

A Basic Introduction to FRBR_{OO} and PRESS_{OO}

Patrick Le Boeuf

Performing Arts Department, National Library of France, Paris, France.

E-mail address: patrick.le-boeuf@bnf.fr



Copyright © 2015 by Patrick Le Boeuf. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License: <http://creativecommons.org/licenses/by/3.0/>

Abstract:

In addition to FRBR, FRAD, and FRSAD, two conceptual models for bibliographic information were developed these last years: FRBR_{OO} (a reformulation of the FRBR/FRAD/FRSAD models) and PRESS_{OO} (more specifically devoted to bibliographic information about continuing resources). This paper provides the explanation that is necessary for a good understanding of the formalism used in these two models. It then shows how some of the basic constructs of FRBR/FRAD/FRSAD are rendered in the FRBR_{OO} model. In a very brief third section, it highlights some core aspects of PRESS_{OO}.

Keywords: FRBR_{OO}, PRESS_{OO}, CIDOC CRM, Conceptual models for cultural heritage information

Among the working groups that are affiliated to the IFLA FRBR Review Group, there is the ‘Working Group on FRBR/CRM dialogue’, which was formed in 2003. Since then, this Working Group has been developing the FRBR_{OO} model, which is introduced on the IFLA Web site as ‘an object-oriented formulation of FRBR’.¹ Version 1 of this model was released in 2009, and version 2 in 2013.² Admittedly, the definition of this model is not easy to read; it is a document that does not lend itself to linear reading, from first to last page. The present talk aims therefore at providing an overview of the model, so that the audience can then look in the available documentation for more detail about how FRBR_{OO} models bibliographic information.

PRESS_{OO} is a more recent initiative; it was developed on the basis of FRBR_{OO} in 2013-14, not by IFLA, but by the ISSN International Centre in collaboration with the National Library of France, in order to fill the gaps left by FRBR in the modelling of bibliographic

¹ <http://www.ifla.org/node/928>.

² All final and draft versions are available from http://www.cidoc-crm.org/frbr_drafts.html. The RDF representation of FRBR_{OO} recommended by IFLA is available from the Open Metadata Registry, at <http://metadataregistry.org/schema/show/id/94.html>. An OWL representation, published by the University of Erlangen-Nuremberg, is available from <http://erlangen-crm.org/efrbroo>. An RDFS representation is available from the CIDOC CRM Web site, at <http://www.cidoc-crm.org/rdfs/FRBR2.1-draft.rdfs>.

information devoted to continuing resources.³ Version 1 was published in June 2014.⁴ As this second model is more limited in scope and does not result from an IFLA decision, this talk will provide a much briefer overview of it.

1 THE FORMALISM USED IN FRBR_{OO}

The first thing to do, if someone is familiar with the FRBR/FRAD/FRSAD models and finds FRBR_{OO} puzzling, is to get accustomed to the formalism used in FRBR_{OO}. Although the structure and terminology of FRBR_{OO} may seem, at first sight, very different from the structure and terminology of FRBR/FRAD/FRSAD, it is actually very easy to discover the same fundamentals in both families of models.

FRBR, FRAD and FRSAD are defined in three distinct texts⁵ that follow the same pattern: they define a number of *user tasks* that all other elements listed in the models should, to various degrees, contribute to fulfilling; they define a number of *entities* or basic notions in the universe of bibliographic discourse;⁶ they assign lengthy lists of *attributes* to each one of those entities, and these attributes serve to provide a description of any single thing that is classified as belonging to a given entity; and they enumerate the various *relationships* that should or may hold between individual things that are classified as belonging to those various entities.

1.1 User tasks

FRBR_{OO} does not mention any user task. This first difference can be explained by the context in which FRBR originated, in the early 1990s—over twenty years ago! The avowed purpose of the whole study on *Functional Requirements for Bibliographic Records*⁷ was to cut cataloguing costs by finding some ‘scientific’ way of assessing the value and utility of each single information element present in a catalogue record. User tasks were only meant to serve that purpose, by justifying any of the attributes and relationships declared in the model: each of them has to be recorded in some way, since it helps to fulfil a given user task among those listed at the beginning of the study. These user tasks are not an integral part of the *model* properly said; but they found their way into RDA, the cataloguing code that is based on FRBR.⁸ Indeed, their presence is more justified in a cataloguing code than in the definition of a conceptual model.

³ ‘Certain aspects of the model merit more detailed examination. The identification and definition of attributes for various types of material could be extended through further review by experts and through user studies. In particular, the notion of “seriality” and the dynamic nature of entities recorded in digital formats merit further analysis.’ IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records : Final Report* (München : K.G. Saur, 1998), p. 5.

⁴ *PRESS_{OO}: Extension of CIDOC CRM and FRBR_{OO} for the modelling of bibliographic information pertaining to continuing resources*. Version 1.0 ([Paris] : International ISSN Centre, 2014). Available from http://www.issn.org/wp-content/uploads/2014/02/PRESSoo_1-02.pdf or http://www.bnf.fr/documents/pressoo_v1.pdf.

