

Identifying resources: FRBR and accessibility

Chris Oliver

Head, Resource Description and Metadata Services, University of Ottawa Library, Ottawa, Canada.

christine.oliver@uottawa.ca



Copyright © 2016 by Chris Oliver. This work is made available under the terms of the Creative Commons Attribution 4.0 International License: <http://creativecommons.org/licenses/by/4.0>

Abstract:

This paper will outline some of the key aspects of the FRBR family of conceptual models that support resource discovery especially for persons who are blind, visually impaired, or otherwise print disabled. The FRBR family of models have had a significant influence on the ways in which communities around the globe perceive and understand the bibliographic universe. This paper will focus on two areas where the conceptual models have had an important impact: bibliographic information as data and the precise delineation between content and carrier. The paper focuses on these two areas because they are of particular interest for a user with a print disability who approaches the task of discovering an appropriate resource. FRBR modelling, as expressed in the original models or in the new consolidated model, FRBR-LRM, offers a roadmap for structuring metadata in ways that allow more options for resource discovery in an increasingly global context.

Keywords: FRBR, FRAD, FRSAD, print disabilities, metadata

Introduction

In 1998, IFLA published FRBR, *Functional Requirements for Bibliographic Records*,¹ a conceptual model that changed the way we think about bibliographic information. Two extensions of FRBR were subsequently published, one for authority data in 2009, and one for subject authority data, in 2011: *Functional Requirements for Authority Data*;² *Functional Requirements for Subject Authority Data*.³ The three models together are often referred to as the FRBR family of conceptual models. This paper will look at key aspects of FRBR modelling that support resource discovery especially for persons who are blind, visually impaired, or otherwise print disabled.

The FRBR family of conceptual models are closely inter-related and present a way of understanding the bibliographic universe that is independent of encoding standards or cataloguing rules. They are based on a detailed analysis of real bibliographic and authority information, as recorded by libraries around the world. The conceptual models clarify the structure of the information and how pieces of information are related to each other. They

represent a “commonly shared understanding” of the nature of the bibliographic and authority information recorded by the library community. As IFLA standards, this commonly shared understanding can also be considered to be a globally shared understanding, or a kind of common language for discussing bibliographic and authority information.⁴

The models promote a view of the bibliographic universe where the focus is on what is important to the user. When a cataloguing standard is founded on this underlying theoretical framework, the result is a more user-focused set of instructions. Resource descriptions created according to such standards are structured to promote improved resource discovery for all users.

When the modelling is used in practical implementations for bibliographic data, it has the potential to make it easier for users to identify the content they need and match it with a carrier that delivers the content in a format they can access. FRBR modelling, as expressed in the original models or in the new consolidated model, offers a roadmap for structuring metadata in ways that allow more options for resource discovery in an increasingly global context.

Also, it is significant that the three IFLA groups who developed the models used a well-established modelling technique: entity-relationship modelling, a technique used by those who model data in domains such as software engineering. Thus the resulting conceptual models are understandable beyond the library domain because the specificities of our data are represented using a modelling technique that is widely recognized in other domains. But this decision to use entity-relationship modelling was also an important step in changing the library’s perspective of bibliographic information, from strings of information embedded in paragraphs to data that can be handled efficiently in digital environments.

The three models have had a significant influence on the ways in which communities around the globe perceive and understand the bibliographic universe. This paper will focus on two particular areas: bibliographic information as data and the precise delineation between content and carrier. The delineation between content and carrier will cover both the significance of the expression entity for filtering content as well as the impact of disentangling content from carrier. The paper focuses on these areas because they are of particular interest for a user with a print disability who approaches the task of discovering an appropriate resource.

The models: a quick overview

As background for understanding the impact of FRBR on resource description, the paper includes a quick overview for those unfamiliar with the FRBR family of conceptual models. Since the areas that are the focus of this paper are primarily influenced by the original FRBR model, this overview will focus primarily on FRBR, with some references to the two later extensions.

