UNIMARC and FRBR – can we have both?

Trond Aalberg
Norwegian University of Science and Technology
Trondheim, Norway

Jan Pisanski and
Maja Žumer
University of Ljubljana
Ljubljana, Slovenia

Meeting:

187 — *Advancing UNIMARC: alignment and innovation* — IFLA UNIMARC Programme (UNIMARC)

Abstract:

With the Functional Requirements for Bibliographic Records (FRBR) model the library community has been introduced to a new conceptual framework for bibliographic data. On one hand, the model builds upon current bibliographic standards and practice, but on the other hand it represents a dramatic change: as it is a model for bibliographic information expressed as entities and relationships. Alignment of the UNIMARC standard with FRBR is important to facilitate the implementation of the model in library systems, the exchange of bibliographic data based on the FRBR model, and the reuse of UNIMARC data as machine-interpretable information for a broad range of semantic-aware applications. In this paper we analyze the capability of UNIMARC to code simple and complex FRBR structures, and discuss different alternatives for coding these and attempt to identify the most appropriate solution.

Background

Projects and prototypes which focus on the extraction of FRBR entities and relationships from MARC-based records have demonstrated many of the possibilities and problems of using MARC formats in this new context. The work, the expression it is realized through, and related persons and corporate bodies, can with some success be inferred by processing existing records, but most interpretations will produce errors because the data was not meant to be processed in this way (Aalberg, 2006, Manguinhas, Freire, Borbinha, 2010). The problem of proper identification of FRBR entities is generally known (e.g., Hegna, Murtomaa, 2002, Hickey, O'Neill, 2005, Pisanski, Žumer, Aalberg, 2009), and is typically caused by cataloguing practice and data requirements that are not aligned with the model.

A different problem, which has received less attention, but is even more fundamental for the alignment of FRBR and MARC standards, is the problem of well-defined representation of more complex entity-relationship structures in MARC. Records describing resources with multiple distinct parts, possibly combined with multiple persons/corporate bodies, subject entries and other related works, are unfortunately commonly represented in ways that are informative and human readable when displayed, but often impossible to interpreted correctly and consistently by software. On one hand this limits the value and usability of these particular records in library systems based on FRBR, on the other hand it is a problem that affects the collection as a whole. The catalogue will be incomplete if we leave out such records and if we attempt to interpret such records we are likely introduce a number of errors. Interpreting a collective title as the single distinct work produces an irrelevant work and ignores the actual works that may be important to end-users. Interpreting alternative titles or titles for related works as distinct content is another source of errors. Even if the number of records that are interpreted incorrectly is low, these errors will typically escalate as we move up the abstraction level. The list of works that different solutions are able to identify for prolific authors such as Agatha Christie, usually includes a high number of incorrectly identified works, and one of the reasons is that the records do not expose the correct FRBR structure for the content.

MARC and FRBR structures

During the early years of MARC, several formats emerged owing to different national requirements and cataloguing practice. Today, UNIMARC and MARC 21 dominate the world of bibliographic information, as libraries seem to acknowledge the advantages of using one of these globally recognized formats rather than national or vendor-specific formats. The ISO 2709 format that defines the syntax and underlying structure of all MARC formats is designed to describe many aspects of a resource using a flat a structure of control fields and data fields. Each data field further has indicators and subfields and records based on the standard may contain rather complex information, but there is no standard formalism for representing a unit of information that describes a set of interrelated entities. However, different methods and practices have been developed within the formalism of the standard to facilitate the coding of more complex bibliographic information in both UNIMARC and MARC 21. Some of these methods are relevant for coding FRBR, as they will give records that can be interpreted consistently as a set of interrelated FRBR entities. Despite an intended interoperability by conversion between these formats, there are differences that are relevant in terms of how the formats can be used to code FRBR structures.

UNIMARC is based on fields that are differently numbered and named from equivalent fields in MARC 21, and has a grouping into blocks that represents a different organization of the information. In UNIMARC we find separate blocks for descriptive information, related titles, and intellectual responsibility. Corresponding information in MARC 21 is found in the main entry fields, title and title-related fields, and added entry fields. Some data fields in UNIMARC are less ambiguous than corresponding data fields in MARC 21. In UNIMARC different fields are always used for names and related titles respectively, whereas e.g. the 700 Added Entry Personal Name in MARC 21 may include a personal name only or a title where the personal name is a part of the title identification. The stricter scope of these fields in UNIMARC also affects how certain FRBR structures are described in the record. The UNIMARC 5xx Related Title block is

only for alternative titles associated with the manifestation being catalogued and will always be a title for the manifestation, the single contained expression or work it realizes. The title of other related works or expressions have to be catalogued using linking fields. A title in a MARC 21 added entry field of MARC 21 would in practice be either an analytical entry title or a title that is either an alternative title for the contained work/expression or a related work/expression. The second indicator can be used to specify that the field is an analytical entry, but the purpose of the field in the context of FRBR is unspecified if there is no value for this indicator.