⁵ There is a current effort, within the IFLA FRBR Review Group, to consolidate all three models in order to result in a single model defined by a single text.

⁶ Over time, unfortunately, these entities have come to sum up, in many people’s minds, the very essence of the model, which only represents a very partial and inaccurate way of understanding it. In the phrase ‘entity-relationship’, it is the term ‘relationship’ that should be stressed, rather than the term ‘entity’.

⁷ IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records : Final Report* (München : K.G. Saur, 1998).

⁸ <http://www.rdatoolkit.org/>.

Arguably, both FRBR/FRAD/FRSAD and FRBR_{OO} could be said to meet only one user need, to fulfil only one user task: *find*.⁹ Most other user tasks can be regarded as subservient variations on that unique basic task: *identify* and *select* are just ways of ensuring that one has *found* precisely what one was looking for, and *obtain* means that one has *found* access to something. The *explore* and *contextualize* tasks are slightly different in essence: they mean that the information recorded in catalogues should enable users to *find* what they were *not* looking for, by displaying the relationships that hold between what they were looking for and other things, the very existence of which they may not even have suspected (the serendipity objective).

The *justify* user task introduced by FRAD is the only one that is truly original, and the only one that truly relates to the ideals of the Semantic Web, since Tim Berners-Lee regards Trust as the top requirement without which the Semantic Web simply cannot exist at all.¹⁰ Ideally, any single information element asserted in a library catalogue should be *justified*, for people to be in a position to *trust* that catalogue—which, obviously, it is totally impossible to achieve, as this would result in a huge, uncontrollable rise of cataloguing costs.

All in all, both FRBR_{OO} and FRBR/FRAD/FRSAD should, therefore, only declare such information elements that are likely to *enable users to*:

- *find* precisely what they were looking for,
- *find* what they were *not* looking for but is related to it in some way (i.e., *explore*),
- and, to a lesser degree, *trust* the information they have *found*.

These are the only three user tasks that ought to be explicitly associated with any conceptual model for bibliographic information (or, more generally speaking, cultural heritage information), if any are actually required.

What it is important to retain from this development, is that FRBR_{OO} does not state any explicit user tasks at all. Users outside the library community are expected to be enabled to do whatever they want with the information that is modelled according to FRBR_{OO}; in the true spirit of Linked *Open* Data, they are not limited to a predetermined, reductive, and more or less arbitrary list of user tasks.

1.2 From entities to classes

In the formalism used in FRBR_{OO}, the term that was chosen to refer to the basic elements in the universe of bibliographic discourse is not *entity*, but *class*. Individual things that belong to a given class are said to be *instances* of that class (e.g., the book on my table is said to be an instance of the Item class; Shakespeare's play *Hamlet* is said to be an instance of the Work class).

In FRBR/FRAD/FRSAD, entities are only identified through a single noun or noun phrase in natural language (e.g., 'Manifestation', 'Corporate Body', etc.). In FRBR_{OO}, classes are identified through the combination of an arbitrary identifier and a label in natural language (e.g., 'F5 Item'). This is meant to facilitate multilingualism, as the identifier of a given class is supposed to remain the same, no matter into which language its label is translated. All class identifiers follow this pattern: they consist of the letter 'F' and an integer. The letter 'F' was arbitrarily chosen because it is the first letter in 'FRBR', and has no meaning by itself.

The label alone is not sufficient to allow readers to grasp the essence of a given class; labels only play the role of mnemonics, and may in some cases even be misleading if their assumed meaning is taken for granted. All implementers of FRBR_{OO} are firmly invited to

⁹ In FRBR, the four user tasks are: *find*, *identify*, *select*, and *obtain*. In FRAD, they are: *find*, *identify*, *contextualize*, and *justify*. In FRSAD, they are: *find*, *identify*, *select*, and *explore*.

¹⁰ This is Tim Berners-Lee's famous 'Semantic Web layer-cake' figure, now ubiquitous on the Web.

read carefully the *scope notes* that accompany the declaration of every single class and explain how it should be used. The phrase ‘scope note’ was chosen, rather than the term ‘definition’ as in FRBR/FRAD/FRSAD, because those paragraphs indicate the general characteristics that instances of a class are assumed to display. It should be kept in mind, though, that the crucial criterion for a thing to be categorised as an instance of a class are the properties that link it to other instances of the same class or of other classes.

A very important notion in FRBR_{OO}, which is absent from FRBR/FRAD/FRSAD¹¹ and which may contribute to make FRBR_{OO} somewhat more puzzling at first sight, is the notion of *inheritance*. This means that some classes are declared to be *subclasses* of others (which, consequently, are declared to be *superclasses* of the former). Any instance of a class that is declared to be a subclass of another class is automatically also an instance of that other class, the superclass (e.g., the F50 Controlled Access Point class is declared as a subclass of the F12 Identifier class, meaning that any thing that is classified as a controlled access point also displays the same characteristics and, more importantly, has the same properties as any identifier, more generally speaking).

