I guess we could make atomic (single) date expressions to be ISO 8601 compliant but we would build our own structure (syntax) around them for ranges, approximations, questionability, etc.??? This may be able to account for most of the requirements apart from non-Gregorian dates (which I think is worth parking to one side for now).

My thoughts:
* I quite like the idea of specifying ISO 8601 profiles - they can pick up where W3CDTF left off (this route has a precedent) and may well be picked up quite widely as there seems to be a lack of competition!
* But then I also like the idea of a new schema using ISO 8601 atomic dates (best of both worlds) - but this will take a lot longer as it will require (generate) much more debate
* We need to keep the DCMI goal of "simplicity of creation and maintenance" front of mind

Douglas Campbell
National Library of New Zealand

[1] http://dublincore.org/documents/usageguide/#whatis
[2] http://www.lyberty.com/meta/date_format.html
[3] http://www.w3.org/TR/xmlschema-2/
[4] http://www.ukoln.ac.uk/metadata/dcmi/date-dccd-odrf/

Date: Thu, 25 Aug 2005 13:13:03 +1200
Reply-To: Douglas Campbell <Douglas.Campbell@NATLIB.GOVT.NZ>
Sender: DCMI Date Working Group <DC-DATE@JISCMAIL.AC.UK>
From: Douglas Campbell <Douglas.Campbell@NATLIB.GOVT.NZ>
Subject: Re: Registering ISO 8601
To: DC-DATE@JISCMAIL.AC.UK

Misha,

As Kelly noted, comma and period are valid, with a preference for comma. They don't state the reason for that preference, but I assume it is because that is more common globally (as is the 24 hour clock apparently)??

ISO 8601:2004 has a number of places where the number of digits is variable (indicated by underlining the character/position that can have zero or more digits) eg. years or decimal fractions. It's tricky to interpret whether the number of digits must be always present or is just a maximum...

Section 3.5 - Expansion
"By mutual agreement of the partners in information interchange, it is permitted to expand the component identifying the calendar year, which is otherwise limited to four digits. This enables reference to dates and times in calendar years outside the range supported by complete representat ions, i.e. before the start of the year [0000] or after the end of the year [9999]."

Section 3.6 - Leading zeros
"If a time element in a defined representation has a defined length, then leading zeros shall be used as required."

Section 3.7 - Mutual agreement
"Some of the representations identified in this International Standard are only allowed by mutual agreement of the partners in information interchange . Such agreement should ensure that fields in which the representation may occur are not allowed to hold other representations that cannot be unambiguously distinguished from the agreed representation."

Section 5 - Date and time format representations
"Underlining of characters in a date and time format representation, to represent zero or more of the underlined characters in the derived date and time representation (in accordance with 3.4.2), is only permitted if, at the time of interchange of the date and time format representation, the number of characters in the derived date and time representation is not known."

Section 4.2.2.4 - Representations with decimal fraction
"If the magnitude of the number is less than unity, the decimal sign shall be preceded by two zeros in accordance with 3.6."
"The interchange parties, dependent upon the application, shall agree the number of digits in the decimal fraction. The format shall be [hhmmss,ss], [hhmm,mm] or [hh,hh] as appropriate (hour minute second, hour minute, and hour, respectively), with as many digits as necessary following the decimal sign. A decimal fraction shall have at least one digit." [NB: in this quote the last s, m, or h in square brackets is underlined.] Unfortunately the example only includes a single decimal place.

So I can't work out whether the size can vary dynamically or leading (or trailing) zeros must be used. I think there is an ISO 8601 discussion list - maybe we should ask the question there?

I went looking for the XML Schema 1.1 draft but I couldn't find it - just a list of requirements from 2003 and a public discussion listserv?

My comments about XML vs ISO probably stem from my [failed] attempt to use XLink within RDF. I had been trying to define rich links within dc:relatio n in RDF/XML using XLink, but it was pointed out to me that XLink is defined relative to the XML syntax not (necessarily) at a higher/semantic level (eg. RDF). Since then I've been more careful about considering what

context a particular thing is relevant to, especially as Dublin Core metadata can exist in many different contexts/syntaxes other than XML.

Thanx,
Douglas

>>> Misha Wolf <Misha.Wolf@REUTERS.COM> 22/08/05 21:19:05 >>>
Hi Douglas,

Thanks for a very thorough analysis. A few quick notes ...

When we wrote W3CDTF in 1997, we included decimal fractions of a second, using the period (".") as the decimal separator. I'm sure that was allowed at the time by ISO 8601. From what you write, they seem to have changed their minds since then. Are you quite sure about that?

