



UNIMARC y FRBR: ¿podemos tener ambos?

Trond Aalberg

Universidad de Ciencia y Tecnología
Trondheim, Noruega

**Jan Pisanski and
Maja Žumer**

Universidad de Ljubljana
Ljubljana, Eslovenia

Traducción de Cayetano

Hernández Muñiz y Lourdes Alonso Viana

Biblioteca Nacional de España

Meeting:

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

Resumen:

Con el modelo Requisitos Funcionales de los Registros Bibliográficos (FRBR), la comunidad bibliotecaria se ha introducido en un nuevo marco conceptual de los registros bibliográficos. Por un lado, el modelo está construido sobre los estándares y prácticas bibliográficas actuales; pero, por otro, representa un dramático cambio, ya que es un modelo donde la información bibliográfica se expresa en términos de entidades y relaciones. Es importante ajustar el estándar de UNIMARC con FRBR para facilitar la implementación del modelo en los sistemas bibliotecarios, el intercambio de datos bibliográficos basados en FRBR, y la reutilización de los datos de UNIMARC como información interpretable por máquina para un amplio espectro de aplicaciones que tienen en cuenta la semántica. En esta ponencia analizamos la capacidad de UNIMARC de codificar estructuras FRBR simples y complejas, y discutimos diferentes alternativas de codificación así como tratamos de identificar la solución más apropiada.

Antecedentes

Los proyectos y prototipos centrados en la extracción de entidades y relaciones FRBR a partir de registros basados en MARC han demostrado muchas de las posibilidades y problemas de usar los formatos MARC en este nuevo contexto. La obra, la expresión a través de la cual se realiza, y las personas y entidades corporativas relacionadas pueden inferirse con cierto éxito al procesar los registros existentes, pero la mayor parte de las

interpretaciones producirán errores porque no se pretendía que los datos se procesaran de esta manera (Aalberg, 2006, Manguinhas, Freire, Borbinha, 2010). El problema de la identificación adecuada de las entidades FRBR es en general conocido (por ejemplo, Hegna, Murtomaa, 2002, Hickey, O'Neill, 2005, Pisanski, Zumer, Aalberg, 2009), y está causado normalmente por la práctica de la catalogación y de los requisitos de los datos que no se ajustan al modelo.

Un problema diferente, que ha recibido menos atención, pero es incluso más fundamental para la coordinación de las normas FRBR y MARC, es el problema de la representación bien definida de estructuras basadas en relaciones entre entidades más complejas en MARC. Los registros que describen recursos con partes múltiples, posiblemente combinadas con múltiples personas/entidades corporativas, entradas de materia y otras obras relacionadas, por desgracia se representan comúnmente de manera que son informativos y legibles para las personas, pero a menudo imposibles de interpretar correcta y consistentemente por el "software". Por un lado, esto limita el valor y la usabilidad de estos registros particulares en sistemas bibliotecarios basados en FRBR, y por otro, es un problema que afecta a la colección como conjunto. El catálogo estará incompleto si excluimos tales registros, y si intentamos interpretarlos es probable que introduzcamos errores. Interpretar un título colectivo como una única obra distinta produce un trabajo irrelevante e ignora las obras reales que pueden ser importantes para los usuarios. Interpretar variantes de título o títulos para obras relacionadas como un contenido distinto es otra fuente de errores. Incluso si el número de registros que se interpretan incorrectamente es bajo, estos errores normalmente aumentarán a medida que subimos en el nivel de abstracción. La lista de obras de autores prolíficos, como Agatha Christie, que diferentes soluciones pueden identificar, normalmente incluye un alto número de obras identificadas incorrectamente, y una de las razones es que los registros no presentan la estructura FRBR correcta para el contenido.