The inheritance mechanism, while it is extremely efficient and economic, may also be deemed to make the model less easy to handle, because all properties that are declared for a given class are not repeated for all of its subclasses. As a consequence, it is not sufficient to look for the properties of a class: one should never forget to look also for the properties that are declared for all the superclasses of that class (and their own superclasses in turn), and that may apply to a single instance of that class (e.g., the *R15 has fragment (is fragment of)* property is declared at the level of the F2 Expression class, but is not repeated at the level of all the subclasses of F2 Expression—like, say, F26 Recording—although they all inherit that property from their superclass).

1.3 Attributes

In FRBR/FRAD/FRSAD, *attribute* is the generic name assigned to all the variables that serve to characterise an instantiation of a given entity and make it unique among all other instantiations of that entity. The sum of all attribute values associated with an instantiation of an entity amounts to a description permitting the identification of that instantiation (e.g., an instantiation of the Manifestation entity is described, among other information elements, through the sum of the ISBD elements associated with it, as all ISBD elements are listed among the attributes of the Manifestation entity).

Information elements that are modelled as mere attributes are those that do not call for further modelling, at least in the context of the bibliographic discourse: they are not deemed, *in that context*, to have any attribute or relationship of their own, and they can be expressed as mere strings (e.g., the Series Statement attribute of the Manifestation entity), as coded values picked from referentials, or both (e.g., the Language attribute of the Expression, which can be expressed as a string and/or as an ISO code). From the moment an information element needs, in the context of bibliographic discourse,¹² to be associated with something more than its mere statement as a string, it has to be modelled as an entity, with its own attributes and/or relationships.

¹¹ It will presumably be introduced in the consolidated model, though.

¹² This restriction is important. In a different context, such as the universe of discourse of linguistics, there would be obviously much more to be said about a given language than the ISO code that represents it, or its name in natural language, and the notion of language could not be modelled as a mere attribute. But in the context of bibliographic discourse, it was deemed quite sufficient to model languages as mere attributes, without providing hooks for further information about them.

In the formalism used in FRBR_{OO}, things are very different: there are simply no attributes at all. But there is a mechanism to associate an instance of a class with a string or a number or any other ‘primitive value’ that does not call for further modelling within the context of cultural heritage information—a universe of discourse which is broader than just bibliographic information, and which calls therefore for more classes than the original FRBR/FRAD/FRSAD.

Many notions that were merely declared as entity attributes in FRBR/FRAD/FRSAD now are modelled as classes in their own right in FRBR_{OO}, because it appeared that they were more complex than initially assumed. The most striking example is that of Date—an attribute which is declared for several entities in FRBR/FRAD/FRSAD. The initial assumption was that a date is a mere literal associated with, say, a Work or a Manifestation. Actually, the semantics conveyed by a date imply that an uncategorised event (which *should* be categorised—e.g., what is meant by ‘Date of the Work’?) occurred at some point in time, and that that point in time is referred to by a conventional phrasing which can vary according to cultural context (i.e., the *same* point in time is identified by several *distinct* dates in the Gregorian calendar, Julian calendar, Hebrew calendar, Islamic calendar, etc.). There is therefore more to be said about a date than just the string that expresses it in an unidentified, implicit calendar: the characteristics of that date and, more importantly, the precise nature of the event which took place at the point in time which is identified by that date and which should not remain implicit, are further information elements that need to be explicitly modelled for the Date attribute to become useful and meaningful in such a broad and unpredictable environment as the Semantic Web.

1.4 From relationships to properties

In traditional databases, it already is important to display the relations that hold between things; it is infinitely more so in a Linked Data environment. FRBR/FRAD/FRSAD declares, therefore, a number of *relationships* between entities (e.g., ‘Work is realized in Expression’).

In FRBR_{OO}, *relationships* are called *properties*—a different name for the same notion.

The same way as classes are referred to, in FRBR_{OO}, through the combination of an identifier and a noun or noun phrase, properties are referred to through an identifier and a verbal phrase (e.g.: ‘R2 is derivative of (has derivative)’).

Property identifiers consist of the letters ‘R’, ‘CLP’, or ‘CLR’, and an integer (e.g.: ‘R2’). The letter ‘R’ was chosen because it is the second one in ‘FRBR’, and does not have any more subtle semantic value. ‘CLP’ stands for ‘class property’, and ‘CLR’ for ‘class R property’; these seemingly difficult notions correspond to the generalisation of properties that are normally attached to physical objects, at the level of the object type to which these physical objects belong. Such explanation may read as rather abstract and abstruse, but actually this is what any cataloguer does every day without even noticing it: they observe that the individual physical item which they are describing has 48 pages, and they indicate in the bibliographic record which they are creating that all items which belong to that same publication are also assumed to have 48 pages: the publication is the abstract type to which the individual physical item belongs, and this *physical* property (to have a certain number of pages) is extended to the *abstract* type.

