Multi-Entity Models of Resource Description in the Semantic Web: A comparison of FRBR, RDA, and BIBFRAME

Thomas Baker, Department of Library and Information Science, Sungkyunkwan University, Seoul, Korea
Karen Coyle, consultant, Berkeley, California, USA
Sean Petiya, School of Library and Information Science, Kent State University, Kent, Ohio, USA

(Preprint)

Abstract

Bibliographic description in emerging library standards is embracing a multi-entity model that describes varying levels of abstraction from the conceptual work to the physical item. Three of these multi-entity models have been published as vocabularies using the Semantic Web standard Resource Description Framework (RDF): FRBR, RDA, and BIBFRAME. The authors test RDF data based on the three vocabularies using common Semantic Web-enabled software. The analysis demonstrates that the intended data structure of the models is not supported by the RDF vocabularies. In some cases this results in undesirable incompatibilities between the vocabularies, which will be a hindrance to interoperability in the open data environment of the Web.*

Data Files

The data files supporting this study are available at: http://lod-lam.slis.kent.edu/wemi-rdf/

Introduction

Most bibliographic metadata on the Web, such as data describing a book, article, or image, follows the implicit model of a single entity (a "resource") with attributes (properties). This model is reflected, for example, in the widely used Dublin-Core-based XML format of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Over the past two decades, however, the library world has developed more differentiated models of bibliographic resources. These models do not see a book as just a book, but as a set of entities variously reflecting the meaning, expression, and physicality of a resource.

Readers with a background in modern library science will recognize Figure 1 as depicting the four "Group 1" entities – Work, Expression, Manifestation, and Item (WEMI) – defined in the specification Functional Requirements for Bibliographic Records (FRBR). The WEMI entities are connected in a daisy chain cascading from the (abstract) Work to a (concrete) information artifact, the Item, and vice-versa. Since its creation by a working group of the International Federation of Library Associations (IFLA) in the 1990s, the FRBR model has been incorporated into Resource Description and Access (RDA),

Published in: Library Hi Tech, v. 32, n. 4, 2014 pp 562-582 DOI:10.1108/LHT-08-2014-0081

designated successor to the standard Anglo-American Cataloging Rules (AACR). In parallel to RDA, the Bibliographic Framework Initiative of the US Library of Congress (BIBFRAME) has adopted a FRBR-like model with just two entities as the basis for its draft successor to the Machine Readable Cataloging (MARC) format: an (abstract) Work is instantiated in a (concrete) Instance.

The emergence since 2000 of the Semantic Web idea, and since 2006 of the Linked Data cloud, have led the maintainers of these multi-entity bibliographic models to publish the models as vocabularies expressed in Resource Description Framework (RDF) and Web Ontology Language (OWL), the standardized Semantic Web languages of the World Wide Web Consortium (W3C). (This paper does not distinguish OWL "ontologies" from other RDF vocabularies.) FRBR and RDA were translated into RDF from source specifications that use non-RDF formalisms, while the BIBFRAME vocabulary was born RDF.

The objective of this paper is to explain, in plain English, what the RDF vocabularies for FRBR, RDA, and BIBFRAME say about the nature of their bibliographic entities. The analysis compares how the bibliographic entities are defined as RDF classes with particular attention to how those classes are differentiated from other bibliographic classes and from the book, article, and image classes of the more widespread single-entity models of bibliographic description. In the cases of FRBR and RDA, the paper evaluates the RDF vocabularies against the non-RDF source specifications.

The paper points out that regardless of their intended uses, RDF vocabularies do not of themselves specify integrity constraints for data validation and quality control. Indeed, the validation of RDF data is widely recognized as a crucial gap in the Semantic Web technology. A W3C activity under development aims at standardizing an approach to RDF validation that takes into account pre-standard work by several companies and initiatives, notably a DCMI working group on RDF-based application profiles for meeting the requirements of cultural heritage. Using RDF vocabularies with such emerging approaches to RDF validation would allow the RDF vocabularies for multi-entity bibliographic models to be significantly simplified while enhancing their usability outside of the library silo. Although the RDF vocabularies reviewed here are occasionally discussed at conferences or on mailing lists, the authors have seen no substantial comparisons of the RDF vocabularies in the literature. If multi-entity bibliographic models are to be the future of bibliographic description, the issues raised here must be resolved before further resources are committed to their use.

Structure of this paper

The paper begins by explaining how the “open-world” context of Linked Data differs, in fundamental ways, from the “closed-world” context of the relational database and document format technologies that have dominated library technology since the 1960s. The paper then presents FRBR, RDA, and BIBFRAME, the RDF expressions of which are reviewed for usability in a Linked Data environment by means of simple, annotated data examples. The examples draw out essential differences in how the vocabularies characterize bibliographic entities — what the entities are, how they relate to each other, and how they relate to entities in other models. A discussion section elaborates on the contrast between hitherto dominant information technologies and the Semantic Web approach, drawing out requirements for data validation which, in the authors’ interpretation, are implied by the RDF expressions of the models.

The final section of the paper draws conclusions about the design of bibliographic data. It argues that RDF vocabularies — the dictionaries of data elements underpinning bibliographic data — should be defined as simply and flexibly as possible. Validation constraints on bibliographic data should be expressed, separately from the underlying vocabularies,in application profiles, whether in the manner of Dublin Core profiles (Coyle and Baker, 2009), in BIBFRAME Profiles (Library of Congress), or in one of several emerging technologies currently under development for RDF validation. RDF validation – the process of validating open-world Semantic Web data for conformance with closed-world constraints – is the focus of a new W3C Working Group on RDF Data Shapes.

The paper concludes by conceptualizing WEMI-like entities not as classes, but as variously shaped data graphs. These graphs may be expressed using flexible RDF vocabularies, yet subject to validation in specific application contexts in accordance with precise, closed-world constraints. It is an approach that aims at minimizing conceptual complexity while maximizing practical flexibility. Adopting an explicitly graph-based interpretation of FRBR-like entities does not so much answer the question of what WEMI-like entities "really" are, in reality, as render the answer to that question relatively insignificant.

How Semantic Web differs from traditional data processing

RDF [1] provides the grammar for a language of data. The language is used to assert "facts" (or "claims") about the world in the form of three-part, sentence-like statements called "triples." As in natural languages, where utterances are meaningful only if they follow a sentence grammar, RDF statements follow a simple and consistent three-part structure of subject, predicate, and object, as in the pseudo-triple "BookA translatedBy PersonB."

RDF statements use Uniform Resource Identifiers (URIs), essentially Web addresses used as identifiers, as names for things. In the example above, URIs would name the specific subject (BookA) and object (PersonB). A third URI would identify the predicate translatedBy. Full descriptions of BookA and PersonB could be built up by adding statements having as subject the URIs of BookA and PersonB (e.g., "BookA hasTitle 'The Blind Watchmaker'" and "PersonB hasName 'Richard Dawkins'). Analogously to paragraphs, RDF statements are aggregated into RDF graphs.

