Representing the FR family in the Semantic Web

Gordon Dunsire, Independent Consultant, Edinburgh, Scotland.
gordon@gordondunsire.com

This is an Author's Original Manuscript of an article submitted for consideration in Cataloging & Classification Quarterly [copyright Taylor & Francis]; Cataloging & Classification Quarterly is available online at http://www.tandfonline.com/10.1080/01639374.2012.679881.

Abstract

Each of the FR family of models has been represented in Resource Description Framework (RDF), the basis of the Semantic Web. This has involved analysis of the entity-relationship diagrams and text of the models to identify and create the RDF classes, properties, definitions and scope notes required. The work has shown that it is possible to seamlessly connect the models within a semantic framework, specifically in the treatment of names, identifiers, and subjects, and link the RDF elements to those in related namespaces.

Keywords

Cataloging standards, Metadata standards, Bibliographic data – interoperability, Data models – cataloging research, Entity-relationships models – cataloging research

Background/Introduction

At its meeting[1] during the 2007 World Library and Information Congress in Durban, South Africa, the FRBR Review Group initiated a project, now known as the FRBR Namespaces Project, "to define appropriate namespaces for FRBR (entity-relationship) in RDF and other appropriate syntaxes". This followed discussion at the Data Model meeting[2] held April 30-May 1, 2007 at the British Library in

London, where one of the recommendations was to undertake the development of an RDA Element Vocabulary. This activity would necessarily involve the FRBR entity-relationship model on which RDA: resource description and access[3] is based, and it seemed appropriate that relevant parts of this activity should be carried out by the FRBR Review Group as maintainers of the model.

RDF, vocabularies, and namespaces

RDF is Resource Description Framework[4], "a standard model for data interchange on the Web." The model is essentially concerned with the representation of data about some thing which itself may be composed of data; that is, data about data, or metadata. Metadata is given in the form of a relationship between a value (data) and the thing it refers to. As only one relationship, one value, and one referrant is allowed in each RDF statement, we can regard such statements as "atomic", in contrast to "molecular" metadata records composed of sets of statements giving values for different aspects of the same referrant. More simply, RDF deals with the constituent parts of what is traditionally regarded as a metadata record.

RDF labels the three components of a statement as the subject (referrant), predicate or property (relationship), and object (value), in that order. ("Predicate" is used in RDF theory, and "property" in RDF practice.) An example is "This article has author Gordon Dunsire"; "this article" is the subject, "has author" is the predicate, and "Gordon Dunsire" is the object. Other examples are "This book has subject 'mathematical physics in Scotland in the 19th century'", "This film has sound mode silent", "This work is a derivation of that work", etc. It is usually possible to represent any bibliographic record as a set of such statements, with some rearrangement of the metadata it contains. An RDF statement is usually called a "triple" to reflect its three-part construction.

RDF is designed to take advantage of the infrastructure of the World-Wide Web and harness its information storage and communication capabilities for machine-processing metadata. To achieve this, the component parts of a triple need some form of identification that is machine-compatible. Using human labels like "this book" or "mathematical physics in Scotland in the 19th century" is not effective because of the inherent ambiguity of human language. We use the same label for different things, and different labels for the same thing, and apply our intelligence to disambiguate them in specific contexts. Computers have no such intelligence, and cannot distinguish, for example, the label "green" in different contexts such as colour, golf, or personal names. Instead, RDF exploits the infrastructure of the Web by using a special form of identifier, the uniform resource identifier or URI. In principle, a URI is any unique combination of letters, numbers, and punctuation, although some punctuation marks are not allowed. An expanded form of URI, the international resource identifier or IRI, is being developed to include non-Latin scripts such as Chinese.

A particular form of URI, the http URI, exploits the existing http uniform resource locator or URL, familiar as the basis of hyperlinks in the Web of documents. An http URI looks just like a URL, but it does not necessarily act like a URL. For example, the URI for the Library of Congress Subject Heading (LCSH) "mathematical physics" is http://id.loc.gov/authorities/subjects/sh85082129. This is not a URL, but in this case it has been made to act like one if it is entered into a normal web browser. This is called "de-referencing"; when the same URI is entered into a semantic web browser, it returns machine-readable RDF statements rather than a human-readable web page.

RDF requires that the subject and predicate parts of a triple are URIs. The object part can be a URI or a literal, such as a string of text characters or a date. The utility of this requirement is discussed later in the paper. The work-horse part of the triple is the predicate or RDF property which relates the object to the subject. Properties are equivalent to the attributes of a metadata schema, such as "author", "title", "subject", etc. It is helpful to turn these labels into verbal phrases to understand them as relationships: "has author", "has title", "is subject of", etc. In fact, these are really attributes of a generalized kind of thing, or class of bibliographic entity, such as "book", "article", "film", etc which can be represented as an RDF class. Types of relationship between such entities are also represented as RDF properties, such as "has summary", "is contained in", "has sequel", etc. These labels for attributes and relationships, together with the labels for the generalized entity classes, can be attached to URIs and are known as an "element set vocabulary", or just "element set". Similarly, labels for values of the object part of a triple, such as "Gordon Dunsire", "mathematical physics", etc. can also be attached to URIs, and are known as a "value vocabulary". Element sets are akin to metadata schemas, and value vocabularies to "authority" controlled terminologies.

These vocabularies, then, consist of machine-readable URIs attached to human-readable labels. It is important to also attach the URIs to human-readable definitions so that the meaning, or semantic, of a URI can be understood by human programmers. The element sets and vocabularies are usually developed and maintained using central management infrastructures, as with metadata schemas and terminologies. It is more efficient to use the same http domain as the basis of the URIs, in the same way as for multi-page websites where one URL can have many sub-folders. A common base domain combined with a unique local identifier is a good way of forming a URI like the LCSH example given above: "http://id.loc.gov/authorities/subjects/" is the base domain and "sh85082129" is the local part of the URI http://id.loc.gov/authorities/subjects/sh85082129. All LCSH URIs have the same base domain, and the collection of all URIs sharing the same base domain is called a "namespace". The scope of the term is often expanded to include all triples with such URIs as the subject.