Estructuras MARC y FRBR

Durante los primeros años de MARC, surgieron varios formatos debido a los diferentes requisitos y prácticas de catalogación nacionales. Hoy, UNIMARC y MARC 21 dominan el mundo de la información bibliográfica, puesto que las bibliotecas parecen reconocer las ventajas de usar uno de estos formatos reconocidos globalmente frente a formatos nacionales o específicos de un proveedor. El formato ISO 2709 que define la sintaxis y estructura subyacente de todos los formatos MARC está diseñado para describir muchos aspectos de un recurso usando una estructura plana de campos de control y campos de datos. Cada campo de datos además tiene indicadores y subcampos, y los registros basados en la norma pueden contener información bastante compleja, pero no hay un formalismo normalizado para representar una unidad de información que describa un conjunto de entidades relacionadas. Sin embargo, se han desarrollado distintos métodos y prácticas dentro del formalismo de la norma para facilitar la codificación de información bibliográfica más compleja tanto en UNIMARC como en MARC 21. Algunos de estos métodos son relevantes para codificar FRBR, puesto que darán registros que puedan ser interpretados de forma consistente como un conjunto de entidades FRBR interrelacionadas. A pesar de una pretendida interoperabilidad por conversión entre estos formatos, hay diferencias que son relevantes en términos de cómo los formatos pueden usarse para codificar estructuras FRBR.

UNIMARC se basa en campos que están numerados y nombrados de forma diferente a los campos equivalentes en MARC 21, y tiene una agrupación en bloques que representa una organización diferente de la información. En UNIMARC encontramos bloques separados para la información descriptiva, títulos relacionados, y responsabilidad intelectual. La información correspondiente en MARC 21 se halla en los principales campos de entrada, campos de título y título relacionado, y campos de entrada secundaria. Algunos campos de datos en UNIMARC son menos ambiguos que los correspondientes campos de datos en MARC 21. En UNIMARC se usan siempre campos diferentes para nombres y títulos relacionados respectivamente, mientras que, por ejemplo, el 700 Entrada Secundaria de Nombre Personal en MARC 21 puede incluir un nombre personal solo o un título donde el nombre personal es una parte de la identificación del título. El alcance más estricto de estos campos en UNIMARC también afecta a cómo ciertas estructuras FRBR son descritas en el registro. El bloque de Título Relacionado UNIMARC 5XX sólo se usa para variantes de título asociados con la manifestación que se está catalogando y siempre será un título para la manifestación, la única expresión u obra contenida que realiza. El título de otras obras o expresiones relacionadas tiene que catalogarse usando campos de relación. Un título en un campo de entrada secundaria MARC 21 en la práctica sería o un título de entrada analítica o un título que es, o bien una variante de título para la obra/expresión contenida, o bien una obra/expresión relacionada. El segundo indicador puede usarse para especificar que el campo es una entrada analítica, pero el propósito del campo en el contexto de FRBR no se especifica si no hay valor para este indicador.

Además, UNIMARC permite que el título propio y la mención de responsabilidad sea repetibles, mientras que MARC 21 tiene subcampos no repetibles para la mención de título y confía en la puntuación ISBD para una estructuración adicional en los campos. Ambos formatos tienen campos para notas de contenido no formateadas, pero MARC 21 tiene un campo adicional para notas de contenido formateadas con subcampos repetibles.

Ambos formatos tienen campos de entradas relacionadas para describir recursos bibliográficos relacionados (o contenidos en o externos a la manifestación), pero hay diferencias inherentes en el uso de estos campos. Con respecto a FRBR, UNIMARC parece diseñado para describir una expresión única contenida y favorece el uso de campos de entradas relacionadas para descripciones estructuradas de múltiples expresiones u obras/expresiones relacionadas. En MARC 21 esto puede lograrse con el uso de entradas secundarias, y los campos de entradas relacionadas son menos relevantes para este fin. Otra diferencia principal entre los campos de relación es el método de relación integrada que encontramos en UNIMARC. Esto permite una descripción más flexible y precisa de las entidades usando cualquier combinación apropiada de campos de datos corrientes antes que el uso del conjunto definido de subcampos relacionados.