The class from which a property goes (e.g., the F28 Expression Creation class for property *R17 created (was created by)*) is said to be its *domain* (this corresponds to the notion of ‘subject’ in an RDF triple, where the property itself corresponds to the notion of ‘predicate’), and the class to which it points (e.g., the F2 Expression class for property *R17 created (was created by)*) is said to be its *range* (this corresponds to an ‘object’ in an RDF

triple). Except for those properties whose range is a ‘primitive value’ (e.g., a string or a number), all properties can be expressed in both domain-to-range and range-to-domain directions. This is the reason why the label that serves as a mnemonic for a property whose range is a primitive value consists of a single verbal phrase (e.g., *CLP57 should have number of parts*), while the label that refers to a property whose range is not a primitive value consists of two verbal phrases (the second one being enclosed within brackets) corresponding to the two directions from domain to range and range to domain (e.g., ‘R17 created (was created by)’, a label that indicates that it should be understood as ‘R17 created’ from F28 Expression Creation to F2 Expression, and as ‘R17i was created by’ from F2 Expression to F28 Expression Creation ; the ‘i’ that is added at the end of the property identifier stands for ‘inverted’: it indicates that the property is read from range to domain).

1.5 FRBR₀₀ as an extension of CIDOC CRM

Perhaps one of the most puzzling aspects of FRBR₀₀ lies in the fact that it is not a standalone model: it was developed as an *extension* of a pre-existing model, namely the CIDOC CRM, and it cannot be used without reference to that model.¹³ Some basic notions, such as the relation between persons and their names, or between works and their titles, were already modelled in CIDOC CRM and are therefore not repeated in FRBR₀₀. Users of FRBR₀₀ are expected to also re-use some of the structures modelled in CIDOC CRM. This is the reason why fragments of the CIDOC CRM definition are reproduced at the end of the definition of FRBR₀₀, so as to allow FRBR₀₀ users not to have to open two distinct documents simultaneously.

Developing a new model as an extension of a pre-existing model means that all the classes of the new model have to be declared as direct or indirect subclasses of classes from the pre-existing one, and all properties of the new model have to be related in some way to properties already declared in the pre-existing one.

The notation conventions are pretty much the same in both FRBR₀₀ and CIDOC CRM. Class identifiers begin with an ‘E’ in CIDOC CRM, instead of an ‘F’, and property identifiers begin with a ‘P’ instead of an ‘R’.

The three core notions that are crucial throughout both CIDOC CRM and FRBR₀₀ are the following:

- the association between any thing in the world and the various appellations that refer to that thing (e.g., personal name ‘Emily Dickinson’, ISBN ‘978-1-58465-674-6’, Dewey decimal number ‘669.096’, etc.);

- the association between any thing in the world and the various categories to which that thing belongs (e.g., the photograph category, the short novel category, the symphony category, etc.);

- and the association between any thing in the world and the various literals that serve to store, in a given system, unstructured or loosely structured information about that thing (e.g., ‘ix-[86] plates in colour’, ‘Includes bibliographical references.’, etc.).

Here are the three properties that express these three core notions:

E1 CRM Entity.¹⁴ P1 is identified by (identifies): E41 Appellation;

¹³ CIDOC CRM is a model that has been developed since 1994 first by the ICOM CIDOC (International Council of Museums, Comité international pour la documentation), then by the CIDOC CRM SIG (Special Interest Group). ‘CRM’ stands for: ‘Conceptual Reference Model’. Version 6 was published in 2015. All versions are available from http://www.cidoc-crm.org/official_release_cidoc.html. CIDOC CRM also is available as ISO standard 21127:2014.

¹⁴ E1 CRM Entity is the generic class that encompasses all classes within the universe of discourse of CIDOC CRM, and therefore of FRBR₀₀ as well.

E1 CRM Entity. P2 has type (is type of): E55 Type;

and

E1 CRM Entity. P2 has note: E62 String.

Another core notion that was propagated from CIDOC CRM to FRBR_{OO}, although it was absent from FRBR/FRAD/FRSAD, is the notion covered by the E7 Activity class. E7 Activity is central in the CIDOC CRM model: it binds together an intentional action of some kind, the person(s) or group of persons who performed that intentional action, the thing that was affected by that intentional action, the place where that intentional action happened, and the time when it happened (e.g., the fact that Emily Dickinson revised her 1864 poem titled *The Wind begun to knead the grass* at Amherst, Mass., in 1873). That notion was only latent in FRBR/FRAD/FRSAD; as a matter of fact, the Date attribute, which was introduced in FRBR/FRAD/FRSAD for so many entities, is an indication that the original IFLA models were aware that all these entities are not described independently from any of their temporal aspects, but that they do appear in time and evolve over time. Consequently, any occurrence of the Date attribute in FRBR/FRAD/FRSAD was modelled in FRBR_{OO} using either the E7 Activity class, a subclass of it, or other subclasses of the more generic E5 Event class¹⁵ (such as E67 Birth and E69 Death in order to account for the dates associated with a person—an attribute declared in FRAD).