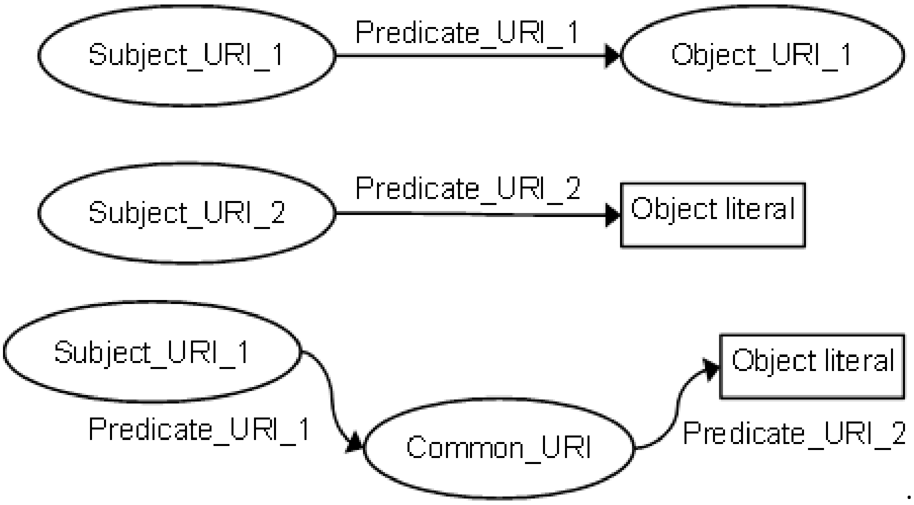
Another advantage of using a common base domain for namespace URIs is that the URI can be shortened using a mnemonic abbreviation, called a "qname", for the base domain. For example, the LCSH base domain above can be abbreviated to "lcsh" and the URI given as lcsh:sh85082129. The URI is reconstituted automatically when it requires processing, using a technique developed for extensible markup language (XML).

RDF graphs

RDF triples can be expressed and displayed in several different ways, known as serializations. A common serialization uses the syntax of XML to represent triples in RDF/XML syntax. Triples can also be displayed as a graph of lines connecting nodes: a single triple is shown as a line or arc, representing the predicate or property, connecting two nodes representing the subject and object.

The connecting arc is usually given an arrow pointing to the object node to distinguish it from the subject.

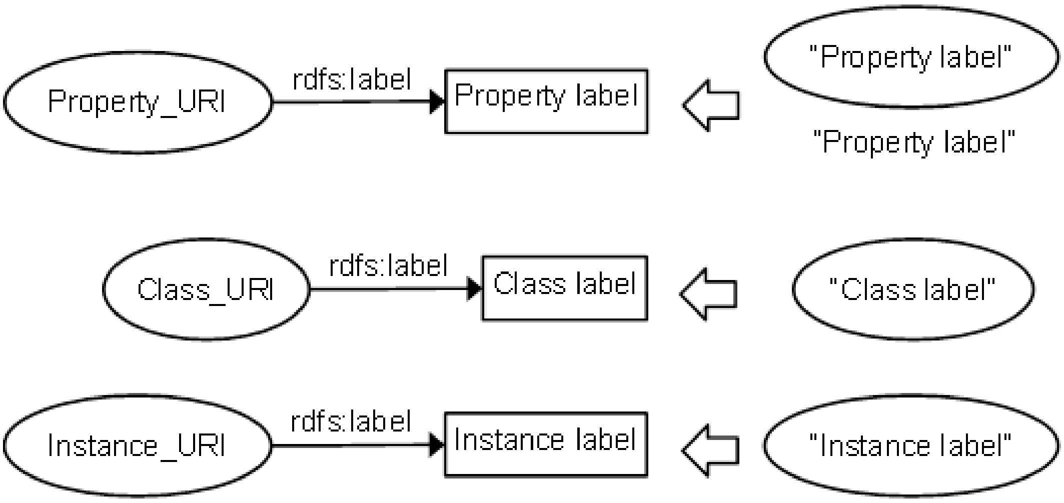
There is a convention, used in the RDF primer[5] and elsewhere, that a node that is a URI is shown as an oval, while a node that is a literal is shown as a rectangle. A subject node is always an oval, because it must be a URI, while an object node may be an oval or a rectangle. Thus any single triple can be displayed as two ovals connected by an arc when the object part is a URI, or as an oval and a rectangle connected by an arc when the object part is a literal. These two basic graphs are shown in Figure 1.



[Figure 1: The two types of graph for any single RDF triple, depending on the object being a URI or literal; and a combined graph created when Object_URI_1 and Subject_URI_2 are the same URI, here given as Common_URI for clarity.]

Although a URI can only refer to one thing, several URIs may refer to the same thing; that is, one thing may have more than one URI. Figure 1 shows how graphs can be combined if two nodes have URIs that identify the same thing; that is, if the URI is the same, or two different URIs are know to refer to the same class, property, concept or value, or specific instance of a person, expression, place, or other entity. The two nodes are merged, forming chains if they are the subject and object of different triples, as in the figure, and forming nets if they are both the subject of different triples.

It is possible for a predicate or property URI to be the subject of a triple; the property used by that triple is thus a property of a property. This is show in the first triple of Figure 2.



[Figure 2: Property, class, and instance URIs related to human-readable labels using the rdfs:label property, with the corresponding proposed display form for labelled RDF graph nodes and links.]

Figure 2 shows three RDF graphs of triples which use the "label" property from the RDF Schema (RDFS) namespace[6] to attach human-readable labels to property, class, and instance URIs. Here, rdfs:label is a URI (and "label" is its label). In the first triple, it is the property of another property with the URI Property_URI. To avoid endless self-referential confusion, RDF graphs in this paper will usually be shown with labels rather than URIs; this will be indicated by enclosing the label in quotes, echoing the convention for indicating literal values.

FR namespaces

The first challenge of the FRBR project was to obtain a base domain for one or more namespaces; multiple namespaces can be created by using the same base domain plus one or more sub-folders before adding the local part to form a specific URI. The project anticipated that namespaces would also be eventually required for other models in the Functional Requirements (FR) family – for Authority Data (FRAD) and subject authority data (FRSAD), and for other bibliographic standards maintained by the International Federation of Library Associations and Institutions (IFLA).

After checking the Domain Name System (DNS) for availability, IFLA was registered as the owner of the http domain "http://iflastandards.info." Although URIs do not need to carry any extrinsic meaning, there are some advantages to embedding a recognizable "brand" in the base domain to promote the organization and indicate trustworthiness in the RDF element sets and value vocabularies it maintains. However, trust cannot be extended to instance triples which use such classes, properties, and concepts because of the AAA principle in RDF: "Anybody can say anything about anything". The source and validity of instance triples, and other areas of provenance, are currently the focus of work to extend the RDF[7].