The focus for the models is the perspective of what users need to successfully discover and access resources that match their requirements, as can be seen in the introduction to FRBR:

The aim of the study was to produce a framework that would provide a clear, precisely stated, and commonly shared understanding of what it is that the

bibliographic record aims to provide information about, and what it is that we expect the record to achieve in terms of answering user needs. (FRBR 1.1)

The models look at bibliographic/authority information from the user's perspective. The focus is not on the cataloguer creating a single record, but on the user searching for a resource within the context of a large catalogue, database, and ultimately, the web.

There are three components in an entity relationship model: entities, attributes or characteristics of the entities, and relationships between the entities. The FRBR entities are the objects of interest to users of bibliographic data, and are divided into three groups:

Group 1 entities: products of intellectual or artistic endeavour
entities: work, expression, manifestation item

Group 2 entities: those responsible for the intellectual or artistic content, the physical production and dissemination, or the custodianship of the entities in the first group
entities: persons, corporate bodies
FRAD and FRSAD added another entity: families

Group 3 entities: subjects
entities: concept, object, event, place + all the entities in groups 1 and 2
FRSAD introduced the generalized entity: *thema* = any entity used as a subject of a work

The group 1 entities are basically the substance of our library collections. The group 2 entities are those agents with a relationship of responsibility for the group 1 entities. Group 3 are subject entities and the model no longer attempts to encompass a categorization of subject entities, since there are many ways to represent subjects. By introducing the generalized *thema* in FRSAD, the FRBR family of models allows each community to divide subjects according to the needs of their domains, cultures, etc.

The group 2 and 3 entities are fairly self-explanatory. The group 1 entities present a challenge because they are both straightforward and puzzling. The terms “work”, “manifestation” and “item” are relatively familiar terms. The model also defines the entity “expression”, adding an important layer between work and manifestation. Their definitions capture the relationships between them: an item is the exemplar of a manifestation; a manifestation is the embodiment of an expression; an expression is the realization of a work. These four entities are conceptually distinct but all four entities are present in each resource. The four entities represent different aspects of the resource, and support user tasks in distinct ways.

Work is a very high-level abstraction and represents the “commonality of content” between various expressions of the work.⁵ The work is “the intellectual or artistic creation” and the expression is “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. An expression is the specific intellectual or artistic form that a work takes each time it is “realized.” ” (FRBR 3.2.2). An expression brings the abstract work into a form that can be communicated. Manifestations embody the expression; they allow the content to be fixed and exchanged. Libraries collect manifestations, they buy and lend them. Manifestations are normally collected because of the content they embody, that is, the expression delivered in that particular manifestation. Item is the concrete entity; an item is owned or loaned. Each manifestation has at least one item exemplifying it, such as in the case of a manuscript, but usually has multiple items, such as the print run of a book with multiple identical copies.

Each entity has a set of characteristics or attributes. The entity is an abstract organizing category about which there are certain types of data. The attributes of an entity are the data usually recorded. The models set out some attributes but they do not claim to be exhaustive.

Examples of attributes:

<i>For an item</i>	item identifier (e.g. barcode number)
<i>For a manifestation</i>	publisher and date of publication
<i>For an expression</i>	language of the content
<i>For a work</i>	form (genre)
<i>For a person</i>	dates of birth and death

Though FRBR listed *name* as an attribute, the FRBR family of models now treats name as a separate entity. FRAD and FRSAD introduced this important step forward in modelling by separating an entity from the name or names associated with it. The decision to make *name/nomen* a separate entity introduced greater flexibility since one entity can have different names. For example, this separation accommodates various ways of handling pseudonyms in different cataloguing traditions. It also allows one to record attributes associated with a name, but not with the person, subject, etc., such as script or time of validity.

Entity relationship modelling has to be understood outside the traditional library record structure. Entities are not held together by being in the same “record”. Libraries working in a MARC environment are accustomed to pieces of information being held together within the shell of a bibliographic or authority record. There is no record in the entity-relationship model. The entities are held together in a meaningful way solely through the relationships. Each model, from the perspective of its area of focus, identifies and defines the key types of relationships between entities that will support the task of resource discovery. Relationships carry information about the nature of the links that exist between entities, enable collocation, and provide pathways to improve navigation and exploration.

