
Date : 18/06/2008

# FRBRoo: enabling a common view of information from memory institutions

by:
**Pat Riva**
Chair of FRBR Review Group, Bibliothèque et Archives nationales du Québec, Canada

**Martin Doerr**
ICS-FORTH, Crete, Greece

**Maja Žumer**
University of Ljubljana, Slovenia

**Meeting:** **156. Cataloguing**
**Simultaneous Interpretation:** English, Arabic, Chinese, French, German, Russian and Spanish

WORLD LIBRARY AND INFORMATION CONGRESS: 74TH IFLA GENERAL CONFERENCE AND COUNCIL
10-14 August 2008, Québec, Canada
http://www.ifla.org/IV/ifla74/index.htm

## *Abstract*

*In 2008 the FRBR/CRM Harmonisation Working Group has achieved a major milestone: a complete version of the object-oriented definition of FRBR (FRBRoo) was released for comment. After a brief overview of the history and context of the Working Group, this paper focuses on the primary contributions resulting from this work.*

- *FRBRoo is a self-contained document which expresses the concepts of FRBR using the object-oriented methodology and framework of CIDOC CRM. It is an alternative view on library conceptualisation for a different purpose, not a replacement for FRBR.*
- *This 'translation' process presented an opportunity to verify and confirm FRBR's internal consistency.*
- *FRBRoo offers a common view of library and museum documentation as two kinds of information from memory institutions. Such a common view is necessary to provide interoperable information systems for all users interested in accessing common or related content.*
- *The analysis provided an opportunity for mutual enrichment of FRBR and CIDOC CRM. Examples include:*
  - *Addition of the modelling of time and events to FRBR, which can be seen in its application to the publishing process*
  - *Clarification of the manifestation entity*
  - *Explicit modelling of performances and recordings in FRBR*
  - *Adding the work entity to CRM*
  - *Adding the identifier assignment process to CRM.*

- *Producing a formalisation which is more suited for implementation with object-oriented tools, and which facilitates the testing and adoption of FRBR concepts in implementations with different functional specifications and in different environments.*

## Background

During the years (1992-1998) that the library community, through IFLA's Cataloguing Section, was creating FRBR (the Functional Requirements for Bibliographic Records), the museum community, working through ICOM-CIDOC (the International Council of Museums -- International Committee on Documentation) was independently engaged in the creation of a conceptual model of the data relevant to the documentation of museum collections. Work on that model began in 1996 under the auspices of the CIDOC Documentation Standards Working Group, and a first version of the CIDOC CRM (Conceptual Reference Model) was released in 1998. In 2000, the CIDOC CRM Special Interest Group (CIDOC CRM SIG) was formed for the purpose of developing the model further. CIDOC CRM has now become an ISO standard, ISO 21127:2006.

The idea that both the library and museum communities might benefit from harmonising the two models was first expressed in 2000, during ELAG's (European Library Automation Group) 24th Library Systems Seminar in Paris. However, it was not until 2003, following the creation of the FRBR Review Group, that a joint group was officially established with the purpose of harmonising these two conceptualisations. The International FRBR/CIDOC CRM Harmonisation Group is both a working group of the IFLA FRBR Review Group, and a subgroup of the CIDOC CRM SIG. It is co-chaired by Martin Doerr (ICS-FORTH, Greece), chair of the CIDOC CRM SIG and the main author of CIDOC CRM, and Patrick Le Bœuf (National Library of France), the first chair of the FRBR Review Group. The Group has held 12 meetings since 2003, detailed reports can be found on the IFLA FRBR Review Group website.

In January 2008 the group completed version 0.9 of *FRBR : object-oriented definition and mapping to FRBR(ER)* and issued it for comment. That document presents the conceptual model known as FRBRoo in a single self-contained document.

## Purpose of FRBRoo and its relationship to FRBR (E-R)

The main task of the Harmonisation Working Group was to express the FRBR model with the concepts, tools, mechanisms, and notation conventions provided by the CIDOC CRM, in this way creating a fully interoperable extension to the CIDOC CRM. There were three reasons for the decision to adopt the CIDOC CRM methodology:

1. The object-oriented formalism is the only one known that allows for integrating multiple conceptual models into one, while still preserving the identity of the constituent parts.
2. As a consequence of generalising over the diversity of museum collections and tasks, the CRM contains more generic concepts than FRBR does.
3. CIDOC CRM is already an ISO standard.