A small ad hoc group within IFLA agreed on a common set of sub-folders to be used with the http domain to delineate separate namespaces as and when required, and taking into account the potential use of the http domain to store documentation and provide web services associated with the namespaces. Thus all IFLA namespaces use a base domain starting "http://iflastandards.info/ns/" followed by a mnemonic for the specific standard. The FR family domains are "http://iflastandards.info/ns/fr/frbr/frbrer"[8] (to distinguish between the entity-relationship and object-oriented versions of the FRBR model, "http://iflastandards.info/ns/fr/frad"[9], and "http://iflastandards.info/ns/fr/frsad"[10]).

Opaque URIs

It is possible to incorporate a human-readable label into the local part of a URI, as with the property rdfs:label with the full URI http://www.w3.org/2000/01/rdf-schema#label. The FRBR Review Group decided to avoid this approach, and use so-called opaque URIs instead; such a URI is not meaningful to a human. The Group wishes to avoid language bias in URIs intended for global application, given that the FR documents have been translated into several language world-wide. Instead, the Group is encouraging translation of the labels, definitions, and scope notes of the RDF classes and properties in each namespace. Spanish translations of labels have already been added; the URI frad:P2005 has the English label "is created by" and the Spanish label "es creador por". The Group also encourages programmers and developers to check definitions of the classes and properties in its namespaces, rather than relying on labels such as "Item" which can be confused with concepts from other models.

RDF representation of the FR family of models

The FRBR Review Group decided from the outset that separate namespaces should be used for each of the FR models, to reflect their historical development. Where FRBR elements are re-used with little or no modification by FRAD and FRSAD, the RDF representations are similarly re-used within the relevant namespaces.

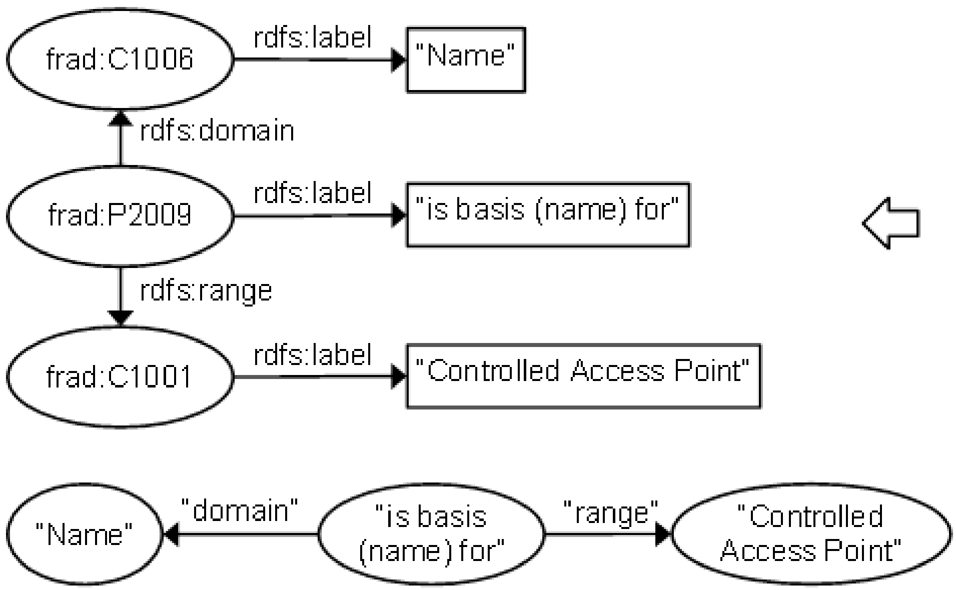
Some details of the work to develop RDF representations of the FR family of models have already been described[11 12]. Generally, an entity such as FRAD's "Name" is represented as a class: frad:C1006 with label "Name". An attribute of an entity, such as FRAD's "dates of usage", is represented as a property: frad:P3025 with label "has dates of usage". A relationship between two entities, such as FRAD's "is basis for", is also represented as a property: frad:P2009 with label "is basis (name) for". Labels for FR classes and properties use the terminology and phrasing found in the models, to allow the documentation to be used as back-up reference for the namespaces.

Domains and ranges

The domain of a property (a property of a property) is an RDF class to which the subject of any triple based on the property is assumed to belong. If property_URI_1 has the domain of Class_A, the triple "this_URI property_URI_1 that_URI" implies that this_URI refers to something that is a member of Class_ A. In the FR namespaces, all properties based on attributes have the domain of the class corresponding to the parent entity; thus FRAD's "dates of usage" property has the class "Name" as its domain.

Similarly, the range of a property is the class to which the object of a triple based on the property is assumed to belong. If property_URI_1 has the range of Class_B, the same example triple implies that

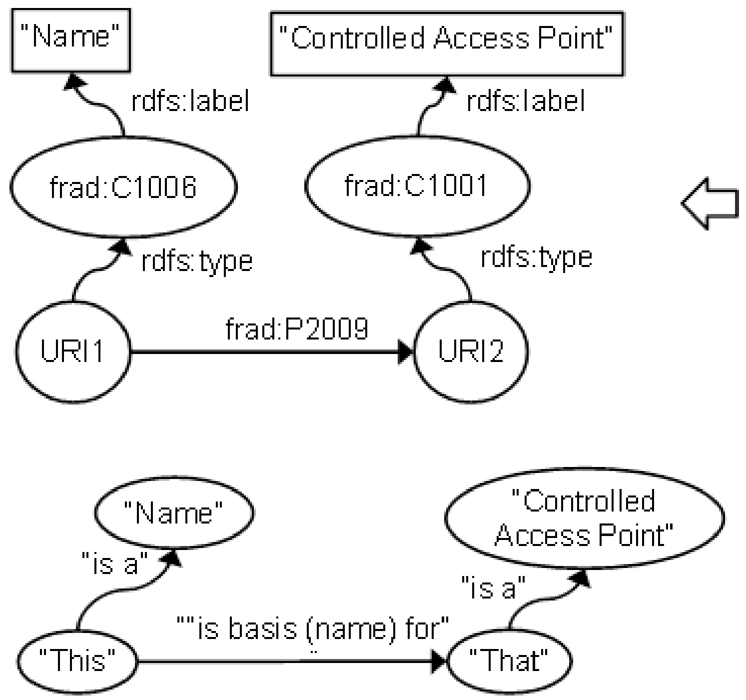
that_URI is a member of Class_B. In the FR namespaces, properties based on entity relationships have domains and ranges of the corresponding classes. FRAD's "is basis (name) for" has the class "Name" as domain, and the class "Controlled Access Point" as range. Figure 3 gives the RDF graph for this property and its domain and range.