RDF statements make use of RDF properties (URIs for the predicates of triples, such as translatedBy) and RDF classes (URIs for types of things, such as Person). Properties and classes are defined in RDF vocabularies. RDF vocabularies defined with many specific axioms are sometimes called ontologies. To avoid confusion, this paper refers to OWL [2] restrictions not as constraints but as axioms. As noted above, this paper uses "RDF vocabulary" and "ontology" interchangeably.

RDF vocabularies are typically published in structured, machine-readable schemas on the Web. Best practice (which is not always followed) dictates that when clicked on in a Web browser, the URIs of properties and classes should resolve to RDF schemas or Web pages holding official definitions of the properties and classes. In that sense, the Web provides the language of data with its dictionary.

RDF vocabularies are published on the Web in order to encourage their re-use. The Dublin Core vocabulary, for example, provides URIs for commonly used properties such as Title and Date, making it easier for consumers of RDF data to understand its meaning and saving countless information providers the trouble of inventing their own synonyms. The ethos of the Linked Data movement encourages such re-use.

RDF is a language designed by humans for processing by machines. The RDF language — the grammar together with available RDF vocabularies — does not itself solve the difficulties of human communication around data and semantics. However, RDF does provide a means for expressing most of the types of information one commonly thinks of as "data." The specific strength of RDF derives from how the triple structure supports the process of connecting dots — of creating "knowledge" — by providing a linguistic basis for expressing and linking data.

RDF was designed in accordance with several principles of relevance to the analysis in this paper:

- **"Anyone can say anything about anything."** The Semantic Web languages were designed to accommodate multiple sources of information reflecting multiple points of view. The grammar of

RDF triples was designed to help merge data from multiple sources by leveraging shared URIs to align layers of information in a unified whole. As it is put in the RDF model specification of 2002: “RDF cannot prevent anyone from making nonsensical or inconsistent assertions, and applications that build upon RDF must find ways to deal with conflicting sources of information. (This is where RDF departs from the XML approach to data representation, which is generally quite prescriptive and aims to present an application with information that is well-formed and complete for the application’s needs.)” The specific strength of the Semantic Web approach, in other words, lies in how its data structure accommodates multiple sources and viewpoints.

- **Open World Assumption (OWA).** As a matter of principle, the information available at any given time may be incomplete. This is about more than just assuming that important servers might temporarily be offline. It is about recognizing that new information might be discovered or made available, or that information reflecting new points of view might be added to the mix. For example, new planets might be discovered, or existing planets might be reclassified (as happened to Pluto). According to the Closed World Assumption, in contrast, the information on hand defines the boundaries of what is known. The classic example of a closed-world system is a personnel database in which the lack of a value for, say, “supervisees” means that for the purposes of the database, the person in question has no supervisees. Where closed-world systems are appropriate for information environments designed to capture all known facts, systems based on the OWA are optimized for environments in which knowledge or scholarly opinion is a moving target and can be expected to evolve, change, or even to contradict itself.

- **Non-Unique Naming Assumption (NUNA).** As a matter of principle, things described in RDF data can have more than one name. Because URIs are used in RDF as names, anything may be identified by more than one URI. For example, the New York Times and Wikipedia may maintain their own URIs to identify the conductor “Lorin Maazel.” To put it another way, things are not assumed to be different because they have different names, as would be the case with a personnel database, where people with different names would be considered separate individuals.

RDF data derives its meaning from the RDF properties and classes it uses. RDF vocabularies define properties and classes in relation to other properties and classes. A property meaning “copyright date” (more specific) may be declared as a sub-property of a property meaning “date” (more general). A class “Primate” may be declared as a sub-class of a class “Mammal”. Such relationships allow a consuming application to infer additional information from the facts asserted in an RDF dataset. For example, if the data says that Lucy is a Primate, a Semantic Web application (called a “reasoner”) will logically conclude that Lucy is a Mammal and may augment the local copy of the dataset with a triple to that effect.

Of particular relevance to the analysis in this paper is the notion that the meaning of properties can be specified with formal domains and ranges. The range of a property defines the type of a value associated with a property, such that the object of a triple using the property may be inferred to be a member of a specific class. The domain of a property serves to indicate the class, or classes, to which the subject of the property belongs. For example, a property meaning “plays game” might have a class Athlete as its domain, and a class Sport as its range, such that whenever the “plays game” property is encountered in an RDF triple, a reasoner will infer that the subject of the triple is an athlete (i.e., an instance of the class Athlete) and the object of the triple is a sport. Given information about formally defined domains and ranges, additional information about the type of thing described by a set of RDF triples can be inferred on the basis of just one of those triples. Note the more general point that RDF instance data is interpreted with reference to the RDF vocabularies used.

How properties are associated with classes in RDF, with RDF domains and ranges, is radically different from how properties are associated with classes in systems based on the Closed-World Assumption, such as relational databases, XML repositories, and object-oriented systems. Closed-world systems prescribe templates to which data must conform, with required structures and integrity constraints that can be used to flag errors in the data. In such a prescriptive system, for example, saying that "plays game" is a property of the class Athlete typically implies that instances of the Athlete class can (or must) have a value for "plays game." Not declaring "plays game" as a property associated with the class Athlete typically means that members of that class may not be described with the property.

In RDF, domains and ranges are not used to bind a set of properties to a class. Indeed, an OWL property cannot be limited for use only with members of a specific class. Rather, RDF domains and ranges enable inferences, such that if a "plays game" property is used to describe a member of the class Person, a reasoner will simply state that the person is also a member of the class Athlete. Where a closed-world system will signal an error if a property is used in a context not intended by its maintainers, an RDF reasoner will simply infer additional information. An RDF reasoner may flag logical inconsistencies, for example in a triple that says "X isDifferentFrom X," however it will not validate that data against a template of integrity constraints.

It is therefore never correct to say that an RDF property is defined "for a class." In accordance with the AAA principle, properties are, on principle, independent of classes. The property for "plays game" is not limited by its formal domain to use only with people declared as members of the class Athlete; rather, the use of "plays game" to describe a person will trigger by inference the additional information that the person is an athlete. This inference then coexists with any other inferences that are supplied by RDF statements in the context of the current application. The flexibility of RDF in this regard simply acknowledges that in reality, things can belong to any number of classes (e.g. "librarian," "gardener," "parent") and that in the open context of the Web, people or communities may legitimately view things from multiple points of view. The paper will return to these fundamental differences of approach in the analysis of examples using RDF vocabularies for bibliographic description.

Multi-entity bibliographic models in RDF

This paper introduces three RDF vocabularies for multi-entity bibliographic models: FRBRer, RDA, and BIBFRAME and presents annotated examples of instance data that use the vocabularies. Each example is interpreted in plain English in accordance with how the underlying RDF vocabularies are defined.

Inasmuch as FRBRer, RDA, and BIBFRAME all define bibliographic entities as RDF or OWL classes, the examples are designed to illustrate the actual effect of the use of these classes in descriptions. Recall that the properties and classes RDF vocabularies are designed to be reusable in multiple contexts, as described above. Indeed, reuse of vocabularies is encouraged by the Semantic Web community as a means of achieving interoperability across multiple information sources. Recall, too, that in RDF, properties cannot be limited to use with a specific class or classes.

