

Declaring FRBR entities and relationships in RDF

Background

The FRBR Review Group meeting of 21 August 2007, held at the World Library and Information Congress in Durban, South Africa, initiated a new project with the specific task "To define appropriate namespaces for FRBR (entity-relationship) in RDF and other appropriate syntaxes."¹

FRBR (entity-relationship), or FRBR(ER), is a reference to the final report of *Functional requirements for bibliographic records*.²

RDF is *Resource description framework*³, a set of specifications developed by the World-Wide Web Consortium (W3C) to provide a means of modelling metadata in a variety of formats (syntaxes). The format usually associated with RDF is Extensible markup language (XML); the RDF/XML syntax is often referred to simply as RDF.

A namespace is an abstract container for a set of terms (names, words, etc.) which provides context for the terms and allows disambiguation for similar terms held in different namespaces. The set of terms is known as a vocabulary. The term "namespace" is variously used to refer to the complete content (terms, definitions, structure) or just the web domain name used to provide machine-readable locations for the content. "Namespaces" may refer to logical divisions of a single namespace, or multiple namespaces.

An XML namespace provides unique labels (identifiers) for the names of elements and attributes used in documents encoded with XML, including RDF/XML. An XML namespace must conform to the recommendation of the W3C⁴.

It is therefore necessary to use XML namespaces for the vocabularies of FRBR entities and relationships in order to be compatible with RDF.

The FRBR vocabularies form part of an ontology, which can be described using RDF schema (RDFS). RDFS⁵ is also a recommendation of the W3C. RDFS requires the use of Uniform Resource Identifiers (URIs, similar to URLs) to designate machine-readable locations for the content of the namespace.

A major impetus behind the project was the formation of the DCMI RDA Task Group.⁶ This group aims to define components of *RDA: resource description and access*⁷ as an RDF vocabulary for use in developing a Dublin Core application

¹ FRBR Review Group meeting report Durban, August 21, 2007. Available at: http://www.ifla.org/VII/s13/wgfrbr/FRBR_RG_Mtg2007.pdf

² Functional requirements for bibliographic records: final report. 1998. Available at: <http://www.ifla.org/VII/s13/frbr/frbr.pdf>

³ Resource description framework (RDF). Available at: <http://www.w3.org/RDF/>

⁴ Namespaces in XML 1.1. 2nd edition, 2006. Available at: <http://www.w3.org/TR/xml-names11/>

⁵ RDF vocabulary description language 1.0: RDF schema. 2004. Available at: <http://www.w3.org/TR/rdf-schema/>

⁶ DCMI/RDA Task Group wiki. Available at: <http://dublincore.org/dcmirdataskgroup/>

⁷ RDA: resource description and access. Available at: <http://www.collectionscanada.gc.ca/jsc/rda.html>

profile. RDA is based on FRBR, and RDA metadata attributes are mapped to FRBR entities. The Scholarly works application profile (SWAP)⁸ also uses elements of FRBR. SWAP is currently developed by the DCMI Scholarly Communications Community⁹.

Requirements

As well as meeting the technical requirements referenced above, some governance and administration requirements must be met if a namespace is to be effective and attractive to professional communities:

- There must be clarity on ownership, governance and administration of the namespace.
- The content of the namespace must be kept current.
- Changes should, nevertheless, be minimised.
- Version control, change logs and change dissemination must be properly managed and operated.
- The base web domain, and the URIs derived from it, must be permanent. If the owner of the namespace changes, the base web domain must be transferred to the new owner.

Investigation

The National Science Digital Library Metadata Registry¹⁰ provides services to the developers of controlled vocabularies, and was chosen as a suitable tool for the project. Specifically, it offers:

- A sandbox or "play" facility for developing familiarity with the registry interface and exploring various approaches to creating namespaces.
- A database approach to storing namespace components which can be mapped to different output syntaxes.
- Database maintenance facilities for adding, amending and deleting namespace components.

These facilities were used to create three separate FRBR vocabularies:

1. FRBR entities
2. FRBR relationships
3. FRBR user tasks

The last vocabulary is an extension to the scope of the project. It was added because:

- User tasks are referenced in RDA, and it may be useful to have formal declarations of their labels and definitions.
- No significant extra effort was required to create the vocabulary and populate it with four terms.

⁸ SWAP. Available at: <http://www.ukoln.ac.uk/repositories/digirep/index/SWAP>

⁹ DCMI Scholarly Communications Community. Available at: <http://dublincore.org/groups/scholar/>

¹⁰ NSDL registry. Available at: <http://metadataregistry.org/>

Desk research was carried out to identify previous initiatives involving FRBR namespaces.

A brief analysis was carried out to identify any overlap between the project's proposed namespaces and those previously created.

Documentation about the project was placed in a subsection of the wiki of the DCMI RDA Task Group¹¹. Subsequently, a note about the project with links to the subsection was included in the wiki frontpage.

Informal project progress reports were circulated to selected members of the FRBR Review Group and the consultants identified at the Durban meeting.

Results

The three FRBR vocabularies^{12 13 14} are currently available in the NSDL sandbox as simple "list" vocabularies rather than schemas. This approach was taken to help understand the relationship between vocabularies and schemas, and the impact of translations of FRBR into other languages. They will be moved to the NSDL registry and declared as provisional schemas over the next few weeks.