[Figure 3: RDF graph for the domain and range of the FRAD property "is basis (name) for", in URI and label display forms.]

It is easy to misinterpret the graph in Figure 3 as being a single statement, and therefore a single triple: "Name is basis (name) for Controlled Access Point". This is, in fact, an entity-relationship structure represented by six RDF triples. The three nodes in the graph do not even form a proper chain, as can be seen by tracing the direction of the property arrows.

In contrast, Figure 4 shows the same property used in an instance triple. The RDF graph is shown in URI and label forms.



[Figure 4: RDF graph for an instance triple using the FRAD property "is basis (name) for", linked to the classes defined as the domain and range of the property, in URI and label display forms.]

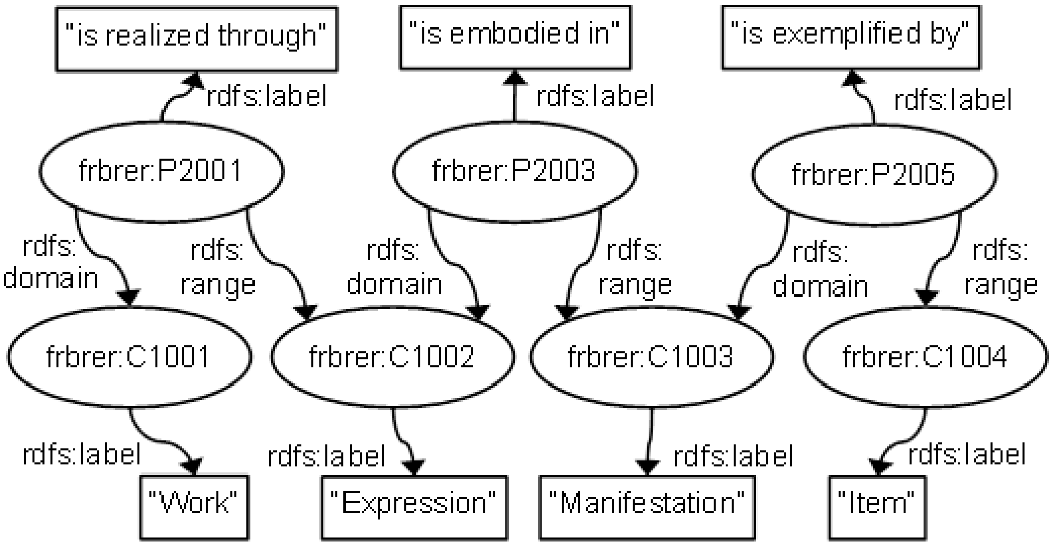
The graph in Figure 4 represents a single instance triple. The domain and range of the property associated with the triple imply that the subject of the triple is a member of the class of the domain, and the object is a member of the class of the range. This is represented using the property rdf:type, which is often labelled "is a". Thus the URI of the subject of the instance triple is a member of the class labelled "Name" and the URI of the object is a member of the class with the label "Controlled Access Point"; "This is a Name" and "That is a Controlled Access Point" are inferred or entailed triples.

The difference between the graphs in Figure 3 and Figure 4 is important. The first graph shows semantic relationships between RDF classes and properties; the second graph shows instance triples based on RDF classes and properties.

In the FR namespaces, no range is given for a property based on an attribute, so it can take either a literal or URI as its object. The FR models do not specify any representation or format for the values of attributes.

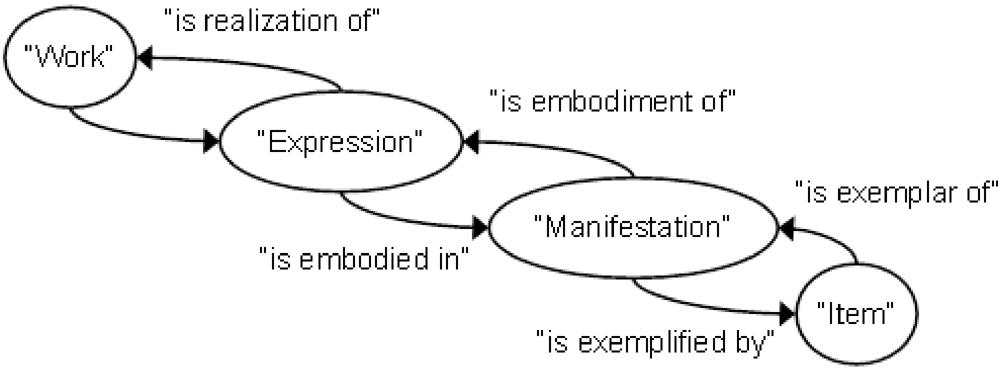
Interpreting entity-relationship diagrams

All three FR models use entity-relationship diagrams which have a similar appearance to RDF graphs, with labelled arrowed lines connecting labelled rectangles. However, there is no direct correspondence between the graphic representations, and each diagram requires decomposition into simple statements before a full RDF graph can be constructed.



[Figure 5: The RDF graph of the FRBR Group 1 entities and primary relationships, showing URIs.]

Figure 5 gives the RDF URI graph for the FRBR Group 1 entities and primary relationships labelled in figure 3.1 of the FRBR model[13]. The bi-directional arrows in the entity-relationship diagram indicate that each RDF property should have an inverse; in RDF graphs, property arcs are uni-directional to preserve the subject-predicate-object order of the triple. The set of inverse properties has the same graph as Figure 5 with the domains and ranges swapped over, and with different, distinctive property labels.



[Figure 6: The RDF graph of the FRBR Group 1 entities and primary relationships and inverses, with human-readable labels.]