We also allowed an arbitrary number of digits after the decimal separator. Why are you proposing that we choose some fixed number of digits?

Ditto for years.

Have you, by any chance, looked at the XML Schema 1.1 draft? I haven't studied it in any detail, but I have noted that quite a lot of work has gone into the date/time sections.

While I wouldn't suggest that we use XML Schema 1.0, due mainly to the lack of a compact syntax (requested by the Library folks), my eyebrows went up when reading your:

> personally I'm not sure about the appropriateness of using a
> format designed for a particular markup language over an
> international Standard

First of all, it is unusual to call XML "a particular markup language", as it is a foundation for the creation of markup languages. AFAIK, it is only the 2nd widely adopted foundation markup language. Furthermore, I think that XML has been adopted as an ISO standard, though I don't know the number and searching the ISO site and Google for "XML ISO standard" generates too many hits.

Misha

Date: Thu, 25 Aug 2005 15:34:54 -0400
Reply-To: "Rebecca S. Guenther" <rgue@LOC.GOV>
Sender: DCMI Date Working Group <DC-DATE@JISCMAIL.AC.UK>
From: "Rebecca S. Guenther" <rgue@LOC.GOV>
Subject: Re: Registering ISO 8601
To: DC-DATE@JISCMAIL.AC.UK

Just a reality check here. I think we said we would separate the 2 issues:
1. Use of the ISO 8601 basic notation (Douglas calls it "short form")
2. Various expressions of dates that are not covered by ISO 8601

This all started with the DC Library Application Profile (yes, that was the requirement), because there are numerous places in library metadata that use the ISO 8601 basic format in millions of records and we wanted a way to say that the encoding scheme used was that alternative specified in ISO 8601 rather than the more eye readable one with hyphens (which is specified in the profile called W3CDTF).

My proposal to the Usage Board at the time asked for just the notation called ISO 8601 basic and left open the question of other kinds of dates that were needed. I proposed to call it "iso8601" and be specific in the definition that all that is meant is that the encoding uses the alternative also known as ISO 8601 basic. That maybe wasn't a good idea because that is when we got into all the issues of all the other things implied by saying ISO 8601. So maybe at the time I should have called it something else and maybe there would have been some mechanism to specify that my encoding scheme is that YYYYMMDD format. But that was like 3-4 years ago and we're still discussing it.

So leaving aside all the other kinds of dates we want to be able to express, can we please just think of a new name for what I mean as the first step? Or shall I pursue this some other way and not through this

DC-Date group? I know I'm not the only one who needs this.

Date: Thu, 25 Aug 2005 21:41:10 +0100
Reply-To: Misha Wolf <Misha.Wolf@REUTERS.COM>
Sender: DCMI Date Working Group <DC-DATE@JISCMAIL.AC.UK>
From: Misha Wolf <Misha.Wolf@REUTERS.COM>
Subject: Re: Registering ISO 8601
To: DC-DATE@JISCMAIL.AC.UK

Hi Rebecca,

First of all, it is good that we did not define some subset of ISO 8601 and call it "iso8601", as this would be extremely misleading, given the widespread use, by DC communities, of quite a different subset of ISO 8601.

Thinking of a name is not the only issue, though. As has been highlighted by the recent correspondence, there are lots of other issues, eg:

- may durations be used?

- which decimal separator is used?

- is a timezone compulsory?

- in those places where the standard allows a variable number of digits, how many digits may be used?

So, I suggest there are (all together) these tasks:

1. Within the ISO 8601 space:

   - disregarding the compact/verbose form, we should decide the answers to the various questions

   - we should then spin off two profiles of ISO 8601 which differ *only* in whether they use the compact or the verbose form

   I say this because, it doesn't make sense to, eg:

   - use a period as a separator in the verbose scheme, but a comma in the compact scheme

   - allow 4-digit years in the verbose scheme, but 17-digit years in the compact scheme

2. Outside of the ISO 8601 space, we should solve the remaining problems.

Misha

Topic: Attributes of DCMI Terms ("Usage Board profile")
Modified: DATESTAMP
Maintainer: Tom Baker
Latest version: http://www.bi.fhg.de/People/Thomas.Baker/ISSUES/profiles-usageboard/

Shepherd: Tom

Discussion in Shanghai:

In Shanghai, I would like to briefly discuss the following issue in order to define a process for moving this forward. At a minimum, I would like to determine the extent to which these issues should be decided in the Usage Board (as opposed to DC-Architecture or the Directorate). We will not have time in Shanghai to discuss any of these points in detail, so UB members need not read the following as carefully as they might have otherwise. At a most general level, I would like to know how urgent we feel it is to clarify aspects of this problem.