Where some of the example scenarios below conform to what the authors interpreted to be the intention of the vocabulary developers ("orthodox" examples), others were designed to test the results of reasoning over what would likely be considered non-conformant data ("unorthodox" examples). For such simple examples, reasoning was done "by hand" on the basis of vocabulary definitions, then checked against the results yielded by loading the instance data and related vocabularies into Stanford's Protégé Desktop software [8] and TopQuadrant's TopBraid Composer [9] and using their built-in reasoners. Both are popular ontology editors that support reasoning. Examples were first tested using the Pellet reasoner in Protégé, with results then confirmed using the SwiftOWLIM inferencing engine in TopBraid. Instance and ontology files for each of the examples is available at website of Kent State University [12] along with additional information about technical changes made to the ontologies required to successfully test the examples (such as modifications to allow titles and dates to be entered as literals).

Annotated examples of instance data

FRBR in RDF

Functional Requirements for Bibliographic Records (FRBR) (IFLA, 2009), a model of the bibliographic universe created in the 1990s as a basis for next-generation cataloging standards, redefined the hitherto unitary bibliographic description as a set of entities and relationships. Some of those entities and relationships reflected authority control practices in library cataloging, such as personal names and subjects. The primary bibliographic entities of Work, Expression, Manifestation and Item (WEMI) had a conceptual basis in library cataloging but had not before been treated as entirely separate "things."

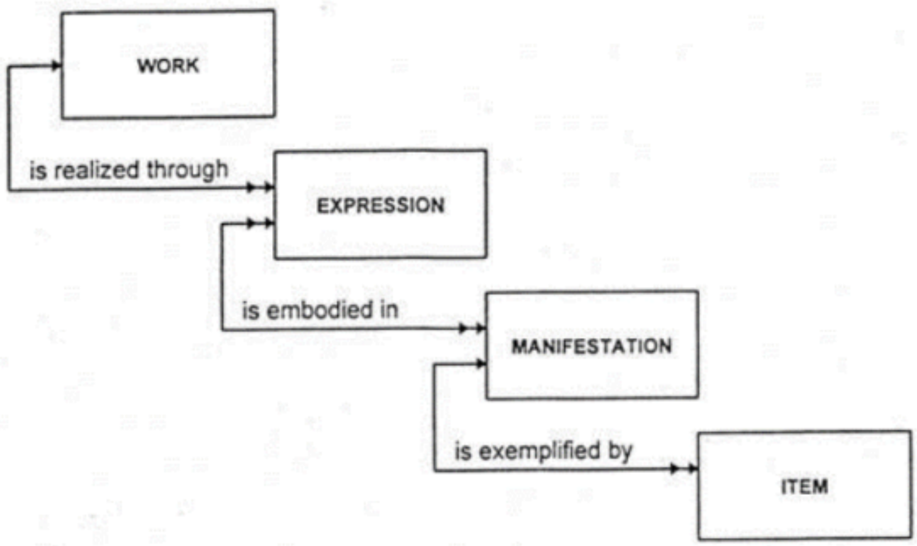

Figure 1. Bibliographic entities in the FRBR model (FRBR, 2008)

The FRBR Study Group was sponsored by IFLA in its cataloging section. The FRBR Study Group presented the model in 1998 in a document of 142 pages, including appendices, though popular understanding of the FRBR model has been shaped less by this text than by its diagrams, and especially of the diagram for the FRBR Group 1 entities: Work, Expression, Manifestation, and Item, or WEMI (see Figure 1 above). Though the full document expresses insights that cannot be reduced to simple formalisms, the FRBR diagrams follow a specific modeling methodology, entity-relation (ER) modeling, as mandated in the 1992 Terms of Reference for the FRBR consultants (Madison, 2005).

ER modeling defines business processes in terms of components (entities), "the key objects… of interest to users of information in a particular domain." (IFLA, 2009) Entities are assigned properties (attributes) and linked among themselves in relationships of dependency, the cardinality of which may be specified in the model. These components provide constraints against which actual data can event ually be validated for conformance.

An ER design process begins by defining a high-level conceptual model and typically culminates in the design of an actual database application. The FRBR specification of 1998 was presented as a high-level conceptual model for the "bibliographic universe." The FRBR Study Group did not specify whether the conceptual model should lead to a record format, to more specific logical database designs, or to something else altogether. Without that follow-on step, FRBR remained a concept without an implementation.

The FRBR conceptual model, as defined in 1998, has since 2008 been expressed by the IFLA Study Group on FRBR in an OWL ontology, "FRBRer" (FRBR entity-relation)[3]. [i]The WEMI entities are interpreted in FRBRer as OWL classes, and the attributes of those entities as properties, with each property having a single domain (that is, each property is associated with one and only one class). FRBRer directly translates constraints in the 1998 ER model into superficially equivalent OWL axioms, such as the one-to-many relationship between works and expressions. In FRBRer, all of the FRBR classes are declared to be disjoint with each other with the exception of the Topic class, which is not disjoint with the bibliographic (FRBR group 1) and agent (FRBR group 2) classes because all FRBR entities can also be subjects of a resource. OWL cardinality axioms are used to define relationships between classes that are superficially analogous to those indicated in the entity-relation analysis. As discussed below, in fact there are significant differences between ER and OWL concepts that make this superficial translation problematic.

Example 1. "Orthodox" data using FRBRer properties

The following data reflects the authors' interpretation of how FRBRer/OWL is intended to be used.

```
ex:ResourceA
    frbrer:P2001 ex:ResourceB ;                                    # isRealizedThrough
    frbrer:P2009 <http://id.loc.gov/authorities/names/n79021164> ;
                                                                   # isCreatedByPerson
    frbrer:P3001 "The adventures of Tom Sawyer" .                  # hasTitleOfTheWork


ex:ResourceB
    frbrer:P2002 ex:ResourceA ;                        # isRealizationOf
    frbrer:P2003 ex:ResourceC ;                        # isEmbodiedIn
    frbrer:P3011 <http://id.loc.gov/vocabulary/iso639-       # hasLanguageOfExpression
2/eng>


ex:ResourceC
    frbrer:P2004 ex:ResourceB ;                    # isEmbodimentOf
    frbrer:P3020 "The adventures of Tom Sawyer" ;  # hasTitleOfTheManifestation
    frbrer:P3055 "1996" .                          # hasDateOfPublicationOrDistribution
```

The data describes three FRBRer entities — a Work, an Expression, and a Manifestation — instantiated as ResourceA, ResourceB, and ResourceC respectively. Each of the properties used in the data is defined with a formal domain that leads a reasoner to infer additional information about the resource described. For example:

- The property isRealizedThrough (frbrer:P2001) has a domain of FRBRer Work (frbrer:C1001), so a reasoner infers that ResourceA is an instance of the class FRBRer Work.
- The property hasLanguageOfExpression (frbrer:P3011) has a domain of FRBRer Expression (frbrer:C1002), so a reasoner infers that ResourceB is an instance of the class FRBRer Expression.
- The property isEmbodimentOf (frbrer:P2004) has a domain of FRBRer Manifestation (frbrer:C1003), so a reasoner infers that ResourceC is an instance of the class FRBRer Manifestation.