Ambos formatos tienen subcampos para códigos de relación controlados para describir la función de entidades del Grupo 2. En UNIMARC esto parece usarse más sistemáticamente. En MARC 21 hay un campo adicional para término de relación. MARC 21 se ha extendido recientemente con subcampos para información de relación en los campos de entradas secundarias. El subcampo de relación de campos \$8 es único de MARC 21 y puede usarse para relacionar un campo de datos con otros campos de datos

específicos. Hay además otras diferencias menores tales como que los subcampos específicos sean repetibles o no. Aunque ambos formatos pueden incluir información equivalente sobre la misma manifestación, las diferencias también tienen algún impacto sobre la descripción.

Puede encontrarse, tanto en UNIMARC como en MARC 21, soporte para tales estructuras, como los campos de relación, el uso de indicadores, campos que pueden (o no) representar entidades separadas y distintas, códigos de relación, etc. Sin embargo, estas características satisfacen sólo en parte los requisitos, y es necesario un enfoque sistemático para identificar las posibilidades actuales y los requisitos futuros para representar el modelo FRBR usando MARC. Además, incluso si los formatos pueden soportar una codificación más explícita correspondiente al modelo FRBR en teoría, hay una variedad de prácticas de catalogación, adaptadas a los diferentes requisitos y tipos de material (por ejemplo, Mcgrath, Bisko, 2008).

Estructuras complejas de FRBR en UNIMARC

A continuación usaremos un enfoque pragmático para analizar la capacidad de UNIMARC para codificar estructuras FRBR complejas. Primero presentamos una estructura sencilla y después damos ejemplos sobre estructuras más complejas, discutimos alternativas diferentes para codificarlas y tratamos de identificar la solución más adecuada. El formalismo subyacente en nuestro análisis es que FRBR modela el universo bibliográfico en un catálogo como una gráfica¹ de nodos escritos de forma diferente (entidades FRBR) con enlaces insertados (relaciones) entre ellos. Cada registro bibliográfico siempre representará una parte de este universo. El único requisito en nuestro análisis es que debería ser posible escribir algoritmos que automáticamente interpretaran una colección de registros y construyeran una gráfica de entidades FRBR que correspondieran a la gráfica FRBR descrita por los registros.

En el modelo FRBR, todas las entidades son unidades de interés distintas, pero interrelacionadas, que queremos poder procesar. Los requisitos para procesar cada registro MARC como un conjunto de entidades interrelacionadas pueden resumirse en las siguientes preguntas:

- ¿Cuál es el conjunto de entidades descritas por cada registro?
- ¿Cuáles son los atributos que describen cada entidad?
- ¿Cuál es la identidad de cada entidad?
- ¿Cómo están conectadas estas entidades?

Las entidades

Un número sustancial de registros en un catálogo de biblioteca tendrá una estructura FRBR bastante simple, y las entidades que están descritas implícita o explícitamente pueden inferirse o de las etiquetas y códigos o de interpretar el modelo mismo. La mayor parte de los registros normalmente describen una publicación única que en el modelo FRBR es considerada una manifestación. Según el modelo sabemos que esta

¹ Realmente un conjunto de gráficas porque puede haber subpartes del universo bibliográfico sin relaciones conocidas con el resto.

manifestación materializa al menos una expresión, que es la realización de sólo una obra. Las personas y otras entidades del Grupo 2 pueden por suerte identificarse a partir de campos específicos y son por esto relativamente fáciles de descubrir. Además, podemos usar otras etiquetas o códigos para identificar obras y expresiones más definidas como las que se describen en las entradas de materia.