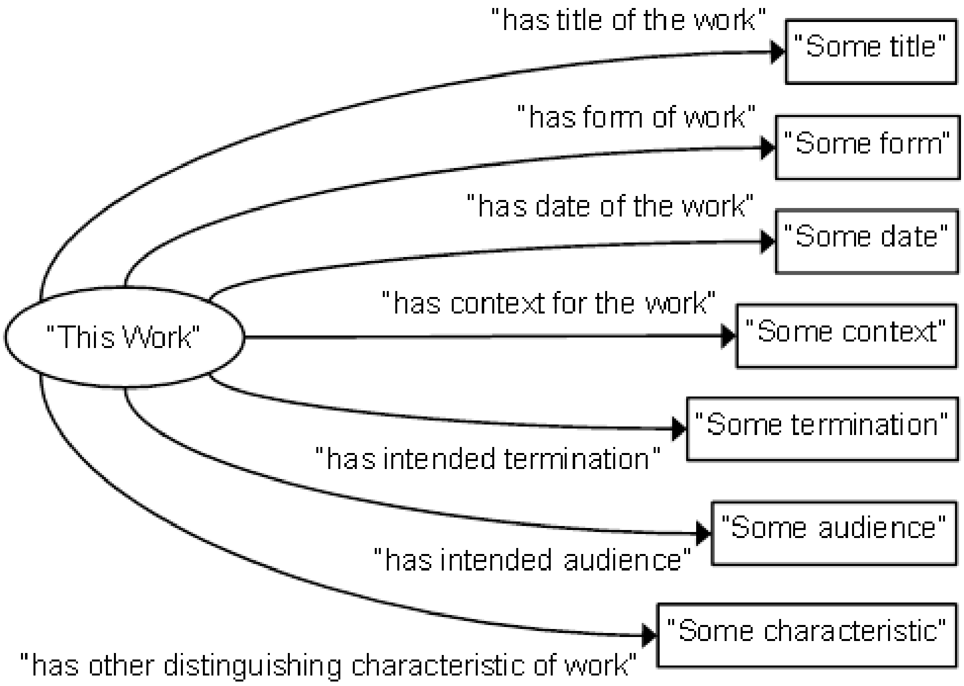
Figure 6 combines the graph of Figure 5 with the corresponding graph for the inverse properties to produce, in label form, an augmented version of FRBR's Figure 3.1. As with Figure 3, the graph in Figure 6 is a representation of an entity-relationship model, and is not a set of three triples joined together in both directions.

The number of arrows in the FRBR entity-relationship diagram indicate one-to-many and many-to-one restrictions on the relationships. These cannot be directly represented using the simple RDF properties of properties given by RDFS. Instead, they have been described using the properties of Web Ontology Language (OWL) as cardinality restrictions.

Attributes and additional relationships

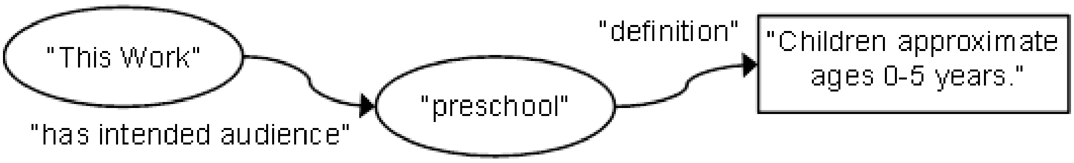
The attributes of each FR entity are represented as RDF properties with the entity class as domain, and no specified range. These properties can be used in triples carrying data about a specific instance of the entity.

Figure 7 gives an RDF graph of instance triples using all the attributes of FRBR's class Work. In each case, the subject of the triple is the URI of a specific work labelled "This Work", and the object is a literal.



[Figure 7: An RDF graph, with labels, of a set of instance triples based on every attribute of the FRBR Work entity.]

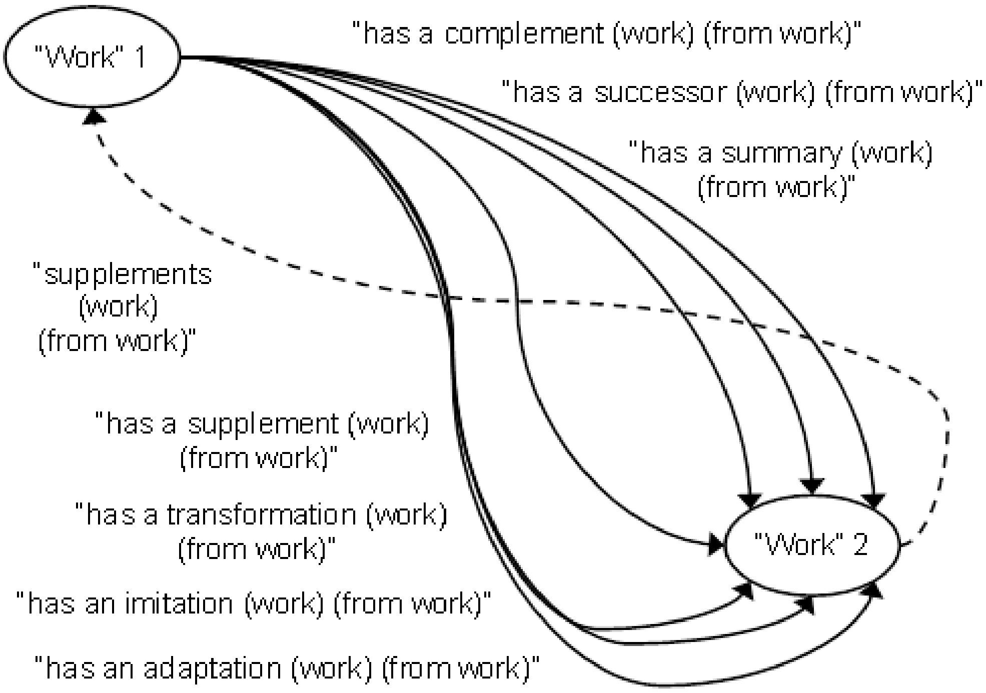
Figure 7 is effectively the graph of a complete "record" for the specified work. Another set of instance triples with different literal values for a different subject URI, thus identifying a different work, would constitute the record for that other work. The RDF properties allow a URI as the object, so it is possible to use the URI of a term or concept from a controlled vocabulary as the object of an instance triple based on, say, the "has intended audience" property, such as the MARC21 Target audience vocabulary[14]. This creates linked data, connecting a specific work to, say, a definition of its intended audience, as shown in Figure 8.



[Figure 8: RDF graph, with labels, of a linked chain of instance triples.]

The FR models describe relationships between entities which are additional to those depicted in the entity-relationship diagrams. These additional relationships are represented as RDF properties with domains and ranges corresponding to the entities they relate. In most cases the entities are of the same type, and so the domain and range of the property is the same RDF class.

Figure 9 shows the RDF graph of possible instance triples based on the FRBR additional relationships between two works. All relationships are bi-directional, requiring a pair of inverse RDF properties for each. Only one pair of reciprocal properties is shown in the graph; for the other relationships, only one uni-directional property is given.