Example 2. "Unorthodox" data with a FRBRer Expression related to more than one FRBRer Work

The FRBR document states that "An expression … is the realization of one and only one work." [FRBR p.13] In the FRBRer vocabulary, the property isRealizationOf is defined using the OWL property mincardinality with a value of "1". This example shows what a reasoner infers when the resource described (ResourceB) is stated to be a "realization of" *two* works (ResourceA and ResourceA1). (Note that the reasoner infers that ResourceB is a FRBRer Expression from the domain of property hasLanguageOfExpression.)

```
ex:ResourceB                                                        # FRBRer Expression
    frbrer:P2002 ex:ResourceA ;                                     # isRealizationOf
    frbrer:P2002 ex:ResourceA1 ;                                    # isRealizationOf
    frbrer:P3011 <http://id.loc.gov/vocabulary/iso639-2/eng> .
                                                            # hasLanguageOfExpression
```

Because of the Non-Unique Naming Assumption, and because an instance of FRBRer Expression is, by definition, a realization of just one FRBRer Work, a reasoner infers that ResourceA and Resource A1 are the same (i.e., that the two URIs are two different names for the same resource).

Such semantics can be used to identify multiple URIs for a single work. However, it seems unwise to hard-wire such a conclusion into the FRBRer vocabulary itself because a reasoner will not know if the URIs actually do represent the same work. The reasoner will reach the same conclusion regardless of whether it is a case of the same work being assigned different URIs, perhaps by different cataloging bodies, or of URIs that actually do identify two different works (e.g., "Hamlet" and "Der Zauberberg").

This example illustrates that the effect of OWL cardinality axioms is to support the inference of additional information. OWL cardinality axioms do not express integrity constraints on data in the manner of closed-world systems. If the property isRealizationOf is used with two different values, as in the example above, a standard reasoner will not flag the multiple values as an error.

Example 3. "Unorthodox" data with clashes between disjoint FRBRer domain classes

There is nothing to prevent someone from creating a set of FRBRer statements on the open Web such as the following:

```
ex:ResourceA
    frbrer:P2001 ex:ResourceB ;                                 # isRealizationOf
    frbrer:P2009 <http://id.loc.gov/authorities/names/n79021164> ;
                                                                  # isCreatedByPerson
    frbrer:P3020 "The adventures of Tom Sawyer" .                 # hasTitleOfTheManifestation
```

The first statement ("ResourceA isRealizationOf ResourceB") supports the inference that ResourceA is an instance of the class FRBRer Work (the domain of isRealizationOf). However the third statement ("ResourceA hasTitleOfTheManifestation 'The adventures of Tom Sawyer'") supports the inference that ResourceA is an instance of the class FRBRer Manifestation (the domain of hasTitleOfTheManifestation).

In FRBRer, the classes Work and Manifestation are formally declared to be disjoint, which means that by definition, no resource in the FRBRer universe is an instance of both FRBRer Work and FRBRer Manifestation. A reasoner will detect that this data violates the axioms of the FRBRer vocabulary. In OWL terms, the data is "inconsistent" and this may preclude operations over RDF data in the open data environment. In the open Web, where anyone can say anything about anything using any available RDF vocabulary, vocabularies defined according to highly specific logical rules, such as disjointness, have the effect of making any data that uses those vocabularies brittle, i.e., difficult to process meaningfully unless it is logically consistent with the orthodoxy, as we illustrate in Example 8.

RDA in RDF

Since the early 2000s, the main activity around FRBR has shifted from the IFLA activities around FRBR to the group preparing a revision of Anglo-American Cataloging Rules (AACR) under the name Resource Description and Access (RDA). The rules of RDA both refer to and are organized according to FRBR entities. As the cataloging rules primarily address description of resources, RDA mainly provides rules for describing the WEMI entities. The RDA documentation extracted an "element set" from the rules as the basis for an RDA-based bibliographic record.

In the absence of a timely push from within the IFLA or RDA communities to express the FRBR model in RDF before 2007, a family of RDF vocabularies emerged in the wider community that were loosely based on, or inspired by, FRBR: notably Indecs (Rust, 2000), FRBR Core (2005) [4], FRBRoo (2006) (International Council for Museums), and FaBiO (2010) (Shotton), as described by Coyle (2014). Work on representing the IFLA and RDA variants of FRBR in RDF began as a result of a meeting at the British Library [5] in May 2007 between representatives of RDA, DCMI, and the Semantic Web communities that agreed on a resolution to express RDA elements as an RDF vocabulary.

The RDF vocabulary for RDA was developed on a website, Open Metadata Registry [13], from the list of RDA elements provided by the RDA group. The first version of the vocabulary did not include bindings to FRBR entities as classes. Because the RDA development group felt strongly that each element must be defined for one and only one FRBR entity, a second vocabulary was developed that followed this precept. This resulted in two variant RDF vocabularies: one "unconstrained" and one "constrained" to FRBR classes. At the time there was no IFLA-sanctioned version of FRBR, so the nine FRBR entities were added to the RDA vocabulary namespace as separate classes, though in contrast to the FRBRer vocabulary, they were not defined as disjoint. Each RDA constrained property has one of these classes defined as its domain. This restriction required the repetition of some properties that had originally been associated in RDA with more than one FRBR entity. For example, "note on extent" became two properties: "note on extent of manifestation" and "note on extent of item."

RDA Constrained and RDA Unconstrained are mapped to each other. Each of the constrained properties has an unconstrained property as its super-property, and all but a few unconstrained properties have one or more constrained properties as sub-properties. There are a total of about 960 constrained properties and slightly over 700 unconstrained. Not included in the unconstrained properties are the relationships that would link FRBR WEMI entities among themselves, such as an Expression to a Work.

A version of RDA in RDF under the namespace URI http://rdaregistry.info was announced in 2014, with separate namespaces for the FRBR classes, each of the WEMI entities, the Agent class, and the unconstrained properties. [6]

Example 4. Orthodox data using official ("constrained") RDA properties

The following example is the same as Example 1, only using corresponding FRBR-constrained properties in RDA.

```
ex:ResourceA
    rdaw:P10078 ex:ResourceB ;                                              # expressionOfWork
    rdaw:P10065 <http://id.loc.gov/authorities/names/n79021164> ;           # creator
    rdaw:P10088 "The adventures of Tom Sawyer" .                            # titleOfTheWork


ex:ResourceB
    rdae:P20231 ex:ResourceA ;                                    # workExpressed
    rdae:P20059 ex:ResourceC ;                                    # manifestationOfExpression
    rdae:P20006 <http://id.loc.gov/vocabulary/iso639-2/eng>       # languageOfExpression
.


ex:ResourceC
    rdam:P30139 ex:ResourceB ;                                   # expressionManifested
    rdam:P30156 "The adventures of Tom Sawyer" ;                 # titleProper
    rdam:P30011 "1996" .                                         # dateOfPublication
```

As in Example 1, this data describes ResourceA, ResourceB, and ResourceC — instances of the RDA classes Work, Expression, and Manifestation, respectively. Each of the properties is associated with just one WEMI class as its domain.