1.6 Conceptual models vs. ontologies

It is important to keep in mind that FRBR_{OO} was developed as a *conceptual model*, rather than an *ontology*, in the sense in which Semantic Web technologies understand that term.¹⁶ Data cannot be encoded directly in FRBR_{OO} such as it stands, and delivered for consumption on the Semantic Web. FRBR_{OO} describes high-level concepts that are likely to be expressed in a variety of RDF structures. Those structures could be more or less detailed, depending on the needs that the data expressed in such structures is meant to meet. Many distinct ontologies could therefore translate the high-level FRBR_{OO} concepts into statements available on the Semantic Web.¹⁷ This ‘Babel effect’ would be tempered by the fact that FRBR_{OO} would remain the common conceptualization above all those ontologies.

A serious difficulty arises from the fact that RDF is unable to account simply for some aspects of the formalism used in both CIDOC CRM and FRBR_{OO}. In both models, some properties are typed, i.e., they are connected with instances of the E55 Type class that serve to specify their meaning. For example, the link between an activity and the person who performed that activity is modelled through the *P14 was carried out by (performed)* property. That property is typed: it has a property of its own, namely *P14.1 in the role of*, which serves to associate the relation between an activity and an actor with any role that an actor can play with regard to that activity (e.g., a person can perform an instance of F28 Expression Creation as an author, as an editor, as a translator, etc.).¹⁸ RDF is just unable to do that.¹⁹ As a consequence, any ontology that would aim at rendering the FRBR_{OO} concepts as RDF

¹⁵ E5 Event is the superclass of E7 Activity.

¹⁶ In the 1990s, the term ‘ontology’ was still understood as referring to any ‘explicit specification of a conceptualization’, as Tom Gruber put it, and did not imply the use of a particular language. Since 2006 and the buzz around Linked Data, the term tends to be used in a narrower sense, exclusively implying RDF implementations.

¹⁷ The National Library of France has begun to develop an ontology based on FRBR_{OO} in order to express its metadata at the finest possible level of granularity as RDF triples.

¹⁸ This is equivalent to the function performed by relator codes in MARC formats.

¹⁹ Neither the Open Metadata Registry representation of FRBR_{OO} as RDF, nor the University of Erlangen-Nuremberg representation of FRBR_{OO} as OWL provides the possibility to type properties, which is extremely problematic for any implementation project.

statements would have to choose between two possible ways of dealing with that difficulty: it would have to either declare explicitly all subproperties of a typed property that correspond to all the types it can have (e.g., ‘to act as author in an expression creation’, ‘to act as editor in an expression creation’, ‘to act as translator in an expression creation’, etc.), which is extremely cumbersome and unsatisfactory; or deal with such properties as classes, which is a better solution but results in a huge increase in the number of RDF statements. This is a real problem, which can have consequences on the adoption of CIDOC CRM and FRBR_{OO} in the Semantic Web environment.

2 HOW FRBR_{OO} MODELS THE MAIN STRUCTURES OF FRBR/FRAD/FRSAD

2.1 ‘Work is realized through Expression’

This is the first ‘structural relationship’ defined between two classes from the FRBR Group 1 of entities. It reflects the notion that any creation of the human mind can be understood as a collection of concepts conveyed through a collection of signs. But it also reflects the notion that distinct collections of signs can be regarded as conveying collections of concepts that are deemed to be similar enough to have the same functionality, to the point that the differences between them are neglected and they are declared to be just ‘one’ collection of concepts (e.g., an original text is a distinct collection of linguistic signs; a translation of that original text also is a distinct collection of linguistic signs; inevitably, at the level of details, there *must* be differences between the concepts conveyed by these two distinct collections of linguistic signs; but in practice, these differences are overlooked when it comes to performing the functionality of allowing end-users to have access to the main concepts conveyed by the original text, and both the original text and the translation are, by convention, regarded as conveying the ‘same’ collection of concepts).

There is, therefore, a semantic duality in this first ‘structural relationship’: any collection of signs conveys a distinct collection of concepts, but several collections of signs also can be said to convey collections of concepts that are ‘similar’ enough to be regarded as one and the ‘same’ collection of concepts.

FRBR_{OO} accounts for both this ambiguity and duality. On the one hand, FRBR_{OO} declares a very general property:

```
F1 Work. R3 is realised in (realises): F22 Self-Contained Expression,
```

which is just as vague and ambiguous as the ‘Work is realized through Expression’ relationship in the original FRBR model. This property is typed, which makes it possible to provide more details about the way an expression realises a work (e.g., a given expression can be said to perform the role of ‘original text’ with regard to a given work). On the other hand, FRBR_{OO} also enables one to distinguish between the notion of ‘work’ understood as the collection of concepts uniquely conveyed by one specific expression, and the fuzzier notion of ‘work’ understood as a cloud of various collections of concepts which, although distinct, are nevertheless regarded as ‘similar’ enough for the differences between them to be held to be negligible. These two notions are reflected in the two following properties:

```
F14 Individual Work. R9 is realised in (realises): F22 Self-Contained Expression,
```

and:

```
F15 Complex Work. R10 has member (is member of): F1 Work.
```

These properties introduce two important subclasses of F1 Work, namely: F14 Individual Work (reflecting a ‘narrow’ understanding of the work class as the collection of concepts conveyed by one and only one collection of signs), and F15 Complex Work (reflecting a more diffuse understanding of the work class as a vague cloud of collections of concepts).

