This is the original English version of a paper presented at the conference «Современные информационные технологии и письменное наследие: от древних текстов к электронным библиотекам» [Modern Information Technologies and Written Heritage: From Ancient Texts to Electronic Libraries] (El'Manuscript-08), held August 25–30, 2008 in Kazan, Tatarstan, Russia. A version in Russian is also available.

Entailment of Entities and Implicature of Attributes in the FRBR Model[1]

Kevin S. Hawkins
University of Michigan, Ann Arbor, Michigan, USA

Introduction

The text encoding community has expended much energy debating the extent to which a data model for text can be formulated. While the use of computers in libraries goes back essentially as far as in the text encoding community, it is only more recently that the library community has attempted to formulate its own data model, one for the most common use of computers in libraries: library catalogs.

However, the most widely discussed model in the library community—that described in *Functional Requirements for Bibliographic Records: Final Report* (FRBR)—is not a full data model, by the report's own admission and by the critiques of others. Nevertheless, the report is often thought of as a data model, with discussions of a hierarchy of entities and inheritance of attributes despite neither of these concepts being made explicit in the report.

This paper acknowledges the FRBR model's inadequacy as a data model and argues that the model should more fully embrace its functional approach to allow for machine learning about bibliographic information based in part on strict inference through entailment but in large part on normativity and assumptions through implicature.

Functional Requirements for Bibliographic Records: Final Report (FRBR)

The FRBR report was first published by the International Federation of Library Associations and Institutions (IFLA) in 1998. It has as one of its two aims "to provide a clearly defined, structured framework for relating the data that are recorded in bibliographic records to the needs of the users of those records" (IFLA 2008, 7). While this was originally aimed at increasing interoperability of catalog records, it has come to be thought of as part of an effort to rethink library catalogs entirely, abandoning the last vestiges of catalog cards.

Most interest in and discussion of FRBR focuses on the entities in "group 1": *work*, *expression*, *manifestation*, and *item*. (Henceforth, these terms will be used only in the FRBR sense.) Definitions and examples are given in the table below:

Entity	Definition	Examples
work	"a distinct intellectual or artistic creation" (IFLA 2008, 17)	Bulgakov's *Master and Margarita*
expression	"the intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of such forms" (IFLA 2008, 19)	the text of the first version, which Bulgakov burned in a stove; the censored version published in *Moskva* magazine in 1966–67; the English translation by Michael Glenny

[1] The author thanks Olga Khomitsevich for translating this paper into Russian (using translations of FRBR terminology by T. A. Bakhturina) and greatly appreciates the help Dave Dubin, Allen Renear, and Richard Urban in the development of this paper.

manifestation	"the physical embodiment of an expression of a work" (IFLA 2008, 21)	the Glenny translation published in paperback by Harper & Row; the Glenny translation published as an audiotape
item	"a single exemplar of a manifestation" (IFLA 2008, 24)	my copy of the Harper & Row paperback edition of the Glenny translation

Note that FRBR explicitly includes non-print media.

The FRBR group-1 entities are often thought of as constituting a hierarchy, with work at the top and item at the bottom (see, for example, Mimno et al. 2005), but the FRBR report is ambiguous on whether these entities form a hierarchy[2] and, furthermore, it has been noted that the idea of a hierarchy for FRBR is problematic (Coyle 2007; Renear and Choi 2006). Nevertheless, it a convenient term for grouping these four entities when speaking loosely.

Besides these entities and others not considered here, the FRBR model also defines attributes. Only certain attributes apply to certain entities; some attributes, such as title, exist for more than one entity, but these are, strictly speaking, different attributes that happen to be homonymous. For example:

1. A work, expression, and manifestation can all have a title, but these titles need not be the same.
2. A manifestation may have a typeface, but this attribute does not apply to any other entity.
3. An expression may have use restrictions, whereas both a manifestation and an item may have access restrictions.

The FRBR attributes are even more clearly descended from cataloging terminology than the FRBR entities, and they strongly reflect cataloging practice and library needs. For example, the title of a work is likely to be a "uniform title" in cataloging, whereas the title of a manifestation is usually that text which appears on the title page. Another example is use restrictions and access restrictions: use restrictions usually derive from copyright, whereas access restrictions apply to a manifestation (in the case of technological anti-circumvention measures) or to an item (in the case of a policy that the item may only be used within the library). This strongly functional approach explains the model's appeal in the library community but also has led to its inadequacy as a formal data model.