Examples of relationships:

<i>Entity</i>	<i>Relationship</i>	<i>Entity</i>
work	<i>created by</i>	person
work	<i>is based on</i>	another work
expression	<i>is a realization of</i>	work
manifestation	<i>published by</i>	corporate body
item	<i>owned by</i>	family
person	<i>translated</i>	expression

This overview highlights a few of significant areas to provide background for the next two sections.

Bibliographic information as data

As mentioned before, the three IFLA groups who developed the models used a well-established modelling technique, entity-relationship modelling. The advantage of choosing a well-known technique is that it made the models understandable for other domains beyond the library. MARC data has tended to keep bibliographic information restricted to the library silo because software developers and database designers were unsure of how to deal with information encoded in this very-library specific schema. The three conceptual models were never intended to be data models, but they present the possibility of developing data models based on their more abstract-level modelling. Entity-relationship modelling opened the door to possibilities for data interoperability between different domains, better designed applications that use bibliographic data, and usability of this data in new technological environments, such as the web.

Entity-relationship modelling also had an important impact on the library's perception of its own bibliographic information. The models were developed out of the analysis of actual cataloguing records, looking first at bibliographic information, and then at authority information. So the models took bibliographic information that had been stored as strings of characters, embedded in paragraphs, and projected them into entity-relationship models. All of a sudden, one could start seeing these strings of characters as having the potential to be treated as data, data that could be manipulated by computer programs, published in new encoding formats, and be discoverable on the web, even in the context of the more complex semantic web of linked data. FRBR demonstrated that bibliographic information could be validly broken down into the types of components that are the basis for schemas, models and diagrams in the world of computing. In fact, if this kind of information can be modelled in entity-relationship models, it can also be represented with other modelling techniques, such as object-oriented modelling, as seen in FRBRoo: an interpretation of the FRBR family of conceptual models, using an object-oriented methodology.⁶

Every piece of information usually recorded by libraries could now be seen as an independent, granular piece of information that could be clearly defined, labelled and differentiated from other types of information. In the past practices of catalogue cards and MARC records, many important pieces of information were buried either in concatenated, long strings of data, or in elements such as general notes that could not be rigorously identified and separated out for the purpose of navigation or data display.

FRBR divided bibliographic information into separate attributes, each attribute defined and associated with only one entity, and into specific and precise relationships between the entities. Each piece of information could now be seen as a data element rather than an embedded string. If one looks at *Resource Description and Access (RDA)*,⁷ a cataloguing standard built on the framework outlined in the FRBR family of models, it is a set of instructions for recording data elements. The information traditionally associated with bibliographic and authority records is divorced from previous encoding practices and arranged as the recording of data elements, each element containing one distinct, precisely-defined type of information about a single attribute or a single relationship. The aim is to record data that is well identified, unambiguous and segmented to a suitable level of granularity for processing in the digital environment. It also means that any data element has

the potential to be used as the basis for an index, as a filter for searches, as a facet for navigation, or as a way to present well-labelled displays for the user. For example, if font size were always recorded in one consistently identified data element, then one could reliably identify/retrieve large print resources. If information about the form of tactile notation were always recorded in one consistently identified data element, then one could reliably separate braille and Moon code resources.

The first step towards improving resource discovery for both a general and a specialized audience is the step of thinking about bibliographic information as data. If a type of information is identified as a significant attribute, and consistently recorded in its own identified data element, this is the first step towards building better discovery systems because the data is made available in a way that is suitable for the automated environments in which we function. By modelling bibliographic information using the entity-relationship technique, the FRBR family of conceptual models played a significant role in shifting the library's perspective from strings to data.

Content and carrier

FRBR's separation into four group 1 entities, work, expression, manifestation and item, allows for a more precise definition of the boundaries between content and carrier. Work and expression are about content; manifestation and item are about carriers. When looking at the relationships between resources, the entities expression and manifestation play a key role in conveying information about the level of difference and similarity between resources.

Expression entity

Content is not assumed to be a single entity, *work*, but is differentiated into two entities, *work* and *expression*. This differentiation allows for a clearer definition of the relationship between content that is similar but not identical.