Furthermore, UNIMARC allows repeatable title proper and statement of responsibility, whereas MARC 21 has non-repeatable subfields for the title statement and relies on ISBD punctuation for further structuring within the fields. Both formats have fields for unformatted contents notes but MARC 21 has an additional field for formatted contents notes with repeatable subfields.

Both formats have linking entry fields for describing related bibliographic resources (either contained or external to the manifestation), but there are inherent differences in the use of these fields. With respect to FRBR, UNIMARC seems to be designed to describe a single contained expression and favours the use of linking entry fields for structured descriptions of multiple contained expressions or related works/expressions. In MARC 21 this can be achieved with the use of added entries and linking entry fields are less relevant for this purpose. Another main difference between the linking fields is the method of embedded linking that is found in UNIMARC. This enables a more flexible and precise description of target entities using any appropriate combination of ordinary data fields rather than the use of the defined set of linking subfields.

Both formats have subfields for controlled relator codes to describe the function of Group 2 entities. In UNIMARC this seems to be more systematically used. In MARC 21 there is an additional field for relator term. MARC 21 has recently been extended with subfields for relationship information in the added entry fields. The $8 field linking subfield is unique for MARC 21 and can be used to link one data field to other specific data fields. There are additionally other minor differences such as whether specific subfields are repeatable or not. Although both formats can include equivalent information about the same manifestation, the differences also have some impact on the description.

Some support for such structures can be found in both UNIMARC and MARC 21, such as the linking fields, use of indicators, fields that may (or may not) represent separate and distinct entities, relator codes etc. However, these features only partly meet the requirements and a systematic approach is needed to identify current possibilities and future requirements for representing the FRBR-model using MARC. Furthermore, even if the formats may support a more explicit coding corresponding to the FRBR model in theory, there is a variety of cataloguing practices, adapted to different requirements and material types (e.g. Mcgrath, Bisko, 2008).

Complex FRBR structures in UNIMARC

In the following we will use a pragmatic approach to analyze the capability of UNIMARC to code complex FRBR structures. First we present a simple structure and then we give some examples on more complex structures, discuss different alternatives for coding these and attempt to identify the most appropriate solution. The underlying formalism

in our analysis is that FRBR models the bibliographic universe in a catalogue as a graph[1] of distinct typed nodes (FRBR entities) with typed links (relationships) between them. Each bibliographic record will always represents a particular part of this universe. The sole requirement in our analysis is that it should be possible to write algorithms that automatically interpret a collection of records and construct a graph of FRBR entities that corresponds to the intended FRBR graph described by the records.

In the FRBR model, all entities are distinct, but interrelated, units of interest that we want to be able to process. The requirements for processing each MARC record as a set of interrelated entities can be summarized in the following questions:

- What is the set of entities described by each record?
- What are the attributes describing each entity?
- What is the identity of each entity?
- How are these entities connected?

The entities

A substantial number of records in a library catalogue will have a rather simple FRBR structure and entities that are implicitly or explicitly described can be inferred either from the tags and codes or by interpreting the model itself. Most records typically describe a single publication that in the FRBR model is considered to be a manifestation. According to the model we know that this manifestation embodies at least one expression, which is the realization of only one work. Persons and other Group 2 entities can fortunately be identified from specific fields and are for this reason relatively easy to discover. Additionally, we can use other tags or codes to identify more distinct works and expressions such as the ones that are described in subject entries.

Unfortunately, this simple default view on a record does not always correspond to the actual FRBR structure that the record represents. Manifestations may embody multiple expressions and there may be other entities in the description that only are related and not contained. Applying the default and simple interpretation that each manifestation embodies a single expression is not only a coarse simplification, but it also means that we do not use the FRBR model as it was intended. Identifying the correct set of FRBR entities in a record, however, is challenging. Contained expressions may be described in contents notes and there are fields that in one context may identify an individual work but not in other.