Por desgracia, esta visión simple por defecto de un registro no siempre corresponde a la estructura real FRBR que el registro representa. Las manifestaciones pueden materializar múltiples expresiones y puede haber otras entidades en la descripción que sólo estén relacionadas y no contenidas. Aplicar la interpretación por defecto y simple de que cada manifestación materializa una expresión única no sólo es una burda simplificación, sino que también significa que no usamos el modelo FRBR como estaba previsto. Identificar el conjunto correcto de entidades FRBR en un registro, no obstante, es un reto. Las expresiones contenidas pueden describirse en notas de contenido y hay campos que en un contexto pueden identificar una obra individual pero no en otro.

Los atributos

La información del registro, con excepción de los campos que describen el registro mismo, definirá varias características de las entidades. Los problemas de una estructura única de una manifestación que materializa una expresión única pueden reducirse a una tarea de mapeo donde cada campo o subcampo puede estar asignado a un tipo de entidad específico. El significado y uso de campos específicos puede ser ambiguo en términos de qué tipo de entidad describen, pero esto es más un problema semántico que estructural. Por otro lado, si el registro describe una manifestación que representa dos expresiones, puede ser un problema estructural. Si tenemos un valor para idioma, podemos asumir que se aplica a ambas expresiones. Si tenemos valores múltiples para idioma, ¿son ambos idiomas válidos para ambas expresiones o son expresiones en idiomas diferentes? En este último caso, ¿qué idioma se emplea en qué expresión?

La identidad de las entidades

La información bibliográfica se basa en el uso de identificadores descriptivos donde las unidades de interés se identifican por un conjunto de atributos construidos de acuerdo a reglas de catalogación. Esto no es muy diferente del uso de los atributos clave en bases de datos relacionales, pero los sistemas de información bibliográficos son “sistemas abiertos al mundo” donde es posible referirse a cualquier entidad bibliográfica conocida. Desgraciadamente, incluso si logramos identificar el conjunto correcto de entidades y equiparar todos los valores de atributo al tipo de entidad correcto a menudo no tenemos suficiente información para identificar las entidades automática y coherentemente. La identificación descriptiva usada en catalogación es llevada a cabo por los seres humanos quienes son capaces de aplicar una interpretación contextual y hacer un mejor uso de cualquier información que encuentren. Si identificamos una obra de una expresión traducida, utilizando el autor, el título uniforme y la forma de la obra, estas serán las claves que tendrán que coincidir para que la obra también aparezca en otros contextos. Por desgracia, esto no siempre es así. Como obra, puede identificársela por el título original (como el título uniforme), pero como materia u obra relacionada puede catalogarse por el título por el que es conocida en la lengua del catálogo.