FRBR as a data model

With such an extensive system of entities, attributes, and relationships, the FRBR model clearly provides the basis for a data model of the bibliographic universe. However, it is far from complete. Ambiguity over hierarchy is not the only way in which the framework underspecifies its underlying data model. The report admits to using entity-relationship analysis (IFLA 2008, 6) but gives a clear disclaimer that its conceptual

[2] There is only one use of the term "hierarchical" in the FRBR report, in the introduction: "Further study could be done on the practical implications of restructuring MARC record formats to reflect more directly the hierarchical and reciprocal relationships outlined in the model" [IFLA 2008: 6].

model "does not carry the analysis to the level that would be required for a fully developed data model" (IFLA 2008, 3). Renear and Choi (2006) provide an overview of difficulties fitting the FRBR model into a strict entity-relationship model.

Those interested in formulating a more rigorous version of the FRBR model, which could be used as a formal ontology by computer applications, are developing an object-oriented definition of FRBR called $FRBR_{OO}$ (International 2008), which attempts to harmonize FRBR with an existing ontology for the museum community called the CIDOC Conceptual Reference Model (CRM) (ICOM/CIDOC 2007). CIDOC CRM is considerably more complex than the FRBR model but provides a number of advantages, including modeling the way in which bibliographic entities come into existence.

I acknowledge that $FRBR_{OO}$ provides a much more solid foundation for developing a true data model of the bibliographic universe. However, I believe the functional approach of the original FRBR model has merit for enhancing common uses of library catalogs. Below I offer proposals for extending and loosening the FRBR model to more closely correspond with its perception and with common conceptions of how bibliographic entities relate to one another.

How do FRBR group-1 entities come into existence?

According to $FRBR_{OO}$, an expression is created, leading to the creation both of a work and a "singleton of a manifestation," which is roughly equivalent to an item in the original FRBR model (International 2008, 14). I find this model of simultaneous instantiation of entities appealing, so I propose a similar approach for the original FRBR model. Let's continue with the example of *Master and Margarita*.

Bulgakov composed the novel on handwritten manuscript pages in 1928 and 1929. I propose that at the point of composition, all four FRBR group-1 entities—work, expression, manifestation, and item—were instantiated. The act of creation *entails* the existence of all four entities.

Many works, however, have more than one expression. In the case of *Master and Margarita*, the author actually created multiple versions, each of which constitutes an expression in the FRBR model. Later scholarly editions and translations also constitute new expressions. For an expression to be published, it is typeset or otherwise prepared for distribution: each distinct version for publication is a new manifestation. Finally, copies of this manifestation are made: each copy is an item. There is a cascade down the hierarchy, often with more than one expression of the work, more than one manifestation of each expression, and more than one item for each manifestation. All of these related entities constitute a "bibliographic family," to use a common term (Taylor 2007, 73–78).

Note that derivative works—sequels, fan fiction, and other works inspired by the original—may be created as well. These fall outside the scope of this paper.

At what level in the hierarchy do FRBR attributes apply?

$FRBR_{OO}$ shares with the original FRBR model a significant deficiency: there is no explanation for how an attribute comes to be assigned to an entity. For example, how does a work, expression, or manifestation acquire its title? How are access restrictions on an item related to access restrictions on a manifestation? A data model for bibliographic entities needs inference rules that allow a machine to determine—or at least guess—that which is not stated explicitly.

It seems to me more useful for users if we instead think of the FRBR attributes as potentially existing at any level of the hierarchy, "inherited" from previously existing entities within the bibliographic family yet likely to be "overridden" during the instantiation of the new entity. This leads to use of attributes in a way more associated with everyday thinking. Let's again consider Bulgakov:

Bulgakov wrote his first draft manuscript, instantiating not only an item but a manifestation, expression, and work. His creation has attributes like title, author, and access restrictions which apply equally at each entity level, even though at this point there is only a single item that exists in the real world.