The attributes

Information in the record, with the exception of fields that describe the record itself, will describe various characteristics of the entities. For the simple structure of a manifestation that embodies a single expression, this is a problem that can be reduced to a mapping task where each field or subfield can be assigned to a specific entity type. The meaning and usage of specific fields may be ambiguous in terms of what entity type they describe, but this is a semantic problem more than a structural one. On the other hand, if the record describes a manifestation that embodies two expressions, it can be a structural problem. If we have one value for language – we may assume that it applies to both expressions. If we have multiple values for language – are both languages valid for

[1] Actually a set of graphs because there may be subparts of the bibliographic universe with no known relationships to the rest.

both expressions or are expressions in different languages? In the latter case, which language is used in which expression?

The identity of entities

Bibliographic information is based on the use of descriptive identification where the units of interest are identified by a set of attributes constructed according to cataloguing rules. This is essentially not very different from the use of key attributes in relational databases, but bibliographic information systems are "open world systems" where it is possible to refer to any known bibliographic entity. Unfortunately, even if we have been able to identify the correct set of entities and have mapped all attribute values to the correct entity type we often do not have sufficient information to identify entities consistently in an automatic way. The descriptive identification used in cataloguing is intended for humans who are able to apply contextual interpretation and make the best use of whatever information they find. If we identify the work of a translated expression, using author, uniform title and form of work – these are the key values that have to match the work appearing in other contexts as well. Unfortunately, this is not always the case. As a contained work it may be identified by original title (as the uniform title), but as a subject or related work entry it may be catalogued with title it generally is known under in the language of the catalogue.

The connections

The final structural challenge is the relationships. This is both a question of relationship type and identification of the entity endpoints that the relationship connects. Information about the type of a relationship can be implicit or can be deduced from the field tag, indicator or specific subfields such as the relator code in the case of Group 2 entities. For a manifestation that embodies a single expression, the type and endpoints of the relationships are implicit, but again this is not the case if we have more complex structures. A manifestation that embodies two expressions and includes a single personal name subject entry an ambiguous structure. The type of the relationship is known, but we do not know the endpoints. Is the subject entry related to both works or only to one of the works? Making the right connection between entities based on type is only possible if there is a single entity on one side of the relationship. The problem of connecting entities is even more problematic if the relationship type is unknown, which is the case if e.g. relator codes are not used.

Manifestations with single expressions

Manifestations with a single expression are the most trivial case of a bibliographic record and books that contain a single novel are good examples. This is also the most common structure in the records of a library catalogue. In theory, many of these publications are in reality more complex, but even in a FRBR-context we would usually neglect cover art, short introductions and other expressions of minor importance because we have to make a balanced choice of what to manage as distinct entities. Adding entities only for the sake of the model is not justified, as it only will introduce noise in the library system.

Because manifestations embody expressions and each expression always is the realization of one (and only one) work, we can infer that these entities exist. If there is only one expression and consequently one work, the main challenge is to find out which fields and subfields describe which entity as well as to select appropriate key attributes for these entities. Adding Group 2 entities to the picture does not necessarily cause any

problems. Distinct Group 2 entities in UNIMARC are well supported by the use of the 7--Intellectual responsibility block, the use of authority files, and the use of relator codes. The relator code is quite significant as it is the main source of information about how a Group 2 entity is related to a Group 1 entity. If the relator code in a 700 Personal Name entry is 070, we know that this person is the author. In FRBR terminology this will be a *has created* relationship to a work, and because there is only one work, the link is rather obvious. If the relator code in a Personal Name entry is 730 we know that the person is the translator and consequently the link should be to the expression.