At this point, the audience/readers are surely entitled to ask themselves: Are such distinctions needed? Perhaps in the future there will be implementations in which it will be deemed helpful to have them at hand. But in simpler implementations such as needed today, it is acceptable to just use the generic *R3 is realised in (realises)* property from F1 Work to F22 Self-Contained Expression. Although the number of classes and properties declared in FRBR_{OO} may seem impressive and deterring, it is quite permissible to use *only* those that are needed, and leave all the other ones aside—*until* they are needed. There is no obligation to use *all* of FRBR_{OO} at once.

2.2 ‘Expression is embodied in Manifestation’

This is the second ‘structural relationship’ from the original FRBR model. It means that a collection of signs is infixed on physical carriers which are accounted for as elements within a set, in the mathematical sense of the term; and mathematical sets may count only one element—that is to say: a collection of signs can be infixed on a unique physical carrier (e.g., a manuscript), or on several physical carriers that share the same characteristics (e.g., all the exemplars of a publication).

Once again, there is a certain amount of ambiguity in this relationship, such as it is declared in the definition of the original FRBR model. It is meant to account for three distinct notions:

- the relationship between a unique physical object, such as a manuscript, and the signs that are infixed on it;
- the relationship between authorial content and a publication (e.g., the relationship between Dan Brown’s English text of *Da Vinci Code* and a publication of *Da Vinci Code*);
- and the relationship between the overall content of a publication, such as a publisher determined it, and that publication (e.g., the relationship between the complete collection of signs embodied in a publication of *Da Vinci Code*, including cover art and all paratexts that were not authored by Dan Brown, and that publication of *Da Vinci Code*).

For the sake of accuracy, FRBR_{OO} carefully distinguishes between these three notions.

Manuscripts and other unique carriers of expressions are modelled as instances of the F4 Manifestation Singleton class. All instances of F4 Manifestation Singleton are physical things, but the term ‘Singleton’ in the label of that class is a reminder that the Manifestation entity was originally defined, in the FRBR model, as a set, in the mathematical sense of the term. The relationship between an instance of F4 Manifestation Singleton and its content is accounted for through the CIDOC CRM property *P128 carries (is carried by)* (domain: E18 Physical Thing, range: E90 Symbolic Object).

Publications are modelled as instances of the F3 Manifestation Product Type class. F3 Manifestation Product Type is an abstract notion: it corresponds to the common physical characteristics shared by the outcomes of an industrial process (e.g., the fact that all exemplars of the publication identified by ISBN ‘978-1-84408-453-1’ were produced as single printed volumes, made of paper and cardboard, with a hard cover, consisting of 512 pages (twenty of which are numbered in Roman numerals and 492 in Arabic numerals) and eight (unnumbered) leaves of plates, and measuring 24 cm in height).

But publications are not just defined by physical characteristics such as their number of pages, the diameter of the CDs they consist of, or their height. More importantly, they are defined by their content. The overall content of a publication is a distinct collection of signs, and therefore an instance of F22 Self-Contained Expression; but that collection of signs cannot be attributed solely to the author’s (composer’s, film director’s, photographer’s etc.) activity. Indeed, the overall collection of signs found in a given publication is basically the outcome of the publisher’s activity, and it *incorporates* the collection of signs that can be

recognised as authorial content. This is not meant to minimise the importance of authors; this is simply meant to provide a more accurate (or, rather, less inaccurate) account of the publication process, and of the cultural role played by publisher's decisions—a role that the original FRBR model tended to overlook.

The relationships between authorial content, publisher's input and publication are therefore modelled in FRBR₀₀ as follows:

F3 Manifestation Product Type. CLR6 should carry (should be carried by): F24 Publication Expression. P165 incorporates (is incorporated in): F2 Expression.

In this chain of properties, F2 Expression represents the authorial expression, and F24 Publication Expression represents the overall collection of signs determined by the publisher. F24 Publication Expression is a subclass of F22 Self-Contained Expression. In the original FRBR model, an expression is not assumed to exist independently from the work it realises, and the F19 Publication Work class was therefore introduced in FRBR₀₀ in order to account for the collection of concepts realised in the overall collection of signs established by a publisher; but in practice, an implementation of FRBR₀₀ could very well dispense with the F19 Publication Work class, the instantiation of which only needs to be automatically inferred from the instantiation of the F24 Publication Expression class.

Ultimately, the very process of publication is modelled in FRBR₀₀ as a process through which an instance of F24 Publication Expression is created by a publisher; the F30 Publication Event is therefore declared as a subclass of F28 Expression Creation. The *date of publication* notion corresponds to the date assigned, in a given calendar, to the time-span during which an instance of F30 Publication Event occurred:

F30 Publication Event. P4 has time-span (is time-span of): E52 Time-Span. P78 is identified by (identifies): E50 Date.

But this is a very idealistic (and simplistic) view. As a matter of fact, we can never be sure that the date of publication as found on an exemplar of a publication corresponds indeed to the moment when the overall collection of signs found on all exemplars of that publication was actually being established. A date of publication is nothing more than a mere *statement* on a title page, on which we base our *assumptions* regarding the temporal aspects of the publication process. Although it is often regarded as fairly complex, the FRBR₀₀ model remains a(n over?)simplification of reality.