Summary of the issue:

The issue of "DCMI terms describing DCMI terms" has been on the back burner for a long time. We have already in effect defined more than two dozen metadata terms describing various attributes of metadata terms (Name, Label, Definition; types of Status; etc...). However, we have merely documented these terms in Web pages [1,2] -- never have we "declared" the terms formally or assigned them URI references backed by DCMI Namespace Policy. For the purpose of the RDF schemas, we have mapped the handful of attributes most needed for the schemas to existing terms (e.g., in the rdfs namespace maintained by W3C).

We need both to clarify both the status of these terms (perhaps taking the occasion to clean up some of the definitions) and the policy by which the terms will be maintained (if different from the existing DCMI Namespace Policy). We also need to consider whether the terms should be assigned URIs and documented in RDF schemas, as other DCMI metadata terms already are.

According to my notes, we discussed this issue briefly in Ithaca in 2003 and concluded that the following steps would be involved:

1. Define the set of properties and encoding schemes for describing terms.
2. Understand how they relate to existing terms.
3. Ask DCMI Directorate for UB namespace.
4. Set up UB namespace and declare terms as necessary.
5. Define an application profile.

At present, the terms we use are defined in the introductions of two documents - the consolidated document "DCMI Metadata Terms" [1] and the historically complete "DCMI Metadata Terms: a complete historical record" [2]. I have attached a summary of the terms and their definitions below.

I currently see the following issues:

1) We need to look carefully at the RDF schema binding to determine which of the attributes used in [1] and [2] are really needed in the RDF schemas. From my notes, here is a draft mapping, with reference to a hypothetical namespace "dcu:" to hold terms not yet formally declared:

Name:	NOT USED	
Namespace:	rdfs:isDefinedBy	rdf:resource="xxx"
Label:	rdfs:label	xml:lang="en-US"
Definition:	rdfs:comment	xml:lang="en-US"
Type of term:	rdf:type	rdf:resource="http://.../#element"
Status:	dcu:status	rdf:resource="http://.../#recommended"
Date issued:	dcterms:issued	
Comment:	dc:description	xml:lang="en-US"

```
See:               rdfs:seeAlso       rdf:resource="http://..."
References:        dcterms:references rdf:resource="http://.../#W3CDTF"
Refines:           rdfs:subPropertyOf
Qualifies:         dcu:qualifies
Date modified:     dcq:Modified
Decision:          dcu:decision       rdf:resource = "uri"
Version:           dcu:version        rdf:resource = "uri"
Replaces:          NOT USED
Is Replaced By:    NOT USED
Broader Than:      NOT USED
Narrower Than:     rdfs:subClassOf
```

Of course, we need to consider the possibility that not all of the attributes of [1] and [2] would be needed in the RDF schemas.

2) If we accept the mappings of some terms defined in [1] and [2] to existing terms in namespaces maintained by W3C and to DCMI's own Terms namespace, then at a minimum it would appear we would need to declare the following:

```
dcu:status      - Harry needs this for the DCMI Registry
dcu:qualifies
dcu:decision
dcu:version
```

3) In addition, it would appear we need the term

```
dcu:isTranslationOf
```

Harry needs this for the DCMI Registry, and Tom thinks this is needed so that a translation of DCMI term definitions into languages such as Japanese can reference the specific Term Version used as the basis for the translation.

4) The term dcu:status has, in effect, a controlled vocabulary of values:

```
http://dublincore.org/usage/documents/process/#conforming
http://dublincore.org/usage/documents/process/#recommended
http://dublincore.org/usage/documents/process/#registered
```

These are currently defined in the document DCMI Usage Board Process, and the URIs are anchors to specific points in that document. We should consider whether it is a good idea to continue this or whether we would want to declare a status vocabulary, and if so, how their URIs should be formed.

5) The term "Type of Term" (currently mapped in the RDF binding to rdf:type) also has, in effect, a controlled vocabulary of values:

```
http://dublincore.org/usage/documents/principles/#element-refinement
http://dublincore.org/usage/documents/principles/#element
http://dublincore.org/usage/documents/principles/#encoding-scheme
http://dublincore.org/usage/documents/principles/#vocabulary-term
```

6) Work on the DCMI Abstract Model [3] and a formal model for DCMI Application Profiles [4] suggests a need for several other terms, along the lines of:

```
dcu:ApplicationProfile
dcu:PropertyUsage
```

In September 2004, Pete posted a strawman set of terms at http://homes.ukoln.ac.uk/~lispj/cen-cwa/vocab/dcapterms.rdf.