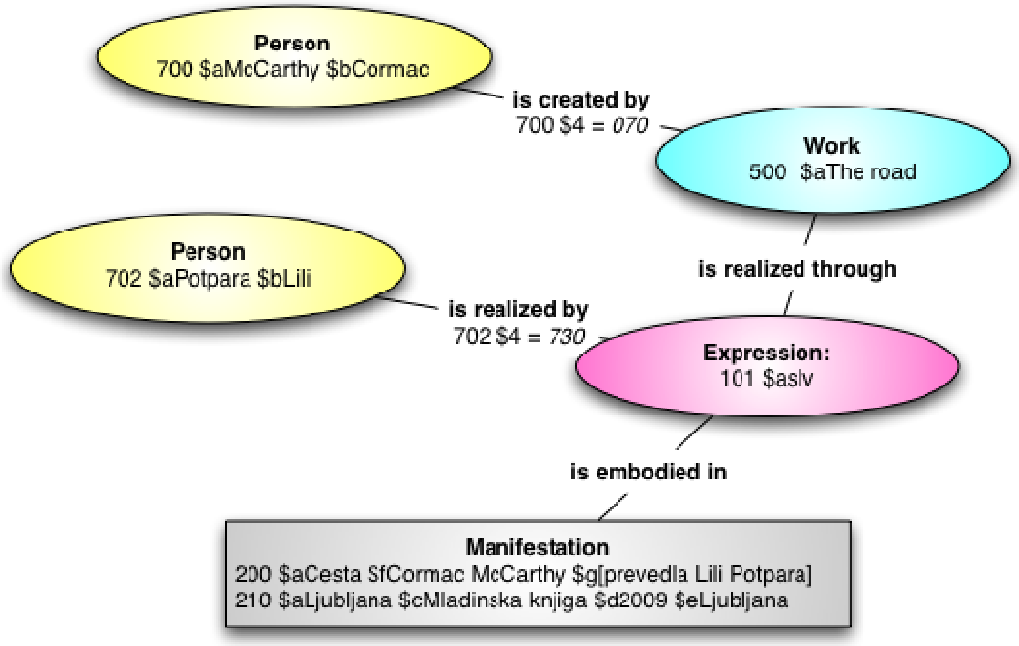

Figure 1: A manifestation that embodies one expression, with examples of significant information from the record used to identify entities and create relationships (Slovenian National Bibliography). Spaces between subfields are added for readability.

The example in Figure 1 illustrates this case by showing a Slovenian translation of *The Road* by the American writer *Cormac McCarthy*. Even with the simple sequence of fields of the MARC format, the structure can be deduced directly from the record. Problems related to such simple structures are mainly caused by insufficiently described and identified entities and lack of relator codes.

Manifestations with two or more expressions

Manifestations may also embody more than one expression: component parts, analytic pieces or constituent units. This is a form of aggregation that is orthogonal to the notion of a manifestation having (sub-) manifestations as parts, or a work that has other works as parts. Such aggregations are often described in the record, but the structure of interrelated entities can be ambiguous due to the limitations of the MARC. The example in Figure 2 shows a publication containing two of Agatha Christie´s novels in Portuguese by two different translators. In this record both the repeated subfields in 200 Title statement and the repeated 304 note indicate that there are two works. Because there is only one author in the record, we can assume that this author has created both works. Information such as language can be related to the expressions applying the same

assumption that the one language value relates to both. The translators are well identified, but there is no explicit link stating which 702 Person entry relates to which expression. This lack of explicit links between entities may not seem to be important at first glance, but it exemplifies a major problem. If the two novels were by different authors, we would have the same kind of problem for the *is created by* relationship which would result in insufficient or erroneous identification of the works.