Later, as additional expressions are created, these attributes are carried down the hierarchy but may change at any point in the cascade. Taking the title as an example, an editor may modify it (at the level of the expression), or a typesetter may change the capitalization (at the level of the manifestation), or the owner of a copy may deface it (at the level of the item). In a case such as a defaced title, it makes sense to speak about the title of that item as defaced, even though FRBR does not allow for the title attribute at the level of an item.

Likewise, consider use restrictions and access restrictions: a use restriction on an expression (like copyright or a license) applies to all manifestations and items derived from that expression, yet if a manifestation of the expression is later released with a Creative Commons license, this revised use restriction would apply to all items of that manifestation.

Normativity and implicature

Indeed, pondering the many levels at which attributes seem to apply leads to a realization of the normative nature of FRBR entities. If a book (an item) is missing a page, it does not cease to exemplify a particular manifestation which is also exemplified by other copies of that edition. If an edition (a manifestation) is published under a different title from previous editions, it does not cease to embody the expression. In the bibliographic universe, there is much information which you can assume but little which, if you learned otherwise, would change your basic assumptions.

As Renear and Choi document, the literature on FRBR makes frequent reference to "inheritance" along the lines of the cascade of entities described above. They show that inheritance is insufficient as an explanation within the FRBR model and suggest "modal notions" as a way around this (Renear and Choi 2006). I propose instead using *implicature* to describe the behavior of attributes. Implicature, a term coined by linguist Paul Grice, denotes that an inference may be drawn but need not be true. Given the existence of an expression, manifestation, or item about which you do not have full information, you can infer attributes from previously existing entities.

It's evident that implicature rather than entailment means a machine cannot draw conclusions but only make statements about what is possible. Still, such statements of possibility could turn out to be quite useful, and they in fact model human behavior when dealing with bibliographic families. When searching for bibliographic entities, you assume they share attributes with other entities in that bibliographic family but are not surprised to find divergences.

Conclusion

While it is clear that FRBR is insufficient as a data model, it is based not only on cataloging practice but also on intuitive notions of how entities in bibliographic families relate to one another. Users searching for bibliographic entities make inferences based on known information, but few definite conclusions can be drawn. If we make tools to assist users in searching for information and make use of the FRBR model, it might be useful to consider entailment of entities and implicature of attributes.

In particular, this revised model allows attributes to exist at nearly any level of the FRBR model. I believe that in a functionally driven model like FRBR, it may be useful to allow for an even looser ontology than that proposed in the report.

Works cited

Coyle, K. 2007. "Hierarchy v. relationships," *Coyle's inFormation: Comments on the digital age, which, as we all know, is 42.* http://kcoyle.blogspot.com/2007/11/use-of-hierarchy-as-organizing.html.

ICOM/CIDOC Documentation Standards Group. 2007. *Definition of the CIDOC Conceptual Reference Model.* Version 4.2.2. http://cidoc.ics.forth.gr/docs/cidoc_crm_version_4.2.2.pdf.

IFLA Study Group on the Functional Requirements for Bibliographic Records. 2008. *Functional requirements for bibliographic records: Final report.* As amended and corr. through February 2008. http://www.ifla.org/VII/s13/frbr/frbr_2008.pdf.

International Working Group on FRBR and CIDOC CRM Harmonisation. 2008. *FRBR object-oriented definition and mapping to $FRBR_{ER}$.* Version 0.9 draft. http://cidoc.ics.forth.gr/docs/frbr_oo/frbr_docs/FRBR_oo_V0.9.pdf.

Mimno, D., Crane, G., Jones, A. 2005. Hierarchical catalog records: implementing a FRBR catalog. *D-Lib Magazine* 11, no. 10 (Oct. 2005). http://www.dlib.org/dlib/october05/crane/10crane.html.

Renear, A., Choi, Y. 2006. *Modeling our understanding, understanding our models: The case of inheritance in FRBR.* http://eprints.rclis.org/archive/00008158/01/Renear_Modeling.pdf.

Taylor, A. 2007. *Understanding FRBR: What it is and how it will affect our retrieval tools.* Wesport, Conn.: Libraries Unlimited.