Example 5. "Unorthodox" (but formally consistent) RDA data with a resource belonging to multiple WEMI classes

On the open Web it is always possible, indeed likely, that RDF vocabularies will be used in ways not foreseen or intended by their creators, whether by mistake or by design. This example shows a use of RDA FRBR-constrained properties with a single defined resource using a mix of WEM elements.

```
ex:ResourceA
    rdaw:P10065 <http://id.loc.gov/authorities/names/n79021164> ;     # creator
    rdaw:P10088 "The adventures of Tom Sawyer" ;                      # titleOfTheWork
    rdae:P20006 <http://id.loc.gov/vocabulary/iso639-2/eng> ;         # languageOfExpression
    rdam:P30156 "The adventures of Tom Sawyer" ;                      # titleProper
    rdam:P30011 "1996" .                                              # dateOfPublication
```

Interpreting the data with a reasoner yields the following:

- The property creator (rdaw:P10065) has a domain of Work, so a reasoner infers that ResourceA is an instance of the class Work.
- The property languageOfExpression (rdae:P20006) has a domain of Expression, so a reasoner infers that ResourceA is an instance of the class Expression.
- The property dateOfPublication (rdam:P30011) has a domain of Manifestation, so a reasoner infers that ResourceA is an instance of the class Manifestation.

ResourceA, in other words, is declared to be an instance of Work, Expression, and Manifestation all at the same time. In contrast to FRBRer, however, the RDA classes Work, Expression, and Manifestation are not disjoint. Nothing in the RDF vocabularies for RDA defines as a logical contradiction for a resource to be an instance of all three classes at the same time. While unorthodox from the standpoint of RDA rules, the data is not flagged as logically inconsistent by a reasoner. In order to detect this as a "mistake," an application would need to rely not on a reasoner but on metadata creation tools or validators using closed-world integrity constraints.

BIBFRAME in RDF

The Bibliographic Framework project (BIBFRAME) at Library of Congress is developing a bibliographic data format based natively on RDF. [7] The project has partnered with some libraries and archives to test the vocabulary and related software as it is developed. Much of the testing being done begins with existing bibliographic records, often in MARC format, and converts these to BIBFRAME. BIBFRAME in general, and the BIBFRAME vocabulary in particular, are "work in progress" and could change considerably in the future. Currently BIBFRAME has about one hundred elements, many less than RDA.

BIBFRAME defines a bibliographic model with two entities: Work and Instance. BIBFRAME Work is approximately equivalent to FRBR Work and Expression in combination. BIBFRAME Instance corresponds roughly to FRBR Manifestation. BIBFRAME also defines an authority entity, which is similar to the library concept of an authority-controlled entry that is maintained outside of the bibliographic record, and an annotation entity, which is described as "decorating" other BIBFRAME resources with additional information. The Work entity represents the intellectual content, and the Instance entity represents the physical embodiment. There does not appear to be a precedent for this two-entity model in the cataloging community.

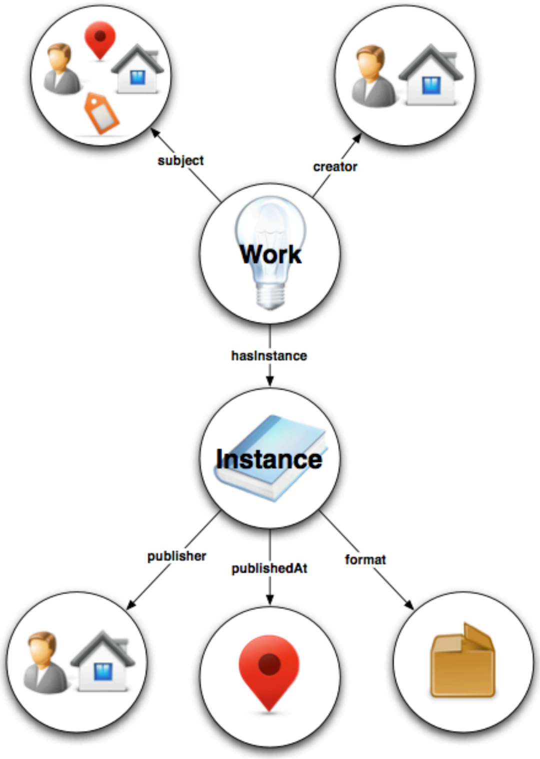

Figure 2. BIBFRAME entity model (Library of Congress, 2014)

The BIBFRAME Project has been criticized in the cataloging community for not using FRBR's four-entity WEMI model. BIBFRAME "profiles" have been announced as a solution. (Library of Congress, 2014) A profile could present a view of BIBFRAME with the same conceptual divisions as WEMI. Profiles can present an alternate view of data by combining the same BIBFRAME properties in a different set of graphs from those defined by the BIBFRAME documentation. Profiles define constraints that are orthogonal to the underlying RDF or OWL vocabularies. Profiles can add constraints that are not in the original vocabulary, however they are additive and cannot override or remove existing meaning.

Example 6. "Orthodox" BIBFRAME data

This example shows how the BIBFRAME vocabulary is used in the examples that are provided by the implementation partners and as output from the BIBFRAME tools [11]. BIBFRAME properties are not associated with FRBR classes as RDF domains, but with the BIBFRAME classes Work and Instance. Those classes, however, are not defined as disjoint. The correct use of properties in describing Works or Instances is documented and enforced outside of the BIBFRAME vocabulary itself, in the BIBFRAME software. Other software making use of this vocabulary may allow different choices to be made.

```
ex:ResourceA
    bf:hasInstance ex:ResourceB ;
    bf:creator <http://id.loc.gov/authorities/names/n79021164> ;
    bf:title "The adventures of Tom Sawyer" ;
    bf:language <http://id.loc.gov/vocabulary/iso639-2/eng> .
ex:ResourceB
    bf:instanceOf ex:ResourceA ;
    bf:instanceTitle "The adventures of Tom Sawyer" ;
    bf:providerDate "1996" .
```

ResourceA and ResourceB are interpreted to be instances of classes Work and Instance, respectively, because the properties used to describe them have been defined with those domains. In fact, this example could consist of the two resources and one property each, and the inferred classes of the two resources would be the same.