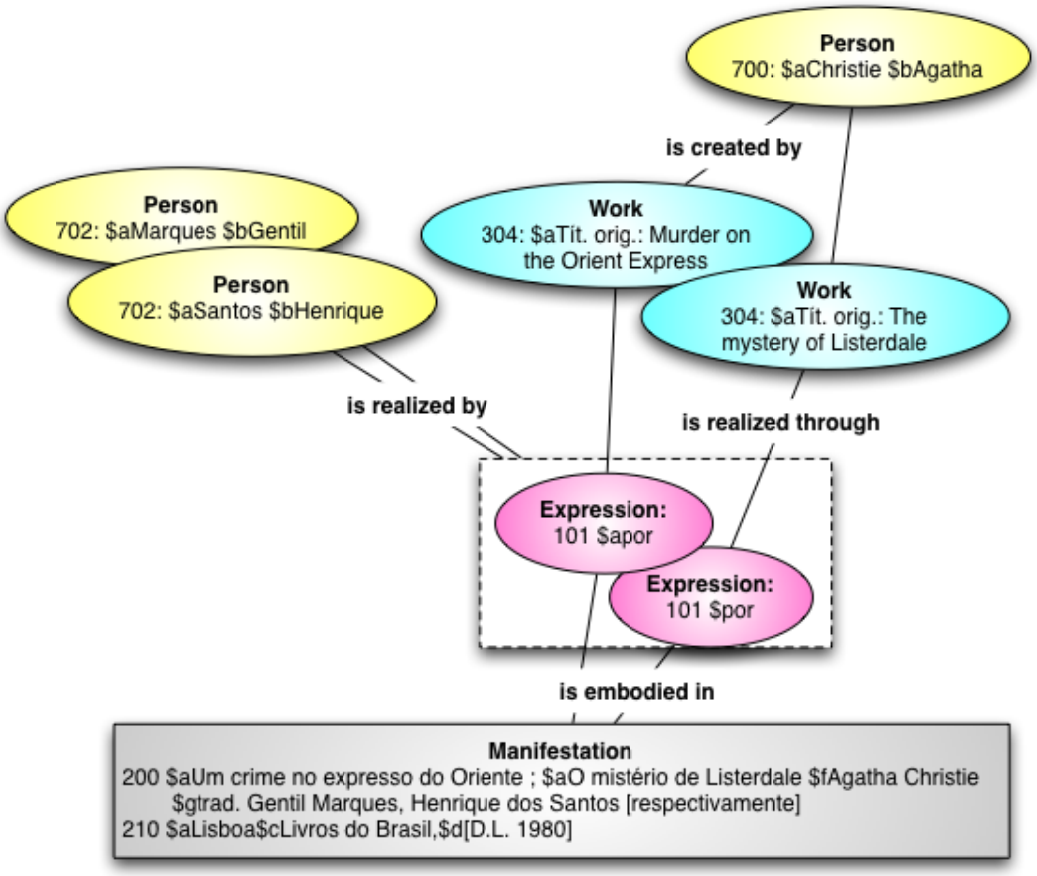

Figure 2: Example showing a manifestation that embodies two expressions (The Portuguese Union Catalogue).

Examples in Table 1 show some records with this kind of structure. The first example describes a collection of crime novels by different authors. The content can partly be identified from the repeatable subfields in the title statement, but the listing of original titles in a 304 notes field is a better source as it includes all contained works. In theory it could be possible to use the sequence titles and personal name entries to link the author to the correct work, but as shown, the order of 70X fields does not follow the sequence of titles in the note.

The second example describes a collection of crime novels by Agatha Christie. This record uses a linking field with the embedded field method that contains all titles and the author using 200 and 700 fields respectively. Additionally it has 327 contents note with the same set of titles. In this case there are multiple titles in the linking entry field that clearly are related to one and the same author.

The third example describes a Swedish translation of the *Lord of the Rings* by J.R.R Tolkien. In this case there are multiple linking entries, one for each expression/work.

The 500 fields contain the original title suitable for work identification, the 200 fields would in this case serve well as titles of the expressions and the inclusion of the author in all entries makes the structure unambiguous. The translator, however, is only mentioned in the statement of responsibility and will be difficult to utilize as part of the expression identification.

101: $apor 200: $aOs crimes da rua morgue $fEdgar A. Poe $gTrad. João Costa$aO cachimbo de Maigret $fGeorges Simennon $gTrad. J. Lima da Costa 210: $aLisboa$cLivros do Brasil,$d1997 304 $aTít. orig.: The murders in the Rue Morgue ; Murder on the links ; La pipe de Maigret ; The bronze door and other stories 700 $aPoe $bEdgar Allan $f1809-1849 701 $aChandler $bRaymond $f1888-1959$4070 701 $aChristie $bAgatha $f1890-1976$4070 701 $aSimenon $bGeorges $f1903-1989$4070 702 $aCosta $bJ. Lima da $4730 702 $aCosta $bJoão Alves da $f1948-$4730 702 $aRodrigues $bFernanda Pinto$4730
101 $aeng 200 $a1950s omnibus $fAgatha Christie 210 $aLondon $cHarperCollins $d2006 327 $aThey came to Baghdad $aDestination unknown $aOrdeal by innocence $aThe pale horse 423 $1200 $aThey came to Baghdad $aDestination unknown $aOrdeal by innocence $aThe pale Horse $1700 $aChristie $bAgatha $f1890-1976 $4070 700 $aChristie $bAgatha $f1890-1976 $4070
101 $aswe 200 $aHärskarringen $fJ. R. R. Tolkien $göversättning av Åke Ohlmarks $gillustrationer av Alan Lee 210 $aStockholm $cNorstedts $d2002 $e[Ljubljana] $gMladinska knjiga 300 $aPrevod dela: The lord of the rings 327 $aSagan om ringen = The fellowship of the ring $aSagan om de två tornen = The two towers $aSagan om konungens återkomst = The return of the king 423 $1200 $aSagan om ringen $1500 $aThe fellowship of the ring $1700 $aTolkien $bJohn Ronald Reuel $4070 423 $1200 $aSagan om de två tornen $1500 $aThe two towers $1700 $aTolkien $bJohn Ronald Reuel $4070 423 $1200 $aSagan om konungens återkomst $1500 $aThe return of the king $1700 1 $aTolkien $bJohn Ronald Reuel $4070 500 $aThe lord of the rings 700 $aTolkien $bJohn Ronald Reuel $4070