In many cases, a work is often realized in only one expression. But works that have formed an important part of our cultural and intellectual history are often realized in many expressions, works such as sacred scriptures and literary classics. A work, such as *Robinson Crusoe*, has many expressions.

Some expressions are translations of the original English text into other languages, such as French and German translations. In these cases, the differentiation between expressions is important to the user: are they searching for the text in a language they can read, or do they require the text from the original expression? Expressions that are translations of the same work realize the same intellectual or artistic creation, but use different languages. Each expression realizes the same work, but every word is actually different in each of these expressions.

Some expressions may be realizations into a different form of expression, for example, *Robinson Crusoe* as a spoken word version (such as an audiobook) instead of one in alphanumeric notation (such as printed book). In the FRBR model, a significant and inherent attribute of the expression is the "form of expression", defined as: the means by which the work is realized (e.g., through alphanumeric notation, musical notation, spoken word, musical sound, cartographic image, photographic image, sculpture, dance, mime, etc.). (*FRBR 4.3.2*) Any change in form results in a different expression. (*FRBR 3.2.2*) An

audiobook version of the work embodies a different expression than the textual version of the work, even if the exact same words are used in both.⁸

If one builds a resource discovery system that is able to focus attention on the expression entity, there are two possible effects on resource discovery results: 1) through the relationship to the same work, the user can be presented with results that show all the available expressions of the work; 2) at the same time, the user can also be presented with results that make clear the variations among these expressions. For a unilingual user, language may be critically important, because the user can only understand French. For a print-disabled reader, the form of expression may be critically important, because they need content expressed as spoken word rather than text. It may be either the relatedness or the differences that are of interest to a particular user.

When a user is not able to use the whole spectrum of available resources, we need to develop systems that allow that user to pinpoint the resources that respond to his/her need. The FRBR family of conceptual models presents a roadmap for identifying the key elements of information that are important to users. By identifying four entities that are important, work, expression, manifestation, and item, FRBR introduced a significant level of differentiation between resources. Cataloguing standards built on the foundation of the FRBR model pay attention to the importance of this data, and provide instructions to ensure the clear and explicit recording of such data. For example, in *Resource Description and Access (RDA)*, it is mandatory to record content type; content type is a categorization of forms of expression.

RDA's definition of content type is a useful insight into the meaning of expression: a categorization reflecting the fundamental form of communication in which the content is expressed and the human sense through which it is intended to be perceived (*RDA 6.9.1.1*). As soon as one is looking for alternative formats, one of the significant questions may be through which sense is the content best delivered for this user. RDA explicitly draws a line connecting expression, communication and human senses.

Identical content delivered in braille notation would also be considered a different expression from the alpha-numeric notation of the same work. It is a notation, but it is a tactile notation rather than an alpha-numeric notation and requires the use of a different sense. Again, the explicit and consistent recording of the form of expression is crucially important for a user who cannot rely on alpha-numeric notation.

A user may come across a spoken word resource with the title *Robinson Crusoe*. But this may not be a reading of the complete text by Daniel Defoe. It may be an abridgement, and if the abridgment does not fundamentally alter the content, it can be considered an expression of the original work, another type of expression. But the spoken word resource with the title *Robinson Crusoe* may also be an entirely different work, albeit one with a relationship to the original. For example, if someone dramatizes the original novel, they have created a new work, but it is a related work, one based on the original *Robinson Crusoe*. The description of a resource needs to include sufficient information that a user clearly understands the relationship of the resource to the original work, either as an expression of that original work, or as another work based on the original one. For one user, a related work may satisfy their requirements, but for the student studying for an exam on Defoe's works, even the abridgment may be inadequate. Ranganathan's fourth law, *save the time of the reader*, applies in the digital age as much as in the traditional print age. In a small, self-contained collection, one might be able to ignore the role of expression-level attributes. But, as we consolidate

catalogues and look to publishing bibliographic information on the web in larger pools of data, insights from the FRBR family of models become increasingly important.