Example 7. "Unorthodox" BIBFRAME data, with a resource that is an instance of two different classes

This example, using BIBFRAME vocabulary, mixes properties intended for describing Works and Instances into a single description of ResourceA.

```
ex:ResourceA
    bf:language <http://id.loc.gov/authorities/names/n79021164> ;
    bf:title "The adventures of Tom Sawyer" ;
    bf:language <http://id.loc.gov/vocabulary/iso639-2/eng> ;
    bf:instanceTitle "The adventures of Tom Sawyer" ;
    bf:providerDate "1996" .
```

The example is unorthodox because in BIBFRAME, instances of Work and Instance are clearly intended to be separate. However, a reasoner will not complain that this data is inconsistent because it does not violate the axioms of the vocabulary. In a well-designed BIBFRAME production environment, "incorrect" data would not be produced in the first place — not because it is disallowed by the vocabulary itself, but because the data input tools would not support its creation. Existing BIBFRAME data would be validated against integrity constraints, perhaps on the basis of an application profile. However, outside of the BIBFRAME application environment, there is no way to prevent the existence of descriptions that use BIBFRAME properties but do not follow the proscribed BIBFRAME data model and structure.

Example 8. Linking FRBRer and BIBFRAME data

This example illustrates the case of linking instances defined using properties from the FRBRer and BIBFRAME vocabularies, where the instances of FRBRer Work and Expression both have a "same as" relationship to an instance of BIBFRAME Work, since the FRBR Work and Expression concepts correspond to the single BIBFRAME entity of Work.

```
ex:ResourceA                              # FRBRer/Work
    frbrer:P2001 ex:ResourceB ;
    frbrer:P2009 <http://id.loc.gov/authorities/names/n79021164> ;
    frbrer:P3001 "The adventures of Tom Sawyer" .

ex:ResourceA owl:sameAs ex:ResourceX .
ex:ResourceB                              # FRBRer/Expression
    frbrer:P2002 ex:ResourceA ;
    frbrer:P2003 ex:ResourceC ;
    frbrer:P3011 <http://id.loc.gov/vocabulary/iso639-2/eng> .

ex:ResourceB owl:sameAs ex:ResourceX .

ex:ResourceC                              # FRBRer Manifestation
    frbrer:P2004 ex:ResourceB ;
    frbrer:P3020 "The adventures of Tom Sawyer" ;
    frbrer:P3055 "1996" .

ex:ResourceX                              # BF/Work
    bf:hasAuthority <http://id.loc.gov/authorities/names/n79021164> ;
    bf:workTitle "The adventures of Tom Sawyer" ;
    bf:language <http://id.loc.gov/vocabulary/iso639-2/eng> .

ex:ResourceY                              # BF/Instance
    bf:instanceOf ex:ResourceX ;
    bf:instanceTitle "The adventures of Tom Sawyer" ;
    bf:providerDate "1996" .
```

The `owl:sameAs` statements on ResourceA and ResourceB state that the instance of BIBFRAME Work, here ResourceX, is equivalent to both the instances of FRBRer Work and FRBRer Expression. While it appears to be logical to combine FRBR and BIBFRAME by treating both the FRBRer Work and Expression as being the same as the BIBFRAME Work, this example fails because FRBRer Work and Expression are defined as disjoint, and therefore cannot be the same as any single resource.

From a BIBFRAME perspective, it is not inconsistent to link a work to more than one other resource, as in:

```
ex:ResourceX # BF/Work
    bf:hasAuthority <http://id.loc.gov/authorities/names/n79021164> ;
    bf:workTitle "The adventures of Tom Sawyer" ;
    bf:language <http://id.loc.gov/vocabulary/iso639-2/eng> .
ex:ResourceX owl:sameAs ex:ResourceA .
ex:ResourceX owl:sameAs ex:ResourceB .
```

Although the BIBFRAME classes Work and Instance are not defined as disjoint, the reasoner still finds inconsistency in this example because of the disjointness between FRBRer Work and Expression. This example shows that bibliographic data which does not use four distinct entities for bibliographic description cannot interact with data described using the FRBRer vocabulary.

Discussion

Bibliographic Entities as Classes

The RDF vocabularies for multi-entity bibliographic models examined here conceptualize bibliographic things as sets of resources belonging to different RDF classes. The FRBRer and RDA Constrained vocabularies define four such classes: Work, Expression, Manifestation, and Item. The simpler but roughly analogous BIBFRAME vocabulary defines just two classes: Work and Instance. This section discusses the role and meaning of these RDF classes in light of the examples above.

RDF classes do not define properties. Associating a property with a single bibliographic class as its RDF domain as in the vocabularies using WEMI, may appear to define a set of properties that are valid for describing instances of the domain class. In RDF, however, the association of properties with domain classes is not restrictive. RDF properties are independent of classes and may in principle be used to describe instances of any class. The inference-based semantics of RDF is informative in nature, adding meaning and context without restricting use. An RDF domain declaration simply allows a reasoner to infer that the resource described by a property is also a member of the domain class. In RDF, as in real life, things may be instances of more than one class; consider a "father" who is also an "opera lover" and an "employee."

RDF classes do not support the validation of data structures. In relational databases, XML environments, object-oriented systems, and record repositories, classes are defined with a fixed set of properties against which data can be validated for conformance. This "closed-world" approach is optimized for situations in which the universe of known information is relatively tightly controlled, as in a database of employee data. In RDF, asserting membership in a class *augment*s the information available about the thing being described. The "open-world" approach of RDF is optimized for merging data from multiple sources in a constantly changing universe where the information available at a given moment may be fragmentary or incomplete. Inferencing can fill in the gaps and explicitly state the mappings needed to query a heterogeneous corpus of data. As will be discussed below, emerging approaches to RDF validation aim at evaluating RDF data according to closed-world constraints by expressing those constraints outside of RDF vocabularies themselves.

OWL axioms are not data validation constraints. Data integrity constraints defined for closed-world systems may mean something quite different from superficially equivalent Semantic Web axioms. As shown in Example 2, the requirement that an Expression express exactly one Work does not translate into the seemingly equivalent cardinality restriction in OWL. If RDF data says that an instance of the class Expression has an "expresses" relationship to two instances of the class Work, an OWL reasoner will follow the Non-Unique Naming principle to conclude that the two instances of Work are the same.

Declaring classes to be disjoint carries heavy consequences. If classes such as Work, Expression, Manifestation, or Item are not declared to be disjoint, it is relatively harmless to claim that something is an instance of such a class, although class declarations should be motivated by specific descriptive needs. On the other hand, declaring classes to be disjoint means that it is logically inconsistent to assert that something is simultaneously a member of any two such classes. This may seem acceptable in controlled situations where clarity of class membership can be assumed, enforced, or imposed. But disjointness creates problems in situations where class membership can be contested, misunderstood, or subject to multiple interpretations, or when data describing instances of the classes must be integrated with data describing instances of entirely different classes. Heavily specified vocabularies carry costs to users downstream, both to consumers of RDF data created using the vocabulary and to creators who want to use a vocabulary in their own data, because deviations from the specification will be flagged by a reasoner as errors of logic. This can result in a silo'd data set that will not play well with other data sources on the open Web.