Table 1: Different manifestations that embody two or more expressions (the first two examples are from the Portuguese union catalogue, the last is from the Slovenian national bibliography).

The third example additionally illustrates another structural issue. The person in the 700 entry can be related to the trilogy "Lord of the Rings" (as a whole) because it makes sense to state that J. R. R. Tolkien was the creator of "Lord of the Rings", which is the work commonly referred to. In the other two examples such a statement would be utterly wrong. Creating a work for the title *1950s omnibus* and associating Agatha Christie as the creator of this work is a wrong statement – she is merely the author of the novels it contains. For the first example it would be equally wrong to associate all authors as creators of the collection. The presence of a 500 uniform title in the last example can indicate that we should consider the whole as a work, but this is an assumption that only works for these examples because the last one is a translation.

Listing the expressions/works using either of the techniques expresses the actual "content" structure, but the use of linking entry fields is by far the most precise method as it enables the inclusion of all information that is needed for the identification of each expression and work. The use of multiple fields is the most appropriate as it allows for differentiation between the entries and is based on a structure that corresponds to the natural grouping of information. One pitfall in the use of any one of these techniques is that there may be information in other fields that applies to all contained expressions, such as language. If there is only one language value it is natural to assume that it is valid for all expression. Two or more language values will only make sense if they are valid for all expressions. Adding language to the embedded links could be a solution for the somewhat rare case of expressions in different languages, but the problem exists for other information as well, such as for subject entries. In MARC 21 the coupling between fields is supported by the use of $8 field linking, which is a potential solution to the problem in UNIMARC as well.

Relationships to other works

The FRBR model defines a variety of relationships between Group 1 entities in addition to the basic *is embodied in* and *is realized through*. Reasons for adding such relationships to a record are to give a better description of the catalogued resource and to enable users to find related works. In the context of FRBR such relationships are even more important because they give a mechanism for supporting new types of user interactions such as exploration and discovery based browsing and searching. The linking entry fields are the main mechanism for creating such relationships in UNIMARC, but just a few of the relationship types described in FRBR are specifically supported. All other relationships will have to be catalogued using the generic 488 linking entry, possibly with the use of a 311 Notes). The latter is unfortunately not a very good mechanism. It is on one hand a textual uncontrolled description intended for display only. On the other hand it will fail if there are different linking entry fields and the note is only valid for one specific field. The embedded linking method is generally superior to the standard subfield technique because it enables the detailed identification of an entity using the exact same set of data fields that would be used if the entity was described using the ordinary data fields in a record. This technique allows for quite precise identification of works, expressions and manifestations but there is a need for a better mechanism to type the relationships.

Even if the linking fields and the use of embedded linking seems well suited for expressing relationships, there is a major challenge in this technique when it comes to manifestations that embody multiple expressions. If all contained expressions are described using linking entry data fields and there is one or more linking entry fields that describe work-to-work or expression-to-expression relationships, there is no way to tell which of the contained expressions the relationship is for. We may assume that it is for all, but this simplification is deemed to produce incorrect interpretations.

Another important type of relationship is that of works having Group 1 or 2 entities as subjects. In UNIMARC there are specific data fields for person, corporate body or family name as subject and for title as subject. The 604 field is for name and title and is based on the embedded field method. Again, there is a problem if the contained expressions (or rather, the corresponding works) have different subjects, such as in the case of essay collections etc. There is no specified way of indicating which subject belongs to which work.

Figure 3 shows the grouping that can be achieved by the use of linking entry fields and the embedded fields method. The various "ordinary" fields describe the manifestation and any information that is valid for all contained expressions/works. Two linking fields are used to describe two different embodied expressions and the works they are realizations of, as well as persons related to specific expressions and works. Because of the linking field tag we know that these are contained expressions related to the manifestation described in the record (using e.g. 423 to follow the convention in the earlier examples). Additional linking entries are for one related work and one subject. The connection between these entries and the entries for the contained expressions, however, are not described and cannot be interpreted precisely.