2.3 'Manifestation is exemplified by Item'

In the case of unique physical carriers (e.g., manuscripts), this relationship is simply not needed, since FRBR₀₀ declares F4 Manifestation Singleton as a subclass of the CIDOC CRM E24 Physical Man-Made Thing: this type of manifestation is *not* an abstract notion that needs to be materialised in a physical object, it *is* by nature something *physical*.

In the case of digital online publications, there is no need to bother with *items*, which consist of unpredictable segments of end-users' hard disk drives (or memory sticks, or whatever), nor even with *manifestations*, since there are no *physical* characteristics that *all* exemplars of such publications are presupposed to share. Acquiring an exemplar of a digital online publication amounts to modifying a segment of a digital carrier by copying onto it an instance of F24 Publication Expression. For the sake of completeness, FRBR₀₀ declares a specific class, F53 Material Copy, for the portions of digital carriers that are physically affected by the acquisition of a digital publication; but, in practice, there is no need to instantiate that class: no library catalogue application in the world would need to implement that part of the model. If it is needed at all, it is in the context of digital storage management; but descriptive metadata management does not call for it.

In the case of ‘classical’, analogue publishing, with physical exemplars, the third ‘structural relationship’ from FRBR is rendered in FRBR_{OO} as:

F5 Item. R7 is example of (has example): F3 Manifestation Product Type.

2.4 Group 1-to-Group 2 relationships

FRBR/FRAD declares specific relationships between each Group 1 entity and all Group 2 entities: ‘Work *is created by* (Person, Family, Corporate Body)’, ‘Expression *is realized by* (Person, Family, Corporate Body)’, ‘Manifestation *is produced by* (Person, Family, Corporate Body)’, and ‘Item *is owned by* (Person, Family, Corporate Body)’.

In FRBR_{OO}, there are almost no such direct links from products to agents: most of these notions are modelled through E7 Activity (or, more precisely, through specific subclasses of E7 Activity).

The notion that a work began to exist at some point in time, prior to which it did not exist yet, is modelled through the F27 Work Conception class. However, I would tend to recommend not to use that class in an implementation. F27 Work Conception corresponds actually to the formation of the initial ideas for a new work in someone’s mind. Both in the scholarly world and in library practice, the notion that a work was ‘created’ is generally understood as the fact that a final clean draft of that work was produced and infixed on some (generally unique) carrier. The production of an expression and its simultaneous infixion on a carrier are modelled in FRBR_{OO} through the F28 Expression Creation class. As for any class in CIDOC CRM and its extensions, it is possible to type an instance of F28 Expression Creation; for instance, we can assign the type ‘major original contribution’ to the instance of F28 Expression Creation that resulted in the original English text of *Great Expectations* on which we base our assumption that *Great Expectations* is a work by Dickens. It is therefore recommended to model the relation between a work and its creator through the following chain of properties:

F1 Work. R19i was realised through (realised): F28 Expression Creation (P2 has type (is type of): E55 Type {major original contribution}). P14 was carried out by (performed): E39 Actor (P14.1 in the role of: E55 Type {[relator code]}).

The notion that an expression was realised by someone is modelled as:

F2 Expression. R17i was created by (created): F28 Expression Creation. P14 was carried out by (performed): E39 Actor (P14.1 in the role of: E55 Type {[relator code]}).

As already mentioned above, the notion that a multi-item manifestation (i.e., a publication, no matter how limited its scope may be) is produced is modelled as a specific case of expression creation: an instance of F30 Publication Event (a subclass of F28 Expression Creation) is an activity the outcome of which is an instance of F24 Publication Expression.

The notion that an item is owned by someone can be modelled through the CIDOC CRM property *P51 has former or current owner (is former or current owner of)* E39 Actor, if nothing is known as to the circumstances under which the owner came into possession of the item, or through the CIDOC CRM chain of properties:

F5 Item. P24i changed ownership through (transferred title of): E8 Acquisition. P22 transferred title to (acquired title through): E39 Actor,

if more detailed information (e.g., the date of the acquisition) is available and deemed historically important enough to be worth recording.

In addition, FRBR_{OO} enables one to express the notion that all the exemplars of a publication were produced by a given manufacturer:

F3 Manifestation Product Type. R26i was produced by (produced): F32 Carrier Production Event. P14 was carried out by (performed): E39 Actor;

and the notion that anyone can modify an item at any point in time:

F5 Item. P31i was modified by (modified): E11 Modification. P14 was carried out by (performed): E39 Actor.

These notions were absent from FRBR/FRAD/FRSAD but may prove quite interesting for bibliographic information about rare and valuable library items.

2.5 Authorities

The FRAD model was developed in order to account for information recorded in authority records. It says that any bibliographic entity is referred to through either a name or an identifier, and that names and identifiers form the basis for controlled access points created and modified by agencies, following specific rules.

