

FRBR-LRM review

April 25, 2016

This review of FRBR-LRM was submitted to the PCC Steering Committee by the PCC Standing Committee on Standards (SCS).

SCS members: Diana Brooking, Chew Chiat Naun, Ryan Finnerty, Stephen Hearn, Ed Jones, Christee Pascale, Isabel Quintana, Kevin Randall, Dave Reser, Lori Robare, Manon Theroux, Jay Weitz, Jodi Williamschen.

SCS consultants: Steven Folsom, Adam Schiff, Tim Thompson.

This report also benefitted from additional comment by Robert Maxwell.

Alignment and objectives

A key objective of FRBR-LRM appears to be to align models among the library, archive, and museum communities. The fact that FRBRoo and PRESSoo are seen as extensions of the CIDOC CRM, and FRBR-LRM “operates at a greater level of generality than FRBRoo” suggests that FRBR-LRM is intended to support this extensibility. This objective appears to underlie some of FRBR-LRM’s more controversial prescriptions, notably its treatment of agency, and deserves fuller elaboration in the document.

FRBR-LRM calls itself “a high-level conceptual model” requiring expansion “or, possibly some omissions” in specific implementations. For example, LRM-R15, which defines the relationship of res to place, says that “in most implementations this relationship would be sub-typed”, and 4.2.2 says that “any external controlled vocabulary” may be used for category attributes.

Although the document discusses expansion in 2.2, FRBR-LRM would benefit from a more comprehensive and clearly defined framework explaining how it would interact with other models and vocabularies in specific implementations. For example, it would be useful to see how FRBR-LRM accommodates the family entity from RDA and its associated attributes such as type of family. It would also be helpful to understand how FRBR-LRM accommodates high-level entity types, such as events, which are not defined under the first, second, or third level of the FRBR-LRM entity hierarchy.

Editorially, SCS feels that the document would benefit from a glossary and more attention to choice of examples.

Fictitious and non-human identities

FRBR-LRM excludes fictitious entities from its definitions of places and persons, and denies that they are capable of agency.

We think this stipulation is unnecessarily restrictive. It leaves the model unable to account satisfactorily for responsibility attributed to fictitious entities. If such a restriction is to be made, it should be at the level of specific implementations and not be dictated by the model.

The same restriction complicates alignment with models such as foaf that have wide currency outside the library community. A foaf:Person representing Hester Prynne is unusable in a FRBR-LRM environment.

We do not share the concern that not restricting responsibility to real entities would lead to the propagation of misleading data. It would be simple enough to introduce a property or subclass to distinguish fictitious entities from real ones. It would also be possible to define attributed responsibility as a relationship between a real or fictitious entity and a work. This would have the advantage of dealing with cases of fictional or dubious attribution to real entities.

FRBR-LRM allows that fictitious entities are res, as is evidenced by the example of Miss Jane Marple under LRM-E1. Nomens are res, but it is insufficient for fictitious entities merely to consist in associations among nomens. Without a common referent, it is impossible to make sense of the fact (for example) that the same fictitious character can be known by different names in different languages or traditions. A nomen cannot carry other fictive attributes that would be needed to satisfy legitimate use cases: for example, a nomen cannot have a birthplace.

The implied restriction of agency to human beings is unhelpful. Non-human real beings are known to be responsible for the creation of and contributions to works and expressions. For instance, there is the well-known case of Congo the chimpanzee, who drew and painted pictures, or Terry the Cairn Terrier who performed in the 1939 film *The Wizard of Oz*. It was not Terry's trainer, Carl Spitz, who appeared in the film as Toto; it was Terry.

Nomens - other issues

In its discussion of schemes (LRM-A26), FRBR-LRM appears to conflate the concept (in the SKOS sense) represented by a nomen with its label. Traditional cataloguing theory makes a distinction between the scheme and the notation which FRBR-LRM does not adequately capture. The omission is a problem because it is necessary to be able to say that two labels apply to the same concept: for example, to say that one is preferred and the other a variant.

FRBR-LRM defines two nomens as being equivalent if they are appellations for the same res (page 43). "Context of use" (LRM-A28) is not taken into account in this definition of equivalency. This has the counterintuitive result that two pseudonyms are equivalent as long as they are used by the same author. This definition is in tension with the discussion of pseudonyms in 5.3, which attempts to analyze bibliographic identities in terms of nomens with distinct contexts of use.

Attributes and relationships

FRBR-LRM entities can have attributes or relationships. The document does not insist on a hard distinction between them, and it may be questioned whether that distinction is useful. For example, location is an attribute for items (LRM-A19), but other place associations are modelled as relationships (LRM-R15). Language is given as an attribute, but might be better modelled as a relationship to a language entity.

FRBR-LRM sets out cardinality for relationships, but does not define repeatability for attributes. It should be possible to say explicitly if an attribute is repeatable; the language of FRBR-LRM, that “it is only a matter of sound sense”, is not sufficiently rigorous.