[Figure 9: RDF graph, with labels, of a set of instance triples based on every additional relationship in FRBR between two works. Only one triple from each inverse pair is shown, with a single example of a reciprocal property.]

The graph in Figure 9 is semantically inconsistent because only one relationship, if any, will be valid between any two works; an imitation is not a summary, supplement, transformation, etc. This restriction has been represented in RDF using the OWL property owl:propertyDisjointWith. If two properties are disjoint, they cannot have the same domain and object when used in instance triples.

The graph becomes self-contradictory only when this restriction is added to it; otherwise, it is perfectly valid.

The FR models give only general characteristics of the entities and relationships they describe. Applications of the models are expected to refine and extend them as appropriate. In RDF, refinement is represented by developing sub-classes and sub-properties with narrower definitions and more specific constraints; an example is the RDA property "Dramatization of (Work)" represented in the RDA namespace[15], which can be considered a sub-property of FRBR's "has an adaptation (Work) (from Work)". Extension is represented by developing additional classes and properties; an example is the RDA "Place of origin of the work" attribute for the Work entity[16], which is not described in the FRBR model.

Consolidating the models

The development of separate namespaces has allowed each of the FR models to be treated on its own merit. However, FRAD and FRSAD use entities and attributes from FRBR, and sometimes redefine them. An example is the entity Corporate Body, which is defined in FRBR as "An organization or group of individuals and/or organizations acting as a unit" and in FRAD as "An organization or group of persons and/or organizations identified by a particular name acting as a unit". According to FRAD, a corporate body must have a name, making the corresponding RDF class semantically distinct from that given in the FRBRer namespace. The FRAD namespace therefore includes its own class for Corporate Body and does not re-use the FRBRer URI. The restriction on the FRAD class is represented in OWL. The FRAD class refines the FRBRer class, and is therefore a sub-class: "frad:C1003 rdfs:subClassOf frbrer:C1006" or "frad:Corporate Body is sub-class of frbrer:Corporate Body". In another example, the FRAD property "has derivative", which relates two works, has a broader definition than the FRBRer property "has a summary (work) (from work)". It is therefore a super-property of the FRBRer property: "frbrer:P2049 rdfs:subPropertyOf frad:P2017" or "has a summary (work) (from work) is sub-property of has derivative".

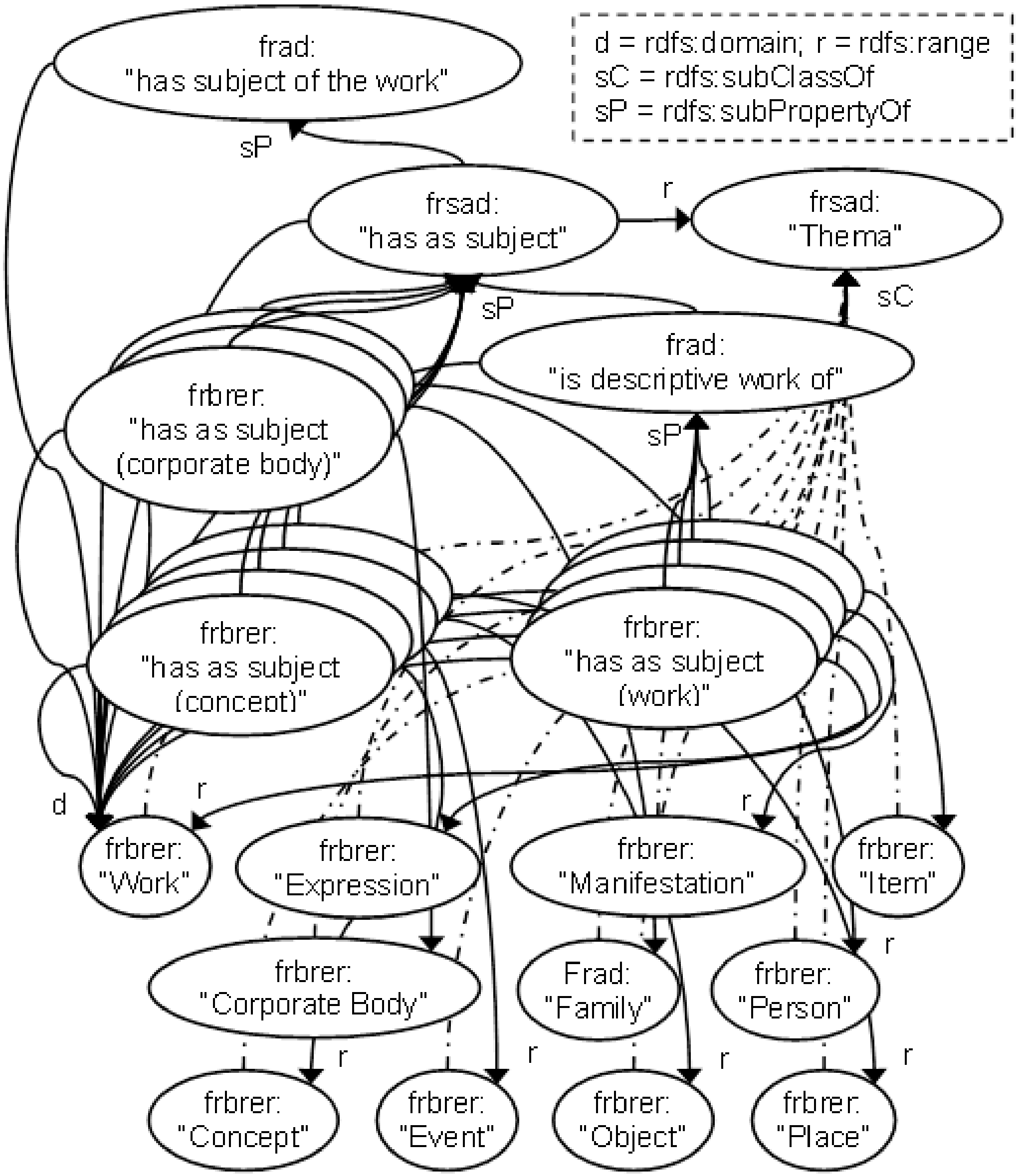
FRAD sometimes gives a different label to a re-used FRBR property, although the definition and other semantic aspects are identical. This is represented in RDF using the property skos:altLabel from the Simple Knowledge Organization System (SKOS) namespace[17]. For example, "frbrer:P3042 skos:altLabel 'has other information associated with the person'", where the existing RDFS label for the FRBRer property is "has other designation associated with the person". Similarly, FRAD may give additional examples of the scope of a definition, which are represented in RDF by an additional scope note. For example, "frbrer:P3003 skos:scopeNote 'That is, the first date associated with the work.'" adds a scope note to the FRBRer property "has date of the work".