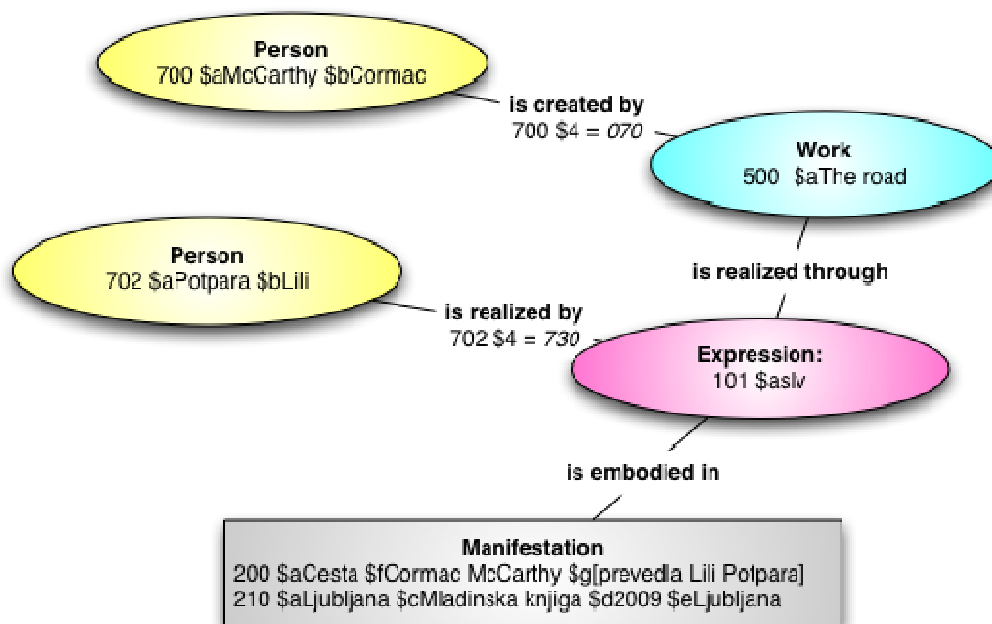
Las conexiones

El desafío estructural final es el de las relaciones. Esto es tanto una cuestión del tipo de relación como de identificación de los términos de las entidades que dicha relación conecta. La información acerca del tipo de relación puede estar implícita o deducirse de las etiquetas de los campos, identificadores o subcampos específicos tales como los códigos de relación en el caso del Grupo 2 de entidades. Para una manifestación que representa una única expresión, el tipo y términos de las relaciones están implícitos, pero de nuevo esto no ocurre en el caso de tener estructuras más complejas. Una manifestación que representa dos expresiones e incluye una única entrada de materia de nombre de persona tiene una estructura ambigua. Conocemos el tipo de relación, pero no los términos, ¿la entrada de materia está relacionada con ambas obras o sólo con una de ellas? Haciendo la conexión correcta entre entidades basadas en la forma solo es posible si hay una única entidad en un lado de la relación. El problema de conectar entidades es incluso más problemático si se desconoce el tipo de relación, que es el caso que se da cuando, por ejemplo, no se usan códigos de relación.

Manifestaciones con expresiones únicas

Las manifestaciones con una expresión única son los casos más triviales de registro bibliográfico y los libros que contienen una sola novela son buenos ejemplos. Esta es también la estructura más común en los registros de un catálogo de biblioteca. En teoría, muchas de estas publicaciones son en realidad más complejas, pero incluso en un contexto FRBR no tendríamos en cuenta generalmente ni el arte de la cubierta, ni las introducciones cortas ni otras expresiones de menor importancia porque tenemos que tomar una decisión equilibrada sobre qué manejar como entidades diferenciadas. Añadir entidades sólo por el uso del modelo no está justificado, sólo introduciría ruido en el sistema de bibliotecas.

Al materializar las manifestaciones expresiones, y ser cada expresión siempre la realización de una, y solo una, obra, podemos inferir que tales entidades existen. Si sólo hay una expresión, y por tanto una obra, el desafío principal es encontrar qué campos y subcampos describen qué entidad así como seleccionar los atributos clave apropiados para estas entidades. Añadir entidades del grupo 2 no tiene por qué causar ningún problema. Entidades del grupo 2 diferenciadas en UNIMARC son bien soportadas por el uso del bloque 7 "responsabilidad intelectual", el uso de ficheros de autoridad, y el uso de códigos de relación. El código de relación es muy significativo ya que es la principal fuente de información sobre cómo una entidad del grupo 2 se relaciona con otra del grupo 1. Si el código de relación en una entrada 700 de nombre de persona es 070, sabemos que esa persona es el autor. En la terminología FRBR será una relación respecto a la obra de *ha creado*, y, al haber sólo una obra, esta relación resulta bastante obvia. Si el código de relación es una entrada de nombre de persona 730, sabemos que la persona es el traductor y por consiguiente debería relacionarse con la expresión.