It is easy to get too close to an issue to see the matter objectively. A great advantage of working with another community is that it requires that all take a step back and not take anything for granted. This allows clarification of the underlying meaning, finally leading to deeper insights. The museum community is a natural fit for this sort of cooperation, as both communities are in charge of cultural heritage, and produce information aimed at facilitating access to that heritage and enabling scholars' reasoning about it. However, a common view of information is necessary to provide interoperable information systems for those users interested in accessing common or related content. With a formal ontology to capture and represent the underlying semantics of bibliographic information, we can facilitate the integration, mediation, and interchange of bibliographic and museum information.

The process of “translating” FRBR into the object-oriented framework was also an ideal opportunity to evaluate the model's robustness, internal consistency, and its extensibility. Calls for “proof” of FRBR are somewhat puzzling as it is a conceptual model and not a mathematical theorem. A model can be viewed as more or less robust, or more or less useful, but not as proven. However, the process of expressing FRBR in a different formalism has had the result of verifying FRBR's internal consistency, and its potential for extensibility and applicability in related contexts. This has been a valuable line of evidence to support the view that FRBR does capture a useful conceptualisation of bibliographic data.

FRBRoo has to be seen as an interpretation of FRBR, it is therefore neither a new version nor a replacement. The library community will still refer to FRBR(ER) and use it in explaining the bibliographic universe. On the other hand, FRBRoo offers an easier path to implementation using object-oriented tools, particularly for the integration of heterogeneous information from the cultural heritage sector. It also enables interoperability with other core ontologies.

In contrast to ER models and other traditional data structures, an ontology aims at describing the world meant by an information system, rather than at being a data structure. It describes how the different things, concepts and processes in a “domain of discourse” are related. Since the ontology is described in a formal or objective way, it can be used to discuss which information elements a system should have, and how they should be connected, in order to create an effective information system that allows for managing a specific task. Thus the ontology is more detailed than is necessary for any particular information system. This richness provides a basis for deciding what the consequences are of neglecting parts of the possible information in terms of the questions the resulting system will be able to answer. Further, an ontology is arranged in hierarchies or levels of generalisation. This allows for recognising optimal simplifications of seemingly unrelated information elements. In this sense, FRBRoo does not address "requirements", rather the concepts suitable in formulating requirements.

FRBR models the outcomes (work, expression...) of processes (such as creation, realisation, planning) but does not deal with the processes themselves. FRBRoo, building on the approach of CRM, focuses on processes. This approach enables reasoning about the circumstances in which, for example, instances of works were conceived or realised. Such circumstances may be the object of research (for example in literary theory), but this research has not been well supported by current bibliographic tools. While we may argue that most libraries do not need to focus on such specialised research, it is important for a general model to support as many needs as possible. In a particular implementation the level of complexity has to be the result of an informed decision.

The current complexity or richness of FRBRoo is not the result of the imagination of its creators in describing the library world, but the result of modelling all the implicit concepts in FRBR(ER) needed to justify the reasons for certain attributes in FRBR(ER). This is formally described by means of a mapping from FRBR(ER) to FRBRoo. Based on a suitable selection of FRBRoo concepts, one can implement very simple information systems that still represent all the key features of FRBR(ER). This will be demonstrated in the near future. Furthermore, one can use FRBRoo in order to compare two different information systems, decide to which degree they conform to the FRBR model, and which is more effective for a particular task.

**Reading FRBRoo and CRM**

The first hurdle is to become familiar with the terminology used in object-oriented models.
An *entity* is called a "class"; a *relationship* is termed a "property". Properties are defined (the term used is "declared") as applying between a domain and a range; that is, they link an instance of the class which serves as the domain (or origin of the link) with an instance of the class which serves as the range (or target of the link). Attributes are also modelled using properties. Classes and properties occur in hierarchies; when a class is declared to be a subclass of another class, then every instance of the subclass is also a valid instance of the superclass, and inherits (in a strict sense) the properties of its superclass.

Understanding the naming conventions

All the classes in FRBRoo have both a name and an identifier constructed according to the conventions used in the CIDOC CRM model. The identifier for a class consists of the letter **F** followed by a sequential number. Properties are also given a name and an identifier, which consists of the letter **R** followed by a sequential number, the number is followed by the letter "**B**" every time the property is mentioned "backwards", i.e., from range to domain. **F** and **R** are to be understood as the first two letters of FRBR and do not have any other meaning. They correspond respectively to the letters **E** and **P** in the CIDOC CRM naming convention, where **E** originally meant "entity" (although in the CIDOC CRM "entities" are now consistently called "classes") and **P** means "property". Whenever CIDOC CRM classes are used in FRBRoo, they retain the name and identifier they have in the CIDOC CRM. A number of properties are identified by the letters **CLP** and a number; "CLP" stands for "CLass Property". Such properties indicate that all the exemplars of a given manifestation "are supposed to" or "should" display the features of the manifestation they belong to.