Even if a work has only one expression, it is still important to identify both the work and expression entities. At the point in time when the description of a resource is created, that work may only have one expression; at that point, the user may not need to be presented with information about both entities. But it remains important to record. At a later date, a new expression may be realized, or a related work may be created. Relationships then need to be recorded, and they must be recorded between the correct and appropriate entities. In a linked data environment, it will be easier to make appropriate links between entities if they have been properly modelled and recorded from the start.

Expression is an important entity because it adds a degree of precision in the delineation of similarities and differences between the content of resources. Expressions of the same work have a high degree of relatedness to the same work, but there are variations in the content. By carefully recording expression relationships, the user can know the exact relationship between the available expression and the work they need to consult; the user can also filter on significant expression attributes, such as language or form of expression, to generate results that are suitable for their requirements.

Manifestation entity

Each expression may be embodied/published in several manifestations.

Manifestations of the same expression embody identical content, though the manifestations are different: for example, identical content but different publishers, or different dates of publication. Looking at the attributes of manifestations, a key attribute is the form of carrier: the specific class of material to which the physical carrier of the *manifestation* belongs (e.g., sound cassette, videodisc, microfilm cartridge, transparency, etc.).(FRBR 4.4.9) The spoken word version of a work can be delivered to the user on a range of different carriers: audio cassette, CD, streaming audio. Each embodiment is a different manifestation.

Differences in the form of carrier can be as critically important as differences in the form of expression. If the content is embodied in a carrier that the user cannot access, such as a floppy disk, then the content is useless. If the user requires a mediation device which is unavailable to them, such as a CD player, then the resource is useless.

If one looks at the relationship between large print and regular print versions of the same identical content, they are both delivered on the same form of carrier, physical volumes, but they have different type sizes. Type size is a manifestation level attribute. A difference in type size is likely to be accompanied by other differences, such as the extent of the carrier, the manifestation identifier, such as ISBN, and maybe also the publisher. The relationship between the two resources would be two different manifestations of the same expression. In the cataloguing of the 20th century, many manifestation attributes were buried in the descriptions, as part of non-specific notes, or jammed into fields with other, completely different data, making it difficult to use the information as a reliable filter. In the FRBR model, type size is an identifiable attribute. For a user who has print disabilities, this piece of information can be of critical importance.

FRBR also lists typeface as an attribute that may be recorded, which may be useful in some cases for a user with print disabilities. However, each of the FR models explicitly states that none of the lists of attributes or relationships are intended to be exhaustive. Applications may also define additional specific attributes and/or relationships. One of the applications of the models, *Resource Description and Access (RDA)*, defines additional data elements that may be useful for a population looking to access alternative formats. For example, RDA includes a data element called encoding format (*RDA 3.19.3*), part of a parsing of FRBR's "file characteristics" into more granular attributes. The cataloguer is encouraged to record the encoding format if it judged to be important for the user when identifying or selecting an appropriate resource. DAISY is a good example of an encoding format that can make a significant difference for the user and should be recorded.

Conclusion

The FRBR models represent an important step forward in terms of our understanding bibliographic information. There are many ways in which the FR models have provided a better understanding, but there are two areas that are especially relevant to users with print disabilities: bibliographic information as data, and the separation of content and carrier.

In the first place, the models demonstrated that long strings of bibliographic information, intended only for humans to read and understand, can in fact be parsed into carefully defined units of information, modelled as entities, attributes and relationships, and treated as data, read by both humans and machines. The models represent a first step in making this data, that was locked into library catalogues, available in a wide range of technological environments, actionable by machines, and suitable for the linked data environment. This results in information about resources being more widely available and discoverable.

Secondly, the models allow for a precise identification and separation of information relating to content as opposed to carrier, especially through the four entities, work, expression, manifestation and item, and the relationships to and between them. This precise delineation makes it possible for the user to identify the nature of the content available in resources and also to filter by content types when a user can only access one particular type of content, such as person who requires audiobooks. Whether a user has a disability or not, it can be important to that user to distinguish between expressions, such as a text versus a spoken word version, or to know that several expressions all realize the same work, and so are equally useful for their purpose. Similarly, it can be important for a user to distinguish between manifestations with different characteristics, whether a CD or streaming audio, large print or regular print, etc. In some cases, a user may find manifestation-level attributes as important as or more important than expression-level attributes. But what is crucial is clarifying the nature of the similarities and differences between resources, and recording this information in a way that it can be used to guide the user to the appropriate resource. When the models are used in practical implementations, they have the potential to make it easier for users to identify the content they need and match it with a carrier that delivers the content in a format they can access.