In places it is not clear what the attribute is an attribute of. Language is given as a direct attribute of an agent (LRM-A23), but the definition speaks of an agent in relation to an expression. The notion of a representative expression raises similar issues: see below.

Transcription

LRM-A16 defines the attribute *manifestation statement* which addresses transcription of manifestation information such as publication statements. The relationship between these manifestation statements and the model’s treatment of distribution and production relationships (LRM-R8/9) would benefit from clarification. We also note that transcription is an issue that arises for many other kinds of data besides manifestation statements and warrants broader treatment.

Notes

Note is an attribute of res (LRM-A2) but the definition of a note as “textual material” seems unduly narrow. The W3C open annotation model acknowledges that many kinds of things can be annotations, and annotations can be applied to many things. The definition should allow for non-textual annotation.

Time-span

The Time-span entity is defined as a temporal event having a beginning, an end and a duration (no matter how short) but at the same time there are attributes for beginning and ending (LRM-A36 and A37) which presumably represent points in time. FRBR-LRM would do better to have separate entities for time interval and time point, as in ISO 8601.

Types of expressions

Some attributes are stated to be applicable only to specific types of expression, e.g. medium of performance (LRM-A11), but the types themselves are not declared within FRBR-LRM, only implied in examples. This leaves the status of the types and attributes within the model uncertain.

Representative expressions

SCS finds the notion of a representative expression problematical. Although it may be possible to demonstrate its usefulness for some communities, it is unclear what considerations would guide the choice of a representative expression, or how it would aid users.

LRM-A5 treats representativity as an attribute of an expression. It would be more satisfactory to treat it as a relationship between a work and an expression, or indeed as a relationship between a work and an expression for a given community. Since works are clusters of expressions grouped according to their perceived utility for a given community, the representative expression will have meaning only within that community.

The model defines representativity as a binary yes/no attribute, but it needs to be clarified whether an expression can be representative in some respects but not others. Alternatively, it may be preferable to record certain high-interest attributes like language and form as attributes related to the work's context or as canonical attributes at the work level rather than of any particular expression.

Aggregates and works

The treatment of aggregating works comes from FRBRoo, which seeks to model the processes that bring about individual works, expressions, and manifestations. However, it is not clear to SCS how far those processes need to be reflected in the FRBR-LRM model itself. What would be useful is clearer exposition of how the component works/expressions appearing in aggregates can be described in the context of a manifestation description, and particularly to how relationships are expressed between an entity and an aggregate resource as compared to an entity and a component resource, and likewise for attributes which could reference a component or the resource as a whole

Serials

The statement that “any serial work can be said to have only one expression and one manifestation” is implausible. It implies that a reproduction of a serial is a separate work from the original, even if the characteristics that define an expression are identical for both. Similarly, a serial may be issued on microform by two different providers; or a serial work may be issued in English and French versions. Holding that these are all cases of distinct works would seem tantamount to excluding serials from the model altogether; we note that serials are excluded from the examples even where they would have been appropriate, as in LRM-R21.

Integrating resources are a notable omission from the model. There is a confusion in the introductory paragraph of 5.6 between serials and journals. Only the latter exhibit the second level of aggregation.

The Wall Street Journal example in FRBR is more complex than paragraph 5 implies, since the midwestern “local edition” actually superseded, over time, an independent newspaper, the Chicago Journal of Commerce. It’s analogous to the example Patrick Le Boeuf used in his paper on PRESSoo at the 2014 IFLA conference in Lyon <http://library.ifla.org/838/1/086-leboeuf-en.pdf> (p. 4), which might actually serve as a better example of the complexity of serial relationships.

Disjoint classes

Ontologically FRBR-LRM, like the FRBR family as a whole, is over-committed with disjoint assertions. This makes it difficult to align with other bibliographic data found on the web. For example, schema.org does not follow strict disjoint WEMI definitions. What this means is that a schema:CreativeWork that has a related subject (Work attribute), a language (Expression attribute), an ISBN (Manifestation attribute) and a copy number (Item Attribute) is not relatable to any of the WEMI classes used in FRBR.

FRBR-LRM needs a better strategy to adhere to minimal ontological commitments. For example, the RDA Unconstrained Properties are an acknowledgement from the RDA developers that library models will need to be less stringent if we want to align our data with other data on the web. The latest versions of CIDOC-CRM and FRBRoo [1][2] have class scope notes/skos:definitions that suggest disjointness, but they do not actually have formal axioms asserting disjointness.

[1] http://www.cidoc-crm.org/docs/frbr_oo/frbr_docs/FRBRoo_V2.4.pdf

[2] http://www.cidoc-crm.org/rdfs/cidoc_crm_v6.2.1-draft-b-2015October.rdfs