The names given to classes and properties are indicative of their meaning, but the real significance of the identifier plus name is to uniquely identify the class or property concerned, and allow one to look it up in the document. The actual meaning of a class or property is as given in its *scope note*. When a class or property is referred to in the text, both its identifier and name is given (F1 Work, rather than just "work").

Anatomy of a class declaration

Each starts with the Class identifier and name presented as a heading in bold face.
"Subclass of:" declares the superclass of the class, if any.
"Superclass of:" is a cross-reference to the subclasses of this class, if any.
"Scope note:" contains the textual definition of the concept the class represents.
"Examples:" introduces a list of examples of instances of this class. If the example is also an instance of a subclass of this class, the unique identifier of the subclass is added in parenthesis. If the example instantiates two classes, the unique identifiers of both classes are added in parentheses. Non-fictitious examples may be followed by an explanation in brackets.
"Properties:" introduces the list of the properties of the class. Each property is represented by its unique identifier, its forward and reverse names, and, after a colon, the range class that it links to.
(See the example F13 Identifier in the appendix)

Anatomy of a property declaration

Similarly, each starts with the Property identifier and name (with the reverse name in parentheses) presented as a heading in bold face.
"Domain:" declares the class for which the property is defined.
"Range:" declares the class to which the property points, or that provides the values for the property.
"Superproperty of:" is a cross-reference to any subproperties the property may have.
"Subproperty of:" declares the superproperty of the property, if any.

"Quantification:" declares the possible number of occurrences for the domain and range for the property. For example: (1:1,0:n) means that a single instance of the domain class is associated via this property to no or any number of instances of the range class.
"Scope note:" contains the textual definition of the concept the property represents.
"Examples:" contains illustrative examples showing how the property should be used.
(See the example R6 carries (is carried by) in the appendix)

Structure of the FRBRoo document

The document starts with an introduction and graphical overview of the model. The class and property hierarchies are presented next, followed by the full declaration of the FRBRoo classes (there are 33 classes) and properties (31 properties and 6 class properties). The next section is a mapping of the entities, attributes and relationships presented in FRBR(ER) to FRBR(OO). This section is particularly useful in understanding how E-R attributes are transformed into properties in the object-oriented framework. Another section reproduces the CRM classes (45) and properties (42) used in the FRBRoo class and property declarations. This demonstrates the extent to which CRM already included modelling of concepts that also occur implicitly in FRBR. Finally, the last section is an appendix on the modelling of the identifier creation process in cataloguing practice. This interesting section is presented as an appendix because it is not within the scope of FRBR, yet a clear understanding of this process was an important step in the development of FRBRoo.

## Enrichment of FRBR from CRM

Clarification of the Group 1 entities

The process of translating the entity-relationship FRBR into the object-oriented framework of CRM required careful analysis of the entities and their relationships in FRBR. For the FRBR group 1 entities (work, expression, manifestation, item) this analysis allowed the concepts behind those entities to be unpacked.

In the resulting FRBRoo, the entities *work*, *expression* and *manifestation* are broken down into multiple classes, each with specific properties. Thus FRBRoo has a class F1 Work, but also declares the subclasses F14 Individual Work, F15 Complex Work, F16 Container Work, F17 Aggregation Work, F18 Serial Work, F19 Publication Work, F20 Performance Work, F21 Recording Work. The class F1 Work is a superclass, encompassing the subclasses as specific cases, each of which has specific processes associated with its creation or assembly. This analysis is a step towards clearer understanding of the issues surrounding aggregates, issues that the Aggregates Working Group of the FRBR Review Group is considering within the E-R framework.

For the expression entity, FRBRoo has the class F2 Expression, with the subclasses F22 Self-Contained Expression, F23 Expression Fragment, F24 Publication Expression, F25 Performance Plan, F26 Recording. Expression fragments facilitate the modelling of the situation where extracts from one expression are reworked and become components of another work. The F24 Publication Expression was introduced to enable the modeling the publishers' intellectual contribution in the preparation of a manifestation.