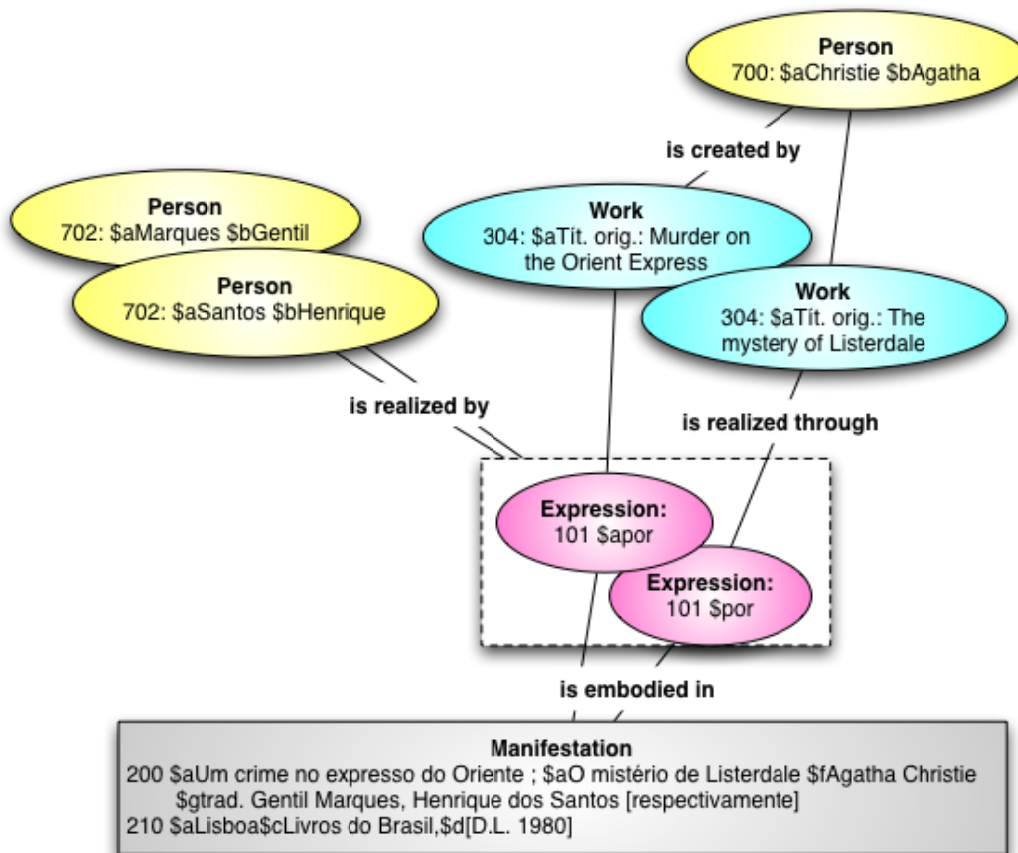
Cuadro 1: Una manifestación que representa una expresión, con ejemplos de información significativa del registro usado para identificar entidades y crear relaciones (Bibliografía Nacional de Eslovenia). Se han añadido espacios entre los subcampos para facilitar la lectura.

El ejemplo del cuadro 1 muestra un ejemplo de este caso mediante la traducción eslovena de *The Road* [La Carretera] del escrito americano *Cormac McCarthy*. Incluso con la secuencia simple de campos del formato MARC, la estructura puede deducirse directamente del registro. Los problemas que conllevan estructuras tan simples están principalmente causados por entidades insuficientemente descritas e identificadas, y por la falta de códigos de relación.

Manifestaciones con dos o más expresiones

Las manifestaciones pueden materializar también más de una expresión: partes integrantes, piezas analíticas o unidades constituyentes. Es una forma de agregación ortogonal a la manifestación que tiene (sub-) manifestaciones como partes, o una obra que tiene otras obras como partes. Dichas agregaciones se describen a menudo en el registro, sin embargo la estructura de las entidades interrelacionadas pueden ser ambigua debido a las limitaciones de MARC. El ejemplo del cuadro 2 muestra una publicación que contiene dos novelas de Agatha Christie en portugués por dos traductores diferentes. En este registro tanto los subcampos repetidos en 200 mencionan de título y la repetición de la nota 304 indican que hay dos obras. Ya que solo hay un autor en el registro, podemos asumir que este autor ha creado ambas obras. Información como el idioma puede relacionarse con expresiones que aplican la misma presunción de que un valor de idioma relaciona ambas. Los traductores están bien identificados, pero no hay una relación explícita que indique qué entrada 702 de persona se relaciona con qué expresión. Esta falta de relaciones explícitas entre entidades puede no parecer importante en un primer vistazo, sin embargo, pone de manifiesto un problema mayor. Si las dos novelas fueran de autores distintos,

tendríamos el mismo tipo de problema que con la relación de *es creado por* que resultaría en identificaciones de las obras insuficientes o erróneas.



