Interoperability and semantics in RDF representations of FRBR, FRAD and FRSAD

Gordon Dunsire

Presented at the Cologne Conference on Interoperability and Semantics in Knowledge Organization "Concepts in Context"

19-20 July 2010

Background (1)

- Functional Requirements for Bibliographic Records (FRBR) published in 1998
  - Developed by a Study Group of the International Federation of Library Associations and Institutions (IFLA)
  - Incomplete – did not address "authorities"
    - Entities used as access points to bibliographic records
- Functional Requirements for Authority Data (FRAD) published in 2009
  - Incomplete – did not address subject authorities

Background (2)

- Functional Requirements for Subject Authority Data (FRSAD)
  - Next presentation ...
- IFLA's FRBR Review Group will develop a consolidated model from the FR "family"
  - Process is now underway
  - Will be informed by the analysis required for Semantic Web compatibility (representation in RDF)

Background (3)

- RDA: resource description and access is based on FRBR and FRAD
  - Bibliographic metadata content guidelines
- DCMI RDA Task Group asked to develop an RDF representation of RDA
  - Following the "London" meeting in April 2007
- Stimulated creation of "FRBR namespaces project"
  - To develop RDF representation of FRBR

Background (4)

- At the same time ... Consolidated edition of International Standard Bibliographic Description (ISBD) in development
    - Structure and content guidelines
    - ISBD/XML Study Group to develop an XML representation of ISBD
    - Decision to use RDF/XML
- IFLA Namespaces Task Group set up to identify requirements/options for support of IFLA standards in the Semantic Web
    - Report submitted; further discussion in Aug 2010

Methodology (1)

- ✧ NSDL Metadata Registry used for basic representation in RDF
  - ✧ Same approach as RDA
  - ✧ URI assignment; labels, definition, scope note, property range and domain
- ✧ FR family are entity-relationship models
  - ✧ Entity => RDF Class
    - ✧ E.g. FRBR Work
  - ✧ Entity attribute => RDF property
    - ✧ E.g. FRBR has-intended-audience
  - ✧ Entity-entity relationship => RDF property
    - ✧ E.g. FRBR is-realized-through (Work-Expression)

Methodology (2)

- Terminology of labels, definitions and scope notes based as closely as possible on source documentation
  - E.g. Property label = "has" + attribute name
- Classes and properties declared for prior model(s) re-used in current model
  - E.g. FRAD re-uses URIs from FRBR namespace
    - E.g. FRBR Expression; FRBR has-key
- No re-use of RDF resources from external community namespaces such as FOAF
  - Equivalences may be established later

Methodology (3)

- ✧ FR attribute properties have domain but no range
  - ✧ Required by the model for extensibility
  - ✧ E.g. FRBR has-key (domain = Work)
- ✧ Relationship properties have domain and range
  - ✧ E.g. FRBR is-realized-through (domain = Work; range = Expression)

General, non-technical issues

- ✦ **Consistency within documentation**
  - ✦ Written for human consumption; evidence of variable phrasing to make it more readable
    - ✦ E.g. "prior", "preceding", "first"
- ✦ **Consistency of labels of RDF resources**
  - ✦ E.g. "has a reproduction" (regular) vs "has reproduction" (irregular)
- ✦ **Documentation refers to sub-types of entity**
  - ✦ E.g. "musical work", "serial"
  - ✦ First pass: sub-type => sub-class
  - ✦ But wrong due to semantic overlap

Opaque URIs

- ✧ Opaque URIs are used
  - ✧ E.g. http://iflastandards.info/ns/fr/frbr/frbrer/1001
    - ✧ frbrer:1001
  - ✧ Not http://iflastandards.info/ns/fr/frbr/frbrer/Work
    - ✧ not frbrer:Work
- ✧ IFLA operates in a multilingual environment
  - ✧ Anglophone bias avoided
  - ✧ Labels, etc. in English (@en), but no problem in accommodating translations
- ✧ Allows subsequent changes to alternative and preferred labels without causing confusion (URI must not change)

Semantic issues

- Do differences in documentation reflect real semantic differences?
- Close examination and detailed discussion required
  - Example (minor): Work
    - FRBR: "A distinct intellectual or artistic creation."
    - FRAD: "A distinct intellectual or artistic creation (i.e., the intellectual or artistic content)."
  - Example (major): Person
    - FRBR: "An individual."
    - FRAD: "An individual or a persona or identity established or adopted by an individual or group."

Ontological issues (1)

- Source documentation only identifies pairs of inverse properties
  - For relationship properties only
    - E.g. is-realization-of/is-realized-through
  - Attribute properties are not inverted because instance triple objects are assumed to be literals
- Analysis of transitive, asymmetric, disjoint, etc. property types required
  - E.g. has-an-alternate is symmetric (implies the inverse is-an-alternate-to is redundant)
- All FRBR classes are mutually disjoint

Ontological issues (2)

- ♦ Relationships between separate FR models are likely to be declared with equivalence and hierarchical properties
  - ♦ E.g. owl:sameAs, rdfs:subClassOf
  - ♦ E.g. FRAD class Corporate Body seems to be a sub-class of FRBR Corporate Body
- ♦ Likely to be published as an addendum to the existing FR models
- ♦ Will inform the consolidated model
  - ♦ Which may also require additional classes and properties

RDA issues (1)

- DCMI RDA Task Group has declared parallel FRBR classes within RDA namespace
  - Could not wait for FRBR RDF resources to be approved
    - RDA to decide whether to substitute FRBR namespace resources, or declare equivalence
- Conflict with FRBR in RDA implementation scenarios?
  - RDA Manifestation "embodies" Work and Expression; FRBR allows only Expression

RDA issues (2)

 ✦ FRBR identifies sets of entities (classes) as Groups 1, 2 and 3, but not intended to be super-classes

   ✦ They simplify the entity-relationship diagrams

   ✦ But relationships are to be interpreted as being between individual entities and not the group

   ✦ So sets of relationships are required

 ✦ RDA is discussing the declaration of such super-classes to simplify and reduce the properties

   ✦ As in the Davis/Newman version of FRBR (2005)

ISBD

- ✧ ISBD has only one class (implied)
  - ✧ Resource: likely to be a super-class of FRBR Work, Expression, Manifestation, and Item
- ✧ Attribute properties, but no relationship properties (between Resources)
  - ✧ No range assumed; no inverse properties
  - ✧ To do: Mapping to FRBR properties
- ✧ ISBD Content form and media type mapped to RDA content and carrier types (all controlled vocabularies)
  - ✧ Via RDA/ONIX framework for resource categorization

Improving interoperability

- At least 3 namespaces in Libraryland will have RDF representations of attributes and relationships
  - FRBR/consolidated; ISBD; RDA
  - Interoperability will improve the quality and quantity of linked-data instances
- Interoperability should be improved by:
  - The Vocabulary Mapping Framework matrix
  - Coherent and consistent management environment of IFLA namespaces
  - Output of W3C Library Linked Data Incubator Group.

Thank you

- gordon@gordondunsire.com
- FRBR Review Group
  - http://www.ifla.org/en/frbr-rg
- ISBD Review Group
  - http://www.ifla.org/en/isbd-rg
- NSDL Metadata Registry
  - http://metadataregistry.org/
- Vocabulary Mapping Framework matrix
  - http://cdlr.strath.ac.uk/VMF/documents.htm
- Library Linked Data Incubator Group
  - http://www.w3c.org/2005/Incubator/lld/