The new consolidated model, FRBR-LRM⁹, brings the three models together into one coherent and consistent entity-relationship model, but it does not make changes to the original four entities that make up a library resource: work, expression, manifestation, and item. These entities, that are essential for the precise identification of content and carrier, remain. For future applications/implementations, it is important to note that the consolidated

model reiterates that attributes listed are not exhaustive. From the introduction to the section on attributes in the draft definition of the FRBR-LRM model:

The attributes listed under each entity are representative and are not in any way to be considered an exhaustive listing of attributes that might be determined to be useful in a particular application. An application can define additional attributes to record additional relevant data or to record data at a greater level of granularity than is illustrated. (FRBR-LRM 4.2.1)

Thus, a specific community of users may need more attributes to describe their resources, or attributes at finer levels of granularity. The underlying FRBR model in no way prevents the definition of attributes required to produce satisfactory resource descriptions for a specific user group. What is crucial is associating newly defined attributes with the correct entity so that the data is consistent with the model, and thus also fits together with data from its own user community and as well as with data from other metadata communities.

FRBR modelling, as expressed in the original models or in the new consolidated model, provides a logical and consistent framework for understanding the principles and structure underlying bibliographic data. By giving us a better understanding of this underlying structure, the models present insights for improved resource discovery, both for general and specific audiences, whether in traditional databases or in the linked data environment. With their international level of recognition and use, the models also promote data interoperability, which is increasingly important in the global context of the web.

References

¹ IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records: Final Report*. (München: K.G. Saur, 1998).
<http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>

² IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). *Functional Requirements for Authority Data: A Conceptual Model*. (München: K.G. Saur, 2009).
<http://www.ifla.org/publications/functional-requirements-for-authority-data>

³ IFLA Working Group on the Functional Requirements for Subject Authority Records (FRASAR). *Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model*. (Berlin/München: De Gruyter Saur, 2011). <http://www.ifla.org/node/5849>

⁴ Pat Riva. "Introducing the Functional Requirements for Bibliographic Records and Related IFLA Developments." *Bulletin of the American Society for Information Science & Technology* 33, no. 6 (2007): pages 9-10. www.asis.org/Bulletin/Aug-07/Riva.pdf

⁵ "A work is an abstract entity; there is no single material object one can point to as the work. We recognize the work through individual realizations or expressions of the work, but the work itself exists only in the commonality of content between and among the various expressions of the work. When we speak of Homer's Iliad as a work, our point of reference is not a particular recitation or text of the work, but the intellectual creation that lies behind all the various expressions of the work." *FRBR 3.2.1*, page 17.

⁶ FRBRoo version 2.4 at the IFLA website (after world-wide review and before receiving approval as an IFLA standard): http://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf

See also Patrick Le Boeuf. *A basic introduction to FRBRoo and PRESSoo*. IFLA Library, 2015.
<http://library.ifla.org/1150/>

⁷ *Resource Description and Access*. Chicago: American Library Association; Ottawa: Canadian Library Association; London: Chartered Institute of Library and Information Professionals (CILIP), 2010-. In *RDA Toolkit*: <http://www.rdatoolkit.org>

⁸ There are also cases where there are more than one expression element differentiating resources. A French translation of Robinson Crusoe in alpha-numeric notation is a different expression from the original English text. The same French translation in spoken word is a different expression from the French translation in text. The French spoken-word expression differs from the original in two expression-level attributes: language and form of expression.

⁹ Draft definition of the FRBR-LRM model available at the IFLA website:
http://www.ifla.org/files/assets/cataloguing/frbr-lrm/frbr-lrm_20160225.pdf

See also Pat Riva and Maja Žumer. *Introducing the FRBR Library Reference Model*. IFLA Library 2015. <http://library.ifla.org/1084/>