The two aspects of the manifestation entity are declared separately. F3 Manifestation Product Type covers those manifestations which are the products of a publication process which produced F5 Items, the physical objects resulting from the publication process. It is a subclass of E55 Type which highlights the abstraction in its nature. In contrast, F4 Manifestation Singleton (meaning a set with a single member) covers those manifestations that were produced as single unique objects, for instance

manuscripts, preparatory sketches, authors' final drafts as sent to a publisher, making it a subclass of E24 Physical Man-Made Thing, a very concrete concept. As an abstraction, a publication cannot be said to have physical characteristics such as “consisting of” a type of material, or having a “number of pages”; these physical characteristics are found by a cataloguer on one of its exemplars, leading the cataloguer to extrapolate them to all the other exemplars with the same origin. This line of reasoning is modelled in FRBRoo by the class properties (CLPs), which are physical properties that apply to an abstract entity only through the physical things that exemplify it.

Modelling of time and events

FRBR takes a “static” point of view by modelling the bibliographic products that result from creation and production processes, without including any explicit modelling of those events. The lack of “event-awareness” in FRBR has frequently been noted. In the CIDOC CRM, temporal entities play a central role, as they are the only means to relate objects (either conceptual or physical) to time-spans, locations, and agents. The FRBRoo classes F27 Work Conception and F28 Expression Creation permit an E39 Actor, a E52 Time and a E53 Place to be linked to the F1 Work, F2 Expression and F4 Manifestation Singleton that are created. Thus we can see *(figure 1)* that first, the process of F27 Work Conception produces an idea, then the process of F28 Expression Creation produces simultaneously an F2 Expression and its first manifestation (in the form of a F4 Manifestation Singleton), which together realise a work (F1).