This process is informing the work of the FRBR Review Group in developing a consolidation of the three FR models, by revealing redundancy betwen models, and cross-categorization of elements, attributes, and relationships. Two major areas of interest have been the treatment of subjects and titles/names in the different models.

Treatment of the subject entity

Each of the FRBR, FRAD, and FRSAD models contain entities, attributes, or relationships associated with subjects. FRBR introduces the "has as subject" relationship between the Work entity and all of the entities defined in Groups 1, 2, and 3. FRBR Group 3 is not represented in RDF as a class, following clarification from the FRBR Review Group: the Groups are used to simplify the entity-relationship diagrams, and are not intended to be super-classes. Instead, 10 separate properties are represented in RDF, all with domain Work and each with one of the Group 1, 2, or 3 entity classes as range, corresponding to Figure 3.3 in the FRBR report. An eleventh property has been added, with the FRAD class Family as range. All 11 properties have reciprocal or inverse properties with the domain and range swapped over. FRAD has the attribute "has subject of the work", represented in RDF as a property with domain Work and no specified range. There is no corresponding inverse property; there is no range to swap over as the domain of an inverse property. Examination of the definition of the FRAD work-to-work relationship "is descriptive work of" reveals that it is also concerned with subjects; it is represented in RDF as a property with domain Work and range Work. The subject entity is, of course, the main focus of FRSAD, which has the relationship "has a subject" represented in RDF as a property with domain Work and range Thema. FRSAD has an inverse relationship, "is subject of", with an RDF property with domain Thema and range Work.

Figure 10 shows the RDF graph of FR classes and properties associated with subjects, and how they can be linked using RDFS properties to produce a semantically coherent super-graph. For clarity, the FRBRer properties are stacked in sets corresponding to Groups 1, 2, and 3, and the RDFS property labels are abbreviated. Figure 10 does not show the inverse FR properties, but they can also be linked using RDFS in a coherent way, and do not affect the inferences that can be made.



[Figure 10: RDF graph of FR classes and properties associated with subjects. Classes and properties are linked with RDFS properties with shortened labels. Inverse FR properties are not shown.]

The graph in Figure 10 uses rdfs:subPropertyOf to link two FR properties. Such a link is coherent if the domain and/or range of the sub-property is a sub-class of the domain and/or range of the super-property, or if the super-property has no domain and/or range. The FR properties all have domain Work, so no incoherency results when linking any of them as sub- and super-properties. However, the FRSAD "has as subject" property must be a sub-property of the FRAD "has subject of the work" property because the former has range Thema and the latter has no range; conversely, the FRAD property cannot be a sub-property of the FRSAD one. The same logic can be used to infer new triples. If the FRBRer properties are represented as sub-properties of FRSAD's "has as subject", then the range of each of the FRBRer properties can be inferred to be a sub-class of Thema, the range of the FRSAD property. Thus the 10 FRBRer classes described in Groups 1, 2, and 3, plus FRAD's Family, are all sub-classes of Thema. The inferred or entailed triples are given in Figure 10 as dot-dashed lines. The graph also shows that the FRAD property "is descriptive work of" is semantically redundant. It carries the same meaning as the FBRBer property "has as subject (work)" and can be removed from the graph without loss of information. It can thus be removed from a future consolidated FR model.

The graph hints at another interesting conclusion: that the FRSAD class Thema is redundant in RDF. Thema is defined as "any entity used as a subject of a work". But anything can be the subject of a work, so the class Thema can have anything as a member, whether it is a class ("the subject of this work is controlled access points in authority records"), an instance ("the subject of this work is the controlled access point for William Shakespeare), or a property ("the subject of this work is the relationship between controlled access points and names"). In other words, Thema is equivalent to the RDFS class Resource, defined as "the class of everything". It is also equivalent to the OWL class Thing. Strictly speaking, Thema should not be represented as a class in the FRSAD namespace as it has no intention; its set of attributes cannot be used to determine whether any one thing is a member of the class or not. It also has no extension; the list of things that are members of the class is infinite. However, there is an argument for retaining Thema in a consolidated FR model for the sake of clarity.

Treatment of titles and names

The treatment of titles and names in the FR models is more complicated than the treatment of subjects. FRBR has "title" as an attribute of the work, expression, and manifestation entities, represented as three properties in RDF with domains "Work", "Expression", and "Manifestation", but no range. FRBR has "name" as an attribute of the person and corporate body entities, extended by FRAD to the family entity, also represented as three properties with domains "Person", "Corporate Body", and "Family", and again with no range. But FRAD treats name as an entity, with the corresponding RDF class "Name" and sub-classes for "Name of a Corporate Body", "Name of a Family", and "Name of a Person". These can be linked to the corresponding FRBR properties by making each of them the range. FRAD also has a sub-class for "Name of a Work", which suggests a crossover between names and titles: the FRAD sub-class can be made the range of the FRBRer "has title of the work" property. FRSAD has the RDF class "Nomen" which is related to "Thema" via the "has appellation" property, and has the definition "any sign or sequence of signs (alphanumeric characters, symbols, sound, etc.) by which a thema (any entity used as a subject of a work) is known, referred to, or addressed as". All names, titles, labels, etc. attached to a subject are therefore instances of the class "Nomen", but as has already been argued, anything can be a subject, so the labels, etc. of everything are instances of "Nomen". This indicates that there may be several areas of redundancy that might be eliminated from a consolidated FR model, and work is ongoing to complete the analysis. What is clear, however, is that the definition of "Nomen" should exclude URIs. If the URI of a specific "Thema" is treated as a specific "Nomen", then that URI is the subject of a triple based on the FRSAD property "has appellation", but it is also the literal object of the same triple: "this_thema_URI has appellation 'this thema URI'". But the triple must have the URI of the specific "Nomen" as its object because of the property's range, so it needs a URI for "this thema URI". At best, this is completely redundant; at worst, it leads to URIs of URIs, identifiers of identifiers, and an infinite regress.