FRBR₀₀ makes a distinction between the F50 Controlled Access Point class, which corresponds to controlled access points such as they are displayed for end-users, and the F35 Nomen Use Statement class, which corresponds to the formal statement, such as stored in an authority file, that connects a controlled access point to the thing it identifies, and that provides further information about that controlled access point (e.g., the script conversion that was used to create it, the other instances of F35 Nomen Use Statement to which it is related in the context of the authority file, the language in which it is expressed, etc.). Instances of F35 Nomen Use Statement are specified by an instance of the F34 KOS (for: Knowledge Organisation System) class, i.e., by the authority file to which authority records belong. The three main properties that model the authority work performed in libraries are the following:

F35 Nomen Use Statement. R38 refers to thema (is thema of): E1 CRM Entity;

F35 Nomen Use Statement. R37 states as nomen (is stated as nomen in): F12 Nomen [of which F50 Controlled Access Point is an indirect subclass];

F35 Nomen Use Statement. R35 is specified by (specifies): F34 KOS.

2.6 Aboutness

The FRSAD model was developed in order to account for the ‘aboutness’ relationship. It says that a work is about a thema, and that a thema is identified by a nomen. These two relationships are already covered by the following CIDOC CRM chain of properties:

F1 Work.²⁰ P129 is about (is subject of): E1 CRM Entity. P1 is identified by (identifies): F12 Nomen.²¹

The nomen for a given thema can be recorded in a given instance of F34 KOS through an instance of F35 Nomen Use Statement, following the same mechanisms already described above, in the ‘Authorities’ subsection.

²⁰ In CIDOC CRM, the actual domain of P129 is E89 Propositional Object, of which the FRBR₀₀ class F1 Work is a subclass.

²¹ In CIDOC CRM, the actual range of P1 is E41 Appellation, of which F12 Nomen is a subclass.

3 PRESS₀₀

PRESS₀₀ is an extension of FRBR₀₀. It obeys the same formalism as FRBR₀₀ and CIDOC CRM, simply declaring subclasses and subproperties of FRBR₀₀ and CIDOC CRM classes and properties, every time this was felt to be needed in order to account with precision for the description of continuing resources.

Identifiers of PRESS₀₀ classes are introduced by a ‘Z’, and identifiers of PRESS₀₀ properties by a ‘Y’. There is no semantic value attached to these two letters: they were simply chosen because they are the last two letters in the Latin alphabet.

In FRBR₀₀, there already is a specific class for continuing resources, namely F18 Serial Work. PRESS₀₀ simply elaborates on the specific properties of the F18 Serial Work class.

Unsurprisingly, much of the PRESS₀₀ model is devoted to the various ‘accidents’ that can happen during the existence of a continuing resource. Continuations, replacements, splits, and mergers are modelled as instances of the Z1 Serial Transformation class, a class that accounts for the fact that one or more than one continuing resource disappears, while one or more than one continuing resource begins to exist. The Z2 Absorption class accounts for the fact that one or more than one continuing resource ceases to exist as such, but is prolonged as part of one or more than one continuing resource which already existed prior to that event. The Z3 Separation class accounts for the fact one or more than one continuing resource begins to exist as an autonomous publication, but already existed prior to that event as part of another continuing resource which does not stop existing after that event.

The Z12 Issuing Rule class represents a crucial notion in PRESS₀₀. Instances of Z12 Issuing Rule correspond to atomic aspects of the overall issuing policy followed by the publisher of a continuing resource, e.g., its frequency. The issuing policy of a continuing resource is likely to vary during the existence of that continuing resource. Such variations are accounted for through the Z5 Issuing Rule Change class.

4 CONCLUSION AND FURTHER DEVELOPMENTS

The overall structure and the number of classes and properties of CIDOC CRM, FRBR₀₀, and PRESS₀₀ may be felt as impressive, perhaps even frightening. The fact that such ‘simple’ and ‘intuitive’ notions as the date of an object or the author of a work are not modelled via a direct link between a date and an object, or a work and a person, but via lengthy chains of properties that go through the notions of ‘event’ or ‘activity’, may be felt as useless, costly, and pedantic. The fact that typed properties are not easily implemented as RDF triples may be felt as deterring. But it is reality itself that is devilishly complex. Conceptual models like CIDOC CRM, FRBR₀₀, and PRESS₀₀ tend to reduce as much as possible that complexity, without losing any of its vital semantic connections. And it should be kept in mind that there is no obligation to use *all* the classes and *all* the properties they declare: any implementation of these models should only retain such constructs that are required for the objectives of that implementation.

The Working Group on FRBR/CRM dialogue plans to release a ‘core’ version of FRBR₀₀, which will keep only those classes and properties that will be deemed sufficient to account for the most significant aspects of bibliographic information. That ‘core’ version, hopefully, will be easier to read, understand, and use, than the complete version (which of course will also remain available for specific needs).

Besides, the consolidation process of the FRBR/FRAD/FRSAD models is expected to result in a tighter, simpler model which will call for a revision of the FRBR₀₀ model, and the release of its version 3. Such developments are due by 2016 or 2017.