Cuadro 2: Ejemplo de una manifestación que materializa dos expresiones (Catálogo Colectivo Portugués)

Los ejemplos de la tabla 1 muestran algunos registros con este tipo de estructura. El primer ejemplo describe una colección de novelas de crímenes de diferentes autores. El contenido puede identificarse parcialmente por los subcampos repetibles en la mención de título, pero el listado de los títulos originales en el campo de notas 304 es una fuente mucho mejor al incluir todas las obras que contiene. En teoría, sería posible usar la secuencia de entradas de títulos y de nombres de persona para relacionar al autor con la obra correcta, pero, como se muestra, el orden de los campos 70X no sigue la secuencia de título en la nota.

El segundo ejemplo describe una colección de novelas de crímenes de Agatha Christie. Este registro usa un campo de relación mediante el método de campo integrado que contiene todos los títulos y el autor usando los campos 200 y 700 respectivamente. Adicionalmente, tiene una nota de contenido 327 con el mismo conjunto de títulos. En este caso, hay múltiples títulos en el campo de relación que enlazan claramente con el mismo único autor.

El tercer ejemplo describe una traducción sueca de *El Señor de los Anillos* de J. R. R. Tolkien. En este caso hay múltiples entradas de relación, una por cada expresión/obra. El campo 500 contiene el título original adecuado para la identificación de la obra, los campos 200 servirían en este caso como títulos de las expresiones, y la inclusión del

autor en todas las entradas hace que la estructura no sea ambigua. El traductor, sin embargo, sólo aparece en la mención de responsabilidad y será difícil usarlo como parte de la identificación de la expresión.

<p>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</p>
<p>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</p>
<p>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</p>

Tabla 1: Manifestaciones diferentes que materializan dos o más expresiones (los dos primeros ejemplos proceden del catálogo colectivo portugués, el último de la bibliografía nacional eslovena)

El tercer ejemplo muestra además otra cuestión estructural. El nombre de persona de la entrada 700 puede relacionarse con la trilogía *El Señor de los Anillos* (como un todo) ya que tiene sentido mencionar que J. R. R. Tolkien fue el creador de *El Señor de los Anillos*, que es la obra a la que comúnmente se le remite. En los otros dos ejemplos dicha mención sería totalmente equivocada. Crear una obra para el título *1950s omnibus* y asociarla a Agatha Christie como creadora de dicha obra sería una mención errónea, ella es meramente la autora de las novelas que la forman. Para el primer ejemplo, sería asimismo erróneo asociar todos los autores como creadores de la colección. La presencia de un 500 de título uniforme en el último ejemplo puede indicar que deberíamos considerar el todo como una obra, pero esta es una asunción que sólo funciona con estos ejemplos porque el último es una traducción.

Listando las expresiones/obras que usan cualquiera de las técnicas tendríamos la estructura de "contenido" actual, pero el uso de los campos de relación es de lejos el método más preciso ya que posibilita la inclusión de toda la información necesaria para la identificación de cada expresión y obra. El uso de campos múltiples es la más apropiada puesto que permite la diferenciación entre las entradas y está basado en una estructura que se corresponde con el agrupamiento natural de información. Una trampa en el uso de cualquiera de estas técnicas es que puede haber información en otros campos que sea común a todas las expresiones que contiene, tales como el idioma. Si sólo hay un valor de idioma, es natural asumir que es válido para todas las expresiones. Dos o más valores de idioma sólo tendrán sentido si son válidos para todas las expresiones. Añadir idioma a los enlaces integrados podría ser una solución para el no tan extraño caso de expresiones en diferentes idiomas, pero el problema existe también para otro tipo de información, tal como las entradas de materia. En MARC 21 el emparejamiento entre campos es soportado por el uso del enlace del campo \$8, que es una solución potencial al problema también en UNIMARC.