Other namespaces

There are other namespaces to which FR classes and properties can be linked. RDA has already been mentioned; as an instantiation of the FRBR model, with plans to incorporate FRAD and FRSAD, it refines and extends the FR namespaces. The consolidated edition of the International Standard

Bibliographic Description (ISBD)[18] describes in the introduction how it is related to FRBR. Indeed, many of the IFLA standards developed to support "universal bibliographic control" are related in some way or another[19]. A namespace for the ISBD element set[20] has been published, and plans are underway to develop a namespace for UNIMARC[21]. The Scholarly Works Application Profile (SWAP)[22] is partly based on FRBR, and includes RDF classes corresponding to several in the FRBRer namespace[23]. SWAP re-uses several properties from the Dublin Core metadata element set, which is refined by the Dublin Core Medata Terms (DCT)[24]. The FR models themselves can be treated as refinements of DCT. For example, the FRAD property "has subject of the work", which is the broadest property in the graph of Figure 10, can be represented as a sub-property of the DCT property "subject".

Methods of creating crosswalks between such graphs and namespaces are discussed in a paper[25] presented at the 2011 Dublin Core conference. The paper provides an illustration of linking the FRBRer property "has extent of the carrier" to the namespaces of RDA, ISBD, DCT and the Bibliographic Ontology (BibO).

Conclusion

It is difficult to think of anything that can be described bibliographically that cannot be connected to a single super-graph of all bibliographic metadata statements. Are there any works that are not related in some way to some other work, for example as derivations or descriptions? Are there any items that are not part of some bibliographic collection to which they can be linked via collection-level description? Every item, in any case, must be linked to a work via a manifestation and expression according to the FRBR model.

RDF representations of the FR models offer a powerful way of relating them to other bibliographic schemas and supporting the transformation of metadata statements from one format to another.

Linking the element sets of each of the FR models has helped the process of consolidating the models by identifying areas of redundancy and semantic incoherency.

The focus in RDF on the atomic metadata statement whose semantics are distinct from the syntax used to carry the (meta)data represents a paradigm shift from the focus on bibliographic records that has driven the development and application of standards for well over a century. The cataloguer of tomorrow may well be involved in creating and maintaining single metadata statements in the context of a very large RDF graph, the Semantic Web, rather than somewhat disconnected, fixed assemblages of metadata in records.

References

[1] IFLA Cataloguing Section. FRBR Review Group. "Meeting report Durban, August 21, 2007". http://www.ifla.org/files/cataloguing/frbrrg/meeting_2007.pdf

[2] British Library Metadata Services Standards. "Data Model Meeting, British Library, London 30 April - 1 May 2007". http://www.bl.uk/bibliographic/meeting.html

[3] Joint Steering Committee for Development of RDA. "RDA: resource description and access". http://www.rda-jsc.org/rda.html

[4] W3C RDF Working Group. "Resource Description Framework (RDF)" (2004). http://www.w3.org/RDF/

[5] W3C. "RDF primer: W3C Recommendation 10 February 2004". http://www.w3.org/TR/rdf-primer/

[6] W3C. " RDF Vocabulary Description Language 1.0: RDF Schema: W3C Recommendation 10 February 2004". http://www.w3.org/TR/rdf-schema/

[7] W3C Provenance Incubator Group. "Provenance XG final report: W3C Incubator Group report 08 December 2010". http://www.w3.org/2005/Incubator/prov/XGR-prov-20101214/

[8] "FRBRer model". http://metadataregistry.org/schema/show/id/5.html

[9] "FRAD model". http://metadataregistry.org/schema/show/id/24.html

[10] "FRSAD model". http://metadataregistry.org/schema/show/id/26.html

[11] Gordon Dunsire and Mirna Willer. "Standard library metadata models and structures for the Semantic Web", *Library hi tech news*, vol. 28, no. 3. (2011): , http://dx.doi.org/10.1108/07419051111145118

[12] Gordon Dunsire. "Interoperability and semantics in RDF representations of FRBR, FRAD and FRSAD", *Concepts in Context: Proceedings of the Cologne Conference on Interoperability and Semantics in Knowledge Organization, July 19th–20th, 2010*. (Ergon Verlag, 2011). [In print]

[13] IFLA Study Group on the Functional Requirements for Bibliographic Records. "Functional requirements for bibliographic records: final report". Amended and corrected. 2009. http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf

[14] "MARC21-008: Target audience".http://metadataregistry.org/vocabulary/show/id/277.html

[15] "Element sets: RDA relationships for works, expressions, manifestations, items: Elements: Dramatization of (Work)". http://metadataregistry.org/schemaprop/show/id/423.html

[16] "Element set: RDA Group 1 elements: Elements: Place of origin of the work". http://metadataregistry.org/schemaprop/show/id/19.html

[17] W3C. "SKOS Simple Knowledge Organization System namespace document - HTML variant: 18 August 2009 Recommendation edition". http://www.w3.org/2009/08/skos-reference/skos.html

[18] "ISBD : International standard bibliographic description". Consolidated ed.(Berlin ; München : De Gruyter Saur, 2011) (IFLA series on bibliographic control ; vol. 44). ISBN 978-3-11-026379-4
[19] Gordon Dunsire. "UNIMARC, RDA and the Semantic Web. Meeting: 135. UNIMARC. World Library And Information Congress: 75th IFLA General Conference And Council, 23-27 August 2009, Milan, Italy". http://www.ifla.org/files/hq/papers/ifla75/135-dunsire-en.pdf
[20] "ISBD elements". http://metadataregistry.org/schema/show/id/25.html
[21] Gordon Dunsire and Mirna Willer. "UNIMARC and linked data. Meeting: 187 - Advancing UNIMARC: alignment and innovation, - IFLA UNIMARC Programme (UNIMARC), World Library And Information Congress: 77th IFLA General Conference And Council, 13-18 August 2011, San Juan, Puerto Rico". http://conference.ifla.org/past/ifla77/187-dunsire-en.pdf
[22] "Scholarly works application profile". http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly_Works_Application_Profile
[23] "Eprints entity type vocabulary encoding scheme". http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_EntityType_Vocabulary_Encoding_Scheme
[24] Dublin Core Metadata Initiative. "DCMI metadata terms". 2010. http://dublincore.org/documents/dcmi-terms/
[25] Gordon Dunsire, Diane Hillmann, Jon Phipps and Karen Coyle. "A reconsideration of mapping in a semantic world", *Proc. Int'l Conf. on Dublin Core and Metadata Applications 2011*. http://dcevents.dublincore.org/index.php/IntConf/dc-2011/paper/view/52/6