7) DCMI's RDF schemas [5] have long asserted the existence of URI references for terms based on the DCMI Namespace http://purl.org/dc/terms/ -- even though, technically, this should not have been possible without going through UB process. These include:

```
http://purl.org/dc/terms/DateScheme
http://purl.org/dc/terms/FormatScheme
http://purl.org/dc/terms/IdentifierScheme
http://purl.org/dc/terms/LanguageScheme
```

```
       http://purl.org/dc/terms/SpatialScheme
       http://purl.org/dc/terms/SubjectScheme
       http://purl.org/dc/terms/TypeScheme

   We would need to formulate a policy for creating,
   maintaining, and identifying such terms - bearing
   in mind that the terms above are already "legacy"
   (i.e., for all we know, there may be applications
   in the world that would break if DCMI were to
   drop or deprecate these terms).

8) Since the addition of

        http://purl.org/dc/dcmitype/MovingImage
        http://purl.org/dc/dcmitype/StillImage

   we have two new attributes for Vocabulary Terms:

        Broader Than
        Narrower Than - currently represented with rdfs:subClassOf

------------------------------------------------------------------------
Usage Board Application Profile (draft)
------------------------------------------------------------------------

Mandatory
   Name           [1] The unique token assigned to the term.
   URI            [1] The Uniform Resource Identifier used to uniquely
                      identify a term.
   Namespace      [2] The Uniform Resource Identifier of the namespace
                      within which the term is defined.
   Label          [1] The human-readable label assigned to
                      the term.
   Definition     [1] A statement that represents the concept
                      and essential nature of the term.
   Type of term   [1] The type of term, such as Element or Encoding
                      Scheme, as described in the DCMI Grammatical
                      Principles.
   Status         [1] Status assigned to term by the DCMI Usage Board,
                      as described in the DCMI Usage Board Process.
   Date issued    [1] Date on which a term was first declared.

When appropriate
   Comment        [1] Additional information about the term
                      or its application.
   See            [1] A link to authoritative documentation.
   References     [1] A citation or URL of a resource referenced
                      in the Definition or Comment.
   Refines        [1] A reference to a term refined by an Element
                      Refinement.
   Qualifies      [1] A reference to a term qualified by an Encoding
                      Scheme.
   Broader Than   [1] A reference from a more general to a more specific
                      Vocabulary Term
   Narrower Than  [1] A reference from a more specific to a more general
                      Vocabulary Term

Version-related
   Date modified  [2] Date on which a term declaration was subsequently
                      modified.
   Decision       [2] A link to the Usage Board decision describing
                      the creation or modification of a term
                      declaration.
   Version        [2] An historical version of a term declaration.
   Replaces       [2] A reference to the immediately precedent
                      historical version of a term declaration.
   Is Replaced By [2] An identifier for the historical version of a
                      term declaration by which this historical version
                      is superseded.

REFERENCES

[1] http://dublincore.org/documents/dcmi-terms/
[2] http://dublincore.org/usage/terms/history/
[3] http://www.ukoln.ac.uk/metadata/dcmi/abstract-model/
[4] ftp://ftp.cenorm.be/public/ws-mmi-dc/mmidc116.htm
[5] http://dublincore.org/2003/03/24/dcq
```