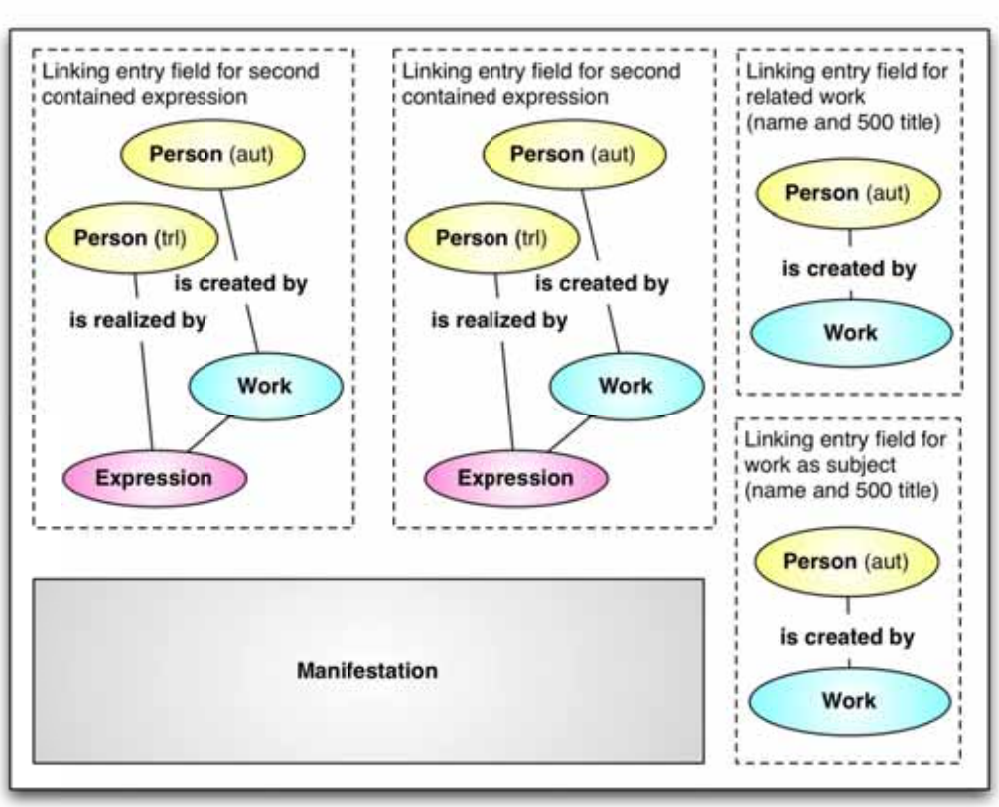

Figure 3: Example of using linking entry fields and the embedded fields method.

Conclusions

In this paper we have discussed and explored some of the problems that exist when coding FRBR in UNIMARC. The problem of identity and attribute assignment mainly originates in cataloguing practice (and cataloguing rules) and is caused by insufficient information (such as lack of titles that properly identifies the works) or data fields that are ambiguous in the context of FRBR. Some of these problems are addressed in RDA and will hopefully be followed up by changes to the data fields in the MARC formats.

The problem of defining and describing a set of entities with correctly linked relationships in UNIMARC is a different issue. The current format is sufficient for describing manifestations with a single expression, but for more complex structures the examples we have shown indicate that there is a variety of solutions in use today, based

on either contents notes, repeatable subfields in the title statement and linking entry fields. The first two use descriptive information, which is not sufficient because they are difficult to process and will result in poor identification of the entities. Fortunately, the use of linking entry fields seems to be an appropriate solution to solve many of the problems related to the more complex structures. It can be seen as a workaround that adds a new depth to the basic MARC structure that seems to fit the FRBR model. This technique mimics the use of multiple linked records, but with the benefit of keeping all information within one record. With the embedded fields method we can include all information needed to describe and identify all entities related to the "branch" of the contained expression.

A particular problem that we have addressed is the ambiguity that occurs when information in a specific field may relate to multiple contained expressions or works, or there are related works/expression that only are valid for specific contained expressions/works. For content related data fields this can be solved by rules for what to include in the embedded fields of the linking entry, but for relationships between works in separate linking entries this solution is impossible. A field linking mechanism comparable to the subfield $8 in MARC 21 could solve the problem, either by adding a subfield for this purpose (that does not conflict with other embedded subfields) or by using the unused first indicator for this purpose (only applicable for connecting linking entry fields).