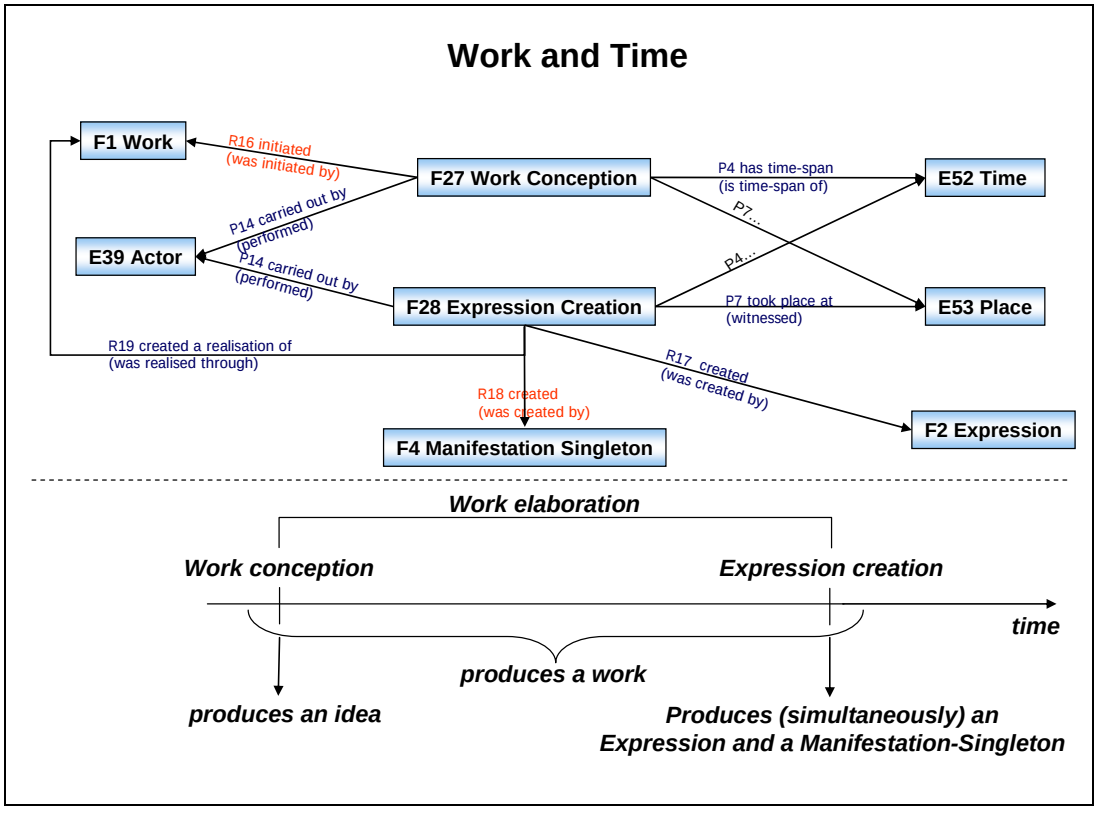


**Figure 1**

In addition, FRBRoo makes explicit the publisher's intellectual contribution, which is not modelled in FRBR. In the diagram *(figure 2)*, the author's contribution is seen on the left, consisting of Work, Expression and Manifestation Singleton. Then on the right, the publisher's contribution to the final product which results in the Manifestation Product Type, which embodies a Publication Expression, which in turn comprises both the author's Expression and the realisation of a Publication Work.

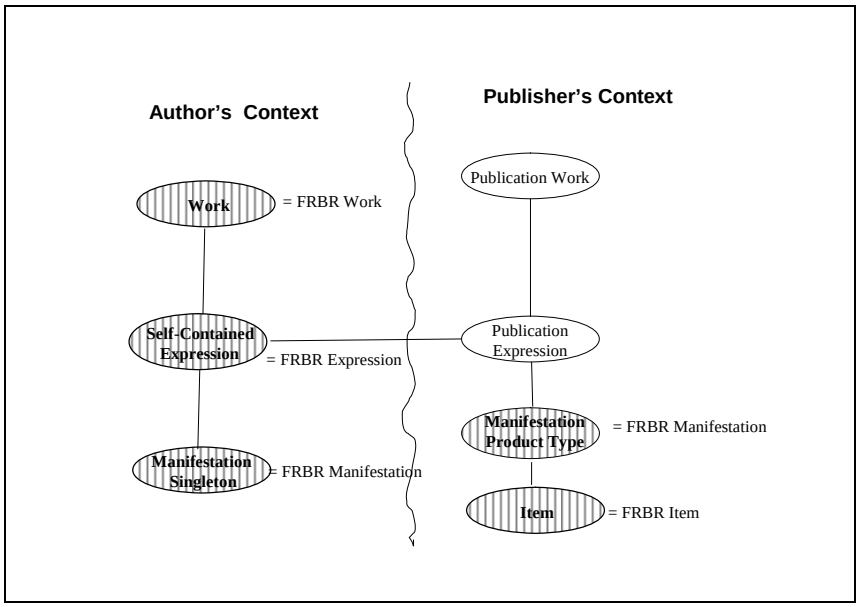


**Figure 2**

Note how the FRBR(ER) group 1 entities are mapped to this diagram: *work* and *expression* appear on the author's side, but *manifestation* and *item* appear on the publisher's side. This is another example of how the harmonisation process with CRM has resulted in drawing out concepts implicit in FRBR.

Modelling of performances and recordings

Another area where events and processes are implicit is in the modelling of performances and recordings. FRBR recognised that recorded performances are new expressions of works, but did not explicitly include the relationships between the source material and its performance. Using FRBRoo we can proceed step-by-step, or event-by-event, demonstrating how successive intellectual processes incorporate the results of the previous process, adding new elements of different natures. These additional contributions "add value" to the previous steps and also transform a source expression into a recording. Consider a play. The author and publisher have already produced a published text. The performers, by taking part in a F31 Performance (an event) according to a F25 Performance Plan, create a new expression. This may be, in turn, incorporated into a F26 Recording during a F29 Recording Event.

## Enrichment of CRM from FRBR

As in any good partnership, there have been benefits in both directions. The analysis provided for bibliographic processes in FRBRoo paved the way for the introduction of refinements into CIDOC CRM, so that the museum community's model could give a better account of mass production phenomena as they apply to certain categories of objects found in museum collections (such as the

printing of engravings), or the relationship between the creation of content (which is immaterial) and its physical carrier. Further, it introduced a basic model of intellectual conception and derivation applicable to all art forms. This required that the concept of work, among others, be integrated into CRM.

Although presented in an appendix to FRBRoo, the identifier assignment process presents features that are generally applicable and prove useful also in CRM. In this context, identifier assignment includes the construction of unique controlled access points based on names associated with specific entities, by an agency applying specific rules. These are concepts that are modelled in FRAD (Functional Requirements for Authority Data). An important implicit step in the process is the determination that a specific manifestation is "representative" of the expression that it embodies, and in turn, that a specific expression is "representative" of the work it realises. Characteristics of the representative manifestation or expression can then be abstracted "upwards" to the expression or the work, and used in creating the identifier.