[6] http://homes.ukoln.ac.uk/~lispj/cen-cwa/vocab/dcapterms.rdf

```
------------------------------------------------------------------------
Strawman vocabulary drafted by Pete Johnston, July 2004
-- http://homes.ukoln.ac.uk/~lispj/cen-cwa/vocab/dcapterms.rdf
   about a hypothetical http://example.org/dcap/
------------------------------------------------------------------------

  dcap:SchemaDocument     http://example.org/dcap/
  dc:title                Schema for the DCAP vocabulary
  dc:description          This schema contains descriptions of the DCAP terms.
                          Terms are declared using RDF Vocabulary Description Language
                          (RDF Schema).
  dc:publisher            http://www.ukoln.ac.uk/#

  dcap:MetadataVocabulary http://example.org/dcap/#
  dc:title                The DCAP Vocabulary
  dc:description          The DCAP Vocabulary provides classes and properties
                          used to describe Dublin Core Application Profiles and Property Usages
                          and related resources.
  dc:publisher            http://www.rdn.ac.uk/#
  dcap:seeAlso            http://www.ukoln.ac.uk/projects/iemsr/wp2/dcap/
  dcap:preferredXMLNamespaceName
                          dcap
  dcap:preferredXMLNamespacePrefix
                          http://example.org/dcap/

  rdfs:Class              http://example.org/dcap/Document
  Label:                  Document

  rdfs:Class              http://example.org/dcap/SchemaDocument
  Label:                  Schema Document

  rdfs:Class              http://example.org/dcap/Agency
  Label:                  Agency

  rdfs:Class              http://example.org/dcap/MetadataVocabulary
  Label:                  Metadata Vocabulary

  rdfs:Class              http://example.org/dcap/AppProfile
  Label:                  Application Profile

  rdfs:Class              http://example.org/dcap/PropertyUsage
  Label:                  Property Usage

  rdfs:Class              http://example.org/dcap/BindingSchema
  Label:                  Binding Schema

  rdfs:Class              http://example.org/dcap/VocabStatus
  Label:                  Vocabulary or Profile Status

  dcap:VocabStatus        http://example.org/dcap/VocabStatus/private
  Label:                  Private

  dcap:VocabStatus        http://example.org/dcap/VocabStatus/draft
  Label:                  Draft

  dcap:VocabStatus        http://example.org/dcap/VocabStatus/proposedRecommendation
  Label:                  Proposed Recommendation

  dcap:VocabStatus        http://example.org/dcap/VocabStatus/recommendation
  Label:                  Recommendation

  rdfs:Class              http://example.org/dcap/TermStatus
  Label:                  Vocabulary or Profile Status

  dcap:TermStatus         http://example.org/dcap/TermStatus/private
  Label:                  Private

  dcap:TermStatus         http://example.org/dcap/TermStatus/unstable
  Label:                  Unstable

  dcap:TermStatus         http://example.org/dcap/TermStatus/testing
  Label:                  Testing

  dcap:TermStatus         http://example.org/dcap/TermStatus/stable
  Label:                  Stable
```

```
  dcap:TermStatus         http://example.org/dcap/TermStatus/deprecated
  Label:                  Deprecated

  rdfs:Class              http://example.org/dcap/Obligation
  Label:                  Obligation

  dcap:Obligation         http://example.org/dcap/Obligation/reserved
  Label:                  Reserved

  dcap:Obligation         http://example.org/dcap/Obligation/optional
  Label:                  Optional

  dcap:Obligation         http://example.org/dcap/Obligation/recommended
  Label:                  Optional (Recommended)

  dcap:Obligation         http://example.org/dcap/Obligation/mandatory
  Label:                  Mandatory


  rdf:Property            http://example.org/dcap/uses
  Label:                  Uses
  rdfs:range              http://www.w3.org/1999/02/22-rdf-syntax-ns#Property

  rdf:Property            http://example.org/dcap/encodingScheme
  Label:                  Encoding Scheme

  rdf:Property            http://example.org/dcap/obligation
  Label:                  Obligation
  rdfs:range              http://example.org/dcap/Obligation

  rdf:Property            http://example.org/dcap/condition
  Label:                  Condition

  rdf:Property            http://example.org/dcap/maxOccurs
  Label:                  Maximum Occurrences

  rdf:Property            http://example.org/dcap/isMemberOf
  Label:                  Is Member Of

  rdf:Property            http://example.org/dcap/seeAlso
  Label:                  See also
  rdfs:range              http://example.org/dcap/Document

  rdf:Property            http://example.org/dcap/version
  Label:                  Version

  rdf:Property            http://example.org/dcap/status
  Label:                  Status

  rdf:Property            http://example.org/dcap/isExpressedBy
  Label:                  Is Expressed By
  rdfs:range              http://example.org/dcap/BindingSchema

  rdf:Property            http://example.org/dcap/preferredXMLNamespaceName
  Label:                  Preferred XML Namespace Name

  rdf:Property            http://example.org/dcap/preferredXMLNamespacePrefix
  Label:                  Preferred XML Namespace Prefix
```

---

5. Attributes for describing DCMI Terms ("UB application profile")

   -- Properties for Describing Terms [Tom]: At some point, we should formally declare terms such as "Status" (as well as the Status types: "recommended", "conforming"...).

      ACTION: Tom to explore possibility of coordinating with SKOS community in this regard. Berlin conference in April will provide opportunity to speak with ISO11179 community as well; will report on results in May.