Bibliographic Entities Seen as Single Resource Descriptions

In practical terms, the ultimate purpose of multi-entity bibliographic models is to distribute the description of bibliographic resources over multiple data sources which can, in principle, be separately maintained, and which can be cited or re-used as needed. For example, it can be useful to retrieve the properties of a FRBR Work for use in a cataloging workflow.

Seen from this perspective, the entities of multi-entity bibliographic models can usefully be conceptualized as being part of a larger graph of bibliographic data that describes a single bibliographic resource. In their "revisioning" of cataloging theory, Murray and Tillett characterize cultural heritage description as yielding "groups of statements that occupy different levels of abstraction." (Murray and Tillett, 2011) Bibliographic description, in other words, may be less about separate entities than about subsets of the description that represent different views of the data. As the FRBR document of 1998 puts it: "Each of the entities defined for the model… serves as the focal point for a cluster of data." (FRBR, 2009) Murray and Tillett characterize WEMI entities as providing variable views of a complex entity that combines meaning, expression, and physicality. In their model, the bibliographic description using WEMI concepts consists of four sets of statements reflecting four complementary views of a resource. A Work graph, then, constitutes the description of a resource "viewed as a work." One might refer to a graph describing just such a bibliographic entity, identified with a URI, as a Single-Resource Description.

With this simple shift of perspective, some of the conceptual problems around multi-entity bibliographic models melt away. In the story of the blind men and the elephant, by analogy, men reach different conclusions about an elephant based on what they can feel. They "see" with their hands four different views of the elephant. In most versions of the story, they stop short of concluding that the elephant is actually four separate entities. The story illustrates that truth may lie in the sum of views of a thing. Separate WEMI descriptions could be combined into a Single-Resource Description. Alternatively, separate WEMI descriptions could be derived from a Single-Resource Description as views.

The FRBR document of 1998 embraces such diversity, at least in principle. Inasmuch as different cultures have different notions of Work with, presumably, different attributes, bibliographic communities may not be expected to share precise definitions for FRBR entities: "Because the notion of a work is abstract, it is difficult to define precise boundaries for the entity. The concept of what constitutes a work and where the line of demarcation lies between one work and another may in fact be viewed differently from one culture to another." (p.16) Catalogers would benefit from alternative definitions of Work better suited to, say, audio-visual materials or art works. (Baca and Clarke, 2007; Van Malssen, 2014)

Multiple "views" may be needed to serve multiple use cases. In today's open-world, Semantic Web context, interoperability no longer requires lock-step conformance with closed-world constraints as it did from the International Cataloging Principles of Paris, in 1961 (IFLA Cataloguing Section) through the MARC format in the 1960s [10], to IFLA's International Standard Bibliographic Description (IFLA, 2007). Inasmuch as RDF is optimized for merging data from multiple sources, reasonably coherent overlap will suffice. In this sense, the implicit goal of the FRBR effort in the 1990s — to create consensus around a single bibliographic model — is no longer required.

Validating RDF Data

What remains is a requirement for quality control: methods for ensuring that RDF data is created according to documentation and templates that can be used to produce consistent output, configure input and display interfaces, ensure the semantic coherence of ingested data, and drive search applications.

The need for methods to control the quality of RDF data is widely acknowledged and is currently being addressed on two fronts. As of August 2014, the World Wide Web Consortium (W3C) is in the process of convening a working group on validating RDF data under the name RDF Data Shapes. The group is looking at several existing approaches to RDF validation, including Top Quadrant's SPIN language (Knublauch, 2014), OWL ICV (Pérez-Urbina, 2014), and Shape Expressions (ShEx, 2014). Some participants in the discussion question whether another formal language, for expressing integrity constraints, is really needed and point out that RDF validation might more simply be implemented on the basis of SPARQL, the existing query language for RDF.

In parallel, a DCMI working group on RDF profiles is re-examining the Description Set Profile Constraint Language (Nilsson, 2008), a draft specification which aimed at providing a formal language for constructing templates for Description Sets (record formats based on RDF graphs). The DCMI work, in turn, has influenced the approach taken for BIBFRAME Profiles (Library of Congress, 2014).

Relating Multi- to Single-Entity Bibliographic Descriptions

In real life, Works, Expressions, and Manifestations are experienced through an encounter with a concrete bibliographic artifact, an Item. On the Web today, such artifacts are commonly described as if they were a single resource, such as a Book or Article, using flat resource models such as Simple Dublin Core.

In FRBR theory, WEMI entities are linked in a chain but do not constitute a meaningful whole. Although some of the canonical FRBR diagrams surround the four entities with a box (for "Group 1"), implying that WEMI entities are somehow encapsulated in a super-entity, the FRBR specification assigns no attributes to any of the three groups as a whole. The requirement for an entity representing the resource as a whole is richly evidenced in the literature: Golub et al (2014) enumerate relationships to a WEM entity seen as a whole (Golub et al 2014). Pisanki and Zumer (2010) present research subjects with cards showing Work attributes (e.g., author and title), Work plus Expression attributes (language), Work and Expression plus Manifestation attributes (publication information), and WEM attributes plus information unique to items; it is indeed hard to imagine a user making any sense of a card containing only, say, Expression attributes.

Partial bibliographic descriptions are considered as somehow adding up to a full description when taken as a whole. If the entities of multi-entity bibliographic models are related to a whole in theory, the nature of that relationship is unclear. Proposals for that relationship include “has aspect” (Dunsire, 2013), “contains” (Murray and Tillett, 2011), and “viewed as” (Murray and Tillett 2011b). Alternatively, one might indeed promulgate rules for conflating multi-level descriptions into Single-Resource Descriptions (for an example, see [14]).

Conclusion

This paper has focused on problems with expressing multi-entity bibliographic models as RDF classes. The analysis has emphasized that RDF vocabularies do not specify validatable data structures on the basis of class membership in support of data quality control. The paper proposes that the RDF vocabularies of bibliographic models be designed to maximize and encourage re-use on the open Web and that requirements for quality control be met by expressing data integrity constraints in one of the emerging technologies for RDF validation. In support of Murray and Tillett, this paper proposes that bibliographic entities be “re-visioned” as graphs. This shift of perspective suggests the following:

- As a rule of thumb, an RDF vocabulary is more reusable the fewer constraints it defines. Some of the RDF vocabularies reviewed here could be helpfully simplified by stripping them of definitions mistakenly thought to support validation. These are the axioms that look like closed-world integrity constraints, notably class membership, cardinality and disjointness, but that actually serve as the basis for misleading inferences. Attempts to use RDF vocabularies as control mechanisms will only serve to isolate those who adhere to them. A simplification of bibliographic vocabularies would make it easier to process library data outside of the library silo and facilitate the uptake of library-world vocabularies in the wider Semantic Web community.

- The integrity constraints needed for quality control of bibliographic data could be documented in Work-level, Expression-level, Manifestation-level, and Item-level profiles (for FRBR), or in Work-level and Instance-level profiles (for BIBFRAME). These profiles could follow either the BIBFRAME or Dublin Core approach to profiles or use another of the emerging approaches to RDF validation. Defining integrity constraints separately from RDF vocabularies results in flexible vocabularies with greater reusability. Different profiles or views can be specified for different contexts, e.g., stricter constraints for data creation and looser constraints for wider data sharing.