The additions and refinements to CIDOC CRM resulting from the elaboration of FRBRoo have been sufficient in number and importance to warrant the preparation of an amendment package for the ISO version of CRM prior to the regular ISO cycle of amendments.

## Conclusion

What are the next steps? During this conference the FRBR Review Group will be discussing FRBRoo and any further issues and comments will be forwarded to the Harmonisation Group. The goal is to have a "version 1.0" ready for approval by the FRBR Review Group and the Cataloguing Section (as well as by the CIDOC CRM SIG) as soon as possible. In parallel, the Group has been working on a "core" FRBRoo, to follow along with the work on a "core" CRM, to show practitioners a simple way in which the key concepts can be turned into an application. As soon as FRAD, and eventually FRSAR, are approved and published, the mapping and harmonisation will be extended to include them.

**Bibliography**

*Definition of the CIDOC Conceptual Reference Model*, version 4.2.4, January 2008. Available at: <http://cidoc.ics.forth.gr/official_release_cidoc.html> (in .doc and .pdf)

Doerr, Martin. 2003. "The CIDOC CRM - An Ontological Approach to Semantic Interoperability of Metadata." *AI Magazine*, 24(3).

IFLA Study Group on the functional requirements for bibliographic records. *Functional requirements for bibliographic records : final report.* Munich, Germany : K.G. Saur, 1998. Also available at: <http://www.ifla.org/VII/s13/frbr/> (in HTML and .pdf)

IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). *Functional requirements for authority data : a conceptual model*. Draft, 2007-04-01. Available at: <http://www.ifla.org/VII/d4/FRANAR-ConceptualModel-2ndReview.pdf>

International Working Group on FRBR and CIDOC CRM Harmonisation. *FRBR : object-oriented definition and mapping to FRBR(ER)*, version 0.9 draft, January 2008. Available at: <http://cidoc.ics.forth.gr/frbr_drafts.html> (in .doc and .pdf) and at: <http://www.ifla.org/VII/s13/wgfrbr/FRBRoo_V9.1_PR.pdf>

ISO Technical Committee 46 "Information and documentation", Subcommittee SC4 "Technical Interoperability". *Information and documentation -- A reference ontology for the interchange of cultural heritage information.* ISO 21127:2006. Geneva : ISO, 2006.

Patrick Le Bœuf and Martin Doerr. "Harmonising CIDOC CRM and FRBR." *International Cataloguing and Bibliographic Control*, v.36, no.4 (Oct./Dec. 2007).

## Appendix: Examples of Class and Property Declarations

### *F13 Identifier*

Subclass of: F12 Name (= E41 Appellation)
Equal to: E42 Object Identifier

Scope note: This class comprises strings assigned to entities in order to identify them uniquely and permanently within the context of one or more organisations. Such codes are typically composed of alphanumeric sequences. The class F13 Identifier is not normally used for machine-generated identifiers used for automated processing unless these are also used by human agents. *[adapted from the Scope Note of CIDOC CRM E42 Object Identifier]*

F13 Identifier covers the notion of "controlled access points" in library practice – both preferred forms and cross references. A cross reference may not identify "uniquely" an entity, but can be shared by two entities; however, as it displays the same structural characteristics as preferred controlled access points, it is still regarded in the model as an instance of F13 Identifier.

Examples: ISSN "0041-5278" (F13)

ISRC "FIFIN8900116" (F13)

Shelf mark "Res 8 P 10" (E42)

"Guillaume de Machaut (1300?-1377)" (F13) [a controlled personal name heading that follows the French rules]

"Guillaume, de Machaut, ca. 1300-1377" (F13) [a controlled personal name heading that follows the AACR2 rules]

"Rite of spring (Choreographic work : Bausch)" (F13)

Properties: R8 consists of (forms part of): F12 Name

### *R6 carries (is carried by)*

Domain: F5 Item
Range: F24 Publication Expression

Superproperty of:
Subproperty of: E24 Physical Man-Made Thing. P128 carries (is carried by): E73 Information Object

Quantification: (1:1,0:n)

Scope note: This property associates an instance of F5 Item with the unique instance of F24 Publication Expression it carries.

Examples: The British Library's holding identified by shelfmark "DSC 9078.177 vol 19" (F5) *R6 carries* The entire content (text, layout, publisher logo, etc.) of the publication entitled "Functional Requirements for Bibliographic Records: final report", issued by publisher named "K. G. Saur" in 1998 (F24)