The vocabularies cannot be formally declared without establishing a web domain name for the basis of both namespaces and identifiers for each term. However, third-party developers testing models involving FRBR elements should be able to use the place-holder domains from the sandbox and provisional "live" schemas. A simple find-replace processing of the domain name in resulting XML documents will usually be sufficient update them.

Definitions and scope notes pertaining to each vocabulary term were taken verbatim from FRBR. As a result, the content is not as regular in layout and tone as would be found in, say, a technical dictionary.

As noted above, SWAP has declared some FRBR elements in namespaces. Its namespace for entity types¹⁵ refers to all six FRBR Group 1 and Group 2 entities. It uses different labels, however, for "Work" (i.e. "ScholarlyWork") and "Item" (i.e. "Copy"), although the definitions are not changed. Several FRBR relationships are also included in its namespace for properties.¹⁶

A complete RDF namespace and schema for FRBR entities and relationships was created in 2005 by Ian Davis and Richard Newman¹⁷. It includes two additional entities (classes): Endeavour, a super-class of all Group 1 entities; and

¹¹ Namespace for FRBR entities/elements in RDF. Available at: http://dublincore.org/dcmirdataskgroup/namespace_20for_20FRBR_20entities_2felements_20in_20RDF

¹² FRBR entities. Available at: <http://sandbox.metadatareregistry.org/vocabulary/show/id/49.html>

¹³ FRBR relationships. Available at: <http://sandbox.metadatareregistry.org/vocabulary/show/id/90.html>

¹⁴ FRBR user tasks. Available at: <http://sandbox.metadatareregistry.org/vocabulary/show/id/69.html>

¹⁵ Eprints EntityType vocabulary encoding scheme. Available at: http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_EntityType_Vocabulary_Encoding_Scheme

¹⁶ Eprints terms. Available at: http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Terms

¹⁷ Expression of core FRBR concepts in RDF. 2005. Available at: <http://vocab.org/frbr/core>

ResponsibleEntity, a super-class of all Group 2 entities. Some definitions are identical to those in FRBR, while others are variations. An RDF namespace for extended FRBR entities and relationships¹⁸ was also created by the same authors in 2005. Both works carry Creative Commons licenses requiring copyright and license notices be kept intact, and credit be given to copyright holder and/or author.

The Davis and Newman works are a very good example of the components of a properly-formed "namespace":

- Version control.
- A human-readable vocabulary of classes (e.g. FRBR entities) and properties (e.g. FRBR relationships).
- A machine-readable schema. In this case, the schema uses Web ontology language (OWL)¹⁹, another W3C initiative. OWL is compatible with RDFS.

The DCMI RDA Task Group has considered the option of creating namespaces for FRBR entities and relationships within its own domain (<http://RDVocab.info>), but prefers to keep its focus as tight as possible (i.e. on RDA vocabularies) and follow semantic web ideas of re-using resources from other namespaces when appropriate. The Task Group is currently not sure whether it needs to use FRBR namespaces, but welcomes this initiative of the FRBR Review Group.

Implications

Much of the technical development required for a FRBR namespace has already been carried out by Davis and Newman. However, the content does not exactly match FRBR(ER), and there issues of ownership, governance and administration need to be resolved.

Implications arising from FRBRoo need further investigation and resolution; these may impinge, for example, on the super-classes created by Davis and Newman.

It may be possible to persuade SWAP to replace its partial FRBR namespace with one under the control of the FRBR Review Group.

The DCMI RDA Task Group would use a FRBR namespace, if required.

There are three basic options for creating the base web domain for a FRBR namespace:

1. Piggy-backing on a related namespace domain (e.g. RDVocab.info)
2. Use an open service for such domains (e.g. PURL)
3. Create a "branded" domain specifically for FRBR (e.g. FRBRVocab.info, IFLA.info, etc.)

The last option encourages views of diversity and collaboration within the semantic web community.

¹⁸ Expression of extended FRBR concepts in RDF. Available at: <http://vocab.org/frbr/extended>

¹⁹ Web ontology language (OWL). Available at: <http://www.w3.org/2004/OWL/>

Recommendations

1. The FRBR Review Group should create a suitable "branded" namespace domain.
2. The FRBR Review Group should accept and operate appropriate commitments to ownership, governance and administration requirements.
3. The creation and implementation of a FRBR namespace should build on the work currently being carried out by the DCMI RDA Task Group related to RDA namespaces, including continuing use, as appropriate, of the NSDL Metadata Registry. The author of this report is co-Chair of the Task Group, and has a specific role in coordinating the work of the Task Group with related activities elsewhere.
4. The FRBR Review Group should consider the implications of FRBRoo for the namespace for FRBR(ER), ensuring that the FRBR namespace based on FRBR(ER) can be integrated with any future development of a FRBRoo namespace. This activity should be carried out in collaboration with the CRM community.
5. A namespace for FRBR(ER) should be implemented before extension to FRBRoo, to allow workflows, administrative policies and practices, etc. to be developed for the simpler case.
6. The FRBR Review Group should initiate discussions with existing users of FRBR vocabularies, specifically Davis and Newman, and the DCMI Scholarly Communications Community.

Gordon Dunsire
25 July 2008