- With unconstrained vocabularies to work with, data creators will more easily be able to adapt the ideas behind multi-level models to previously unforeseen contexts, whether by distinguishing entities differently or by describing them with different attributes and constraints. While such adaptations may not be compatible with the canonical models used in libraries, it would make it easier to promote the vocabularies outside of the cultural heritage area.

If multi-level bibliographic models are to play well in today’s Linked Data cloud, a simple answer is needed the question of how multiple entities such as Work, Expression, Manifestation, Item, and Instance relate to more common (and “common-sense”) entities such as book, article, and musical recording. Metadata describing these resources exists online in databases containing millions of entries. If library data cannot interact with these resources, it will remain in its library-only silo. Through adopting simpler vocabularies with a more flexible approach, libraries can realize the vision, articulated at the birth of FRBR at the Stockholm seminar in 1990, to conserve precious resources by creating shorter, more targeted records. At the same time, they can participate in the grand information exchange that is the Web.

In today’s environment of recombinant metadata, it is no longer necessary to impose consensus in the form of lock-step conformance with closed-world constraints and complex, monolithic formats. Data can be customized for specific needs and contexts, with interoperability achieved to the extent that data graphs share overlapping vocabularies and constraints. In today's environment, interoperability outside of the library silo is not optional but a requirement.

---

*Endnote*s

[1] http://www.w3.org/standards/techs/rdf

[2] http://www.w3.org/standards/techs/owl

[3] http://iflastandards.info/ns/fr/frbr/frbrer.rdf Aug. 13, 2014

[4] http://vocab.org/frbr/core.html Aug. 13, 2014

[5] http://www.bl.uk/bibliographic/meeting.html

[6] https://github.com/RDARegistry/RDA-Vocabularies/zipball/master Aug. 13, 2014

[7] http://bibframe.org/vocab/ Aug. 13, 2014

[8] http://protege.stanford.edu

[9] http://www.topquadrant.com/tools/IDE-topbraid-composer-maestro-edition/

[10] http://loc.gov/marc/

[11] http://bibframe.org/tools/

[12] http://lod-lam.slis.kent.edu/wemi-rdf/

[13] http://rdvocab.info/

[14] http://lists.w3.org/Archives/Public/public-rdf-comments/2012Apr/0001.html

References

Baca, M., & Clarke, S. (2007). "FRBR and Works of Art , Architecture , and Material Culture." In A. Taylor (Ed.), Understanding FRBR: What It Is and How It Will Affect Our Retrieval Tools (pp. 227–242). Libraries Unlimited.

Coyle, K. and Baker, T. (2009) Guidelines for Dublin Core Application Profiles. (Accessed August 19, 2014)

Coyle, K. (2014) "FRBR, Twenty Years On." Cataloging and Classification Quarterly. Accepted for publication. DOI: 10.1080/01639374.2014.94344

Dunsire, G. (2013) "Resource and Work, Expression, Manifestation, Item." http://www.ifla.org/files/assets/cataloguing/isbd/OtherDocumentation/resource-wemi.pdf (Accessed March 26, 2014)

Golub, Koraljka, *et al.* (2014). "Terminology registries for knowledge organization systems." Journal of the American Society for Information Science and Technology (JASIS&T). DOI: 10.1002/asi.23090

IFLA Cataloguing Section. Statement of International Cataloging Principles. http://www.ifla.org/publications/statement-of-international-cataloguing-principles. (Accessed August 19, 2014)

IFLA Study Group on the Functional Requirements for Bibliographic Records. (2009). Functional Requirements for Bibliographic Records. Retrieved from http://archive.ifla.org/VII/s13/frbr/frbr_2008.pdf

IFLA. (2007). *International Standard Bibliographic Description (ISBD)*. München, K.G. Saur.

International Council for Museums. "The CIDOC Conceptual Reference Model: FRBRoo Introduction" (no date) http://cidoc-crm.org/frbr_inro.html

Joint Steering Committee for Development of RDA. RDA Registry. http://www.rdaregistry,info. (Accessed August 19, 2014)

Knublauch, H. (2014) SPIN – Modeling Vocabulary. http://www.topquadrant.com/spin/spin.html (Accessed August 19, 2014)

Library of Congress. Bibliographic Framework Initiative. http://loc.gov/bibframe/. (Accessed August 19, 2014)

Library of Congress. (2014) BIBFRAME Profiles: Introduction and Specification. http://loc.gov/bibframe/docs/bibframe-profiles.html (Accessed August 19, 2014)

Madison, O. (2005) "Terms of Reference for a Study of the Functional Requirements for Bibliographic Records." Le Boeuf, P. Functional Requirements for Bibliographic Records (FRBR): Hype or Cure-All? New York: Haworth Information Press. pp. 35-37

Murray, R.J. and Tillett, B.B. (2011), "Cataloging theory in search of graph theory and other ivory towers", Information Technology and Libraries, December, pp. 168-92. http://ejournals.bc.edu/ojs/index.php/ital/article/view/1868/1706 (Accessed August 19, 2014)

Murray, R.J. and Tillett, B.B. (2011b), Resource description diagram supplement to "Cataloging theory in search of graph theory and other ivory towers". http://arizona.openrepository.com/arizona/handle/10150/139769 (Accessed 17 August 2014).

Nilsson, M. (2008) Description Set Profiles: A constraint language for Dublin Core Application Profiles. http://dublincore.org/documents/dc-dsp/ (Accessed August 19, 2014)

Pérez-Urbina, H. *et al*. Validating RDF with OWL Integrity Constraints. http://docs.stardog.com/icv/icv-specification.html (Accessed August 19, 2014)

Pisanski, J., and Žumer, M. (2010). Mental models of the bibliographic universe. Part 1: mental models of descriptions. *Journal of Documentation*, *66*(5), 643–667. doi:10.1108/00220411011066772

Prud'hommeaux, E., and Seaborne, A. (2008) SPARQL Query Language for RDF. http://www.w3.org/TR/rdf-sparql-query/ (Accessed August 19, 2014)

Rust, G. and Bide, M. (2000). The <indecs> metadata framework; Principles , model and data dictionary. http://www.doi.org/topics/indecs/indecs_framework_2000.pdf (Accessed August 19, 2014)

ShEx (2014). http://www.w3.org/2001/sw/wiki/ShEx (Accessed August 19, 2014)

Shotton, D. and Peroni, S. FaBiO, the FRBR-aligned Bibliographic Ontology http://www.essepuntato.it/lode/http://purl.org/spar/fabio (Accessed August 19, 2014)

Van Malssen, K. (2014) BIBFRAME AV Modeling Study: Defining a Flexible Model for Description of Audiovisual Resources. http://loc.gov/bibframe/pdf/bibframe-avmodelingstudy-may15-2014.pdf (Accessed August 19, 2014)