If we want the linking entry fields to fully support the relationships of FRBR, there is an additional need for either specifically designated linking field entries, or some mechanism to express the type of the relationship by adding a type identifier. The first would yield a large number of linking entry fields, the latter requires some specifically designated subfield for this purpose which is difficult to introduce because it may conflict with the embedded subfields.

We have only described cases related to multiple expressions embodied in a manifestation, but we are aware that there are still more types of structures to explore such as manifestations, expressions and works having multiple parts, amendments and others. Some problems can be solved with linking entry fields, but there may be cases where other solutions are required.

Finally we will emphasise some more general recommendations for the future development of UNIMARC:

A shift in perspective from presentation to processing

Implementing FRBR in a library system means more than simply changing the format or the underlying data model. FRBR represents a logical model of the domain of discourse, and implementations are depending on exchange formats and storage models that reflect the model. The actual requirements for bibliographic records in the future are largely depending on what we want the data to be capable of. Recent developments in library user interfaces have explored the use of FRBR-like groupings of the result set and we are increasingly becoming aware of the benefits of the model. A major challenge for the future is to define the real requirements for the data in terms of the processing and querying we want the data to support, not in terms of how the record should be displayed.

Revise the standard

The FRBR model has influenced RDA, and will consequently influence on the standards for coding bibliographic records in the future. Some changes have been made to the MARC 21 format, and it is important that UNIMARC is able to modernize the format accordingly. Better support for FRBR does not necessarily introduce backwards incompatibility, but lack of support or FRBR surely introduces forward incompatibility.

Define best practice for coding FRBR

Coding FRBR in MARC is not necessarily a question of adding more information to the record and introducing additional cataloguing cost. The main problem is the use of a variety of solutions in different catalogues as well as within the same catalogue. Some cataloguing practices are deemed to be less useful than other in future FRBR implementations. In the current standard there are no examples that relate to FRBR and there is a general lack of guidelines or best practice examples for coding information in a FRBR compatible way. Even without changing the format and without changing the cataloguing rules and practice, it is possible to suggest best practise solutions for a number of common cases.

Proper versioning

Most library catalogues contain records that have been created and aggregated over many years or even decades. The cataloguing practice and formats may have changed over time, which causes minor or major differences. Migrations from national formats to internationally recognized formats may syntactically conform to the target standards, but there are always remains of previous practice that are impossible to migrate. Changing to a different use of the format adapted for the FRBR model is a change that should be captured using some kind of versioning so that we are able to identify these records. The current UNIMARC 005 Version Identifier is simply a date for the last time of update, and does not indicate anything else. If a library catalogue is to contain some records that are tailored to specific ways of expressing and processing FRBR, whereas others are not, a minimum requirement is that we are able to distinguish between these using a version identifier.

References

Aalberg, T. (2006). A process and tool for the conversion of MARC records to a normalized FRBR implementation. *ICADL 2006*, *LNCS*, vol. 4312, 283-292.

Hegna, K. and Murtomaa, E. (2002). Data mining MARC to find: FRBR? *68th IFLA Council and General Conference*, *18-274 August 2002*, *Glasgow*, *Scotland*. http://archive.ifla.org/IV/ifla68/papers/053-133e.pdf

Hickey, T. and O'Neill, E. (2005). FRBRizing OCLC's WorldCat. *Cataloging & Classification Quarterly*, 39 (3/4), 239-251.

IFLA International Federation of Library Associations and Institutions. Study Group on the Functional Requirements for Bibliographic Records (1998). *Functional Requirements for Bibliographic Records: final report*. Munich, Germany: KG Saur

Manguinhas, H., Freire N. and Borbinha, J. (2010). FRBRization of MARC records of multiple catalogs. *Proceedings of ACM/IEEE Joint Conference on Digital Libraries*. 225-234.

Mcgrath, K. and Bisko, L. (2008). Identifying FRBR work-level data in MARC bibliographic records for moving images. *Code4Lib Journal*. 1 (5), December 2008. http://journal.code4lib.org/articles/775

Pisanski, J., Žumer, M. and Aalberg, T. (2009). Frbrisation: towards a bright new future for national bibliographies. *World Library and Information Congress: 75th IFLA General Conference and Council, 23-27 August 2009, Milan, Italy*. http://www.ifla.org/files/hq/papers/ifla75/77-pisanski-en.pdf