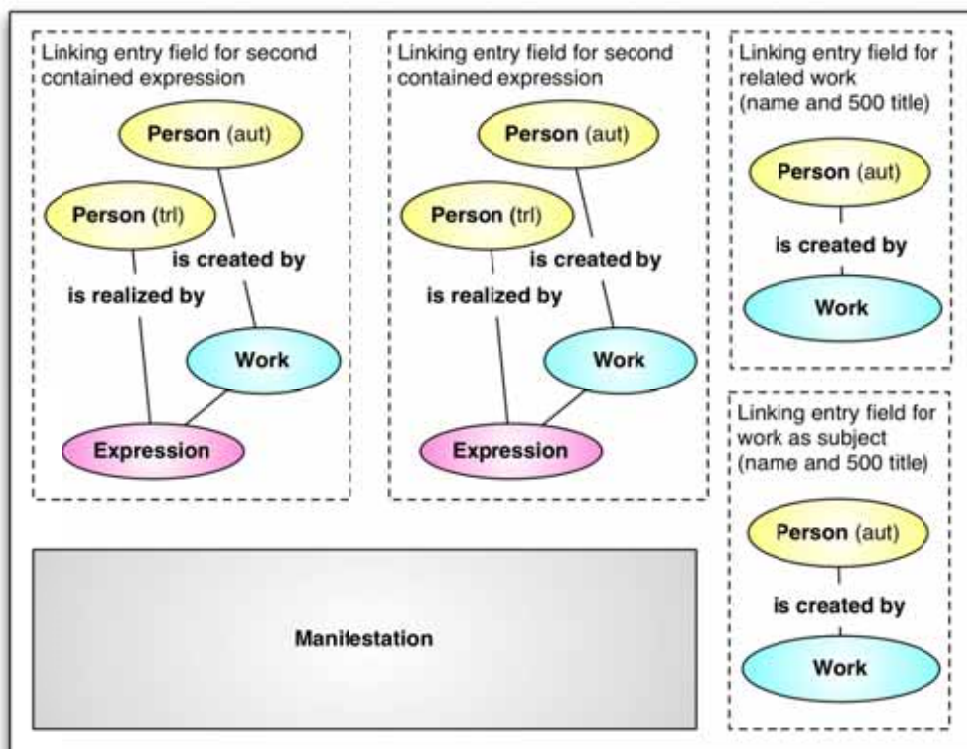
Relaciones con otras obras

El modelo FRBR define una variedad de relaciones entre entidades del grupo 1 además del básico *se materializa en* y *se realiza a través de*. Las razones para añadir tales relaciones a un registro son dar una mejor descripción del recurso catalogado y permitir a los usuarios encontrar obras relacionadas. En el contexto FRBR dichas relaciones son incluso más importantes porque proporcionan un mecanismo de soporte para nuevos tipos de interacciones de los usuarios como son la exploración y el descubrimiento basado en la consulta y en la búsqueda. Los campos de relación son los mecanismos principales para crear estas relaciones en UNIMARC, pero no soporta más que algunos de los tipos de relaciones descritos en FRBR. Todas las demás relaciones tendrán que catalogarse usando una entrada de relación 488 genérica, posiblemente con el uso de un 311 de notas. Este último no es precisamente un muy buen mecanismo. Por un lado es una descripción textual no controlada creada únicamente para ser visualizada. Por otro lado, fallará si hay otros campos de relación y la nota sólo sirve para un campo determinado. El método de enlace integrado es por lo general mejor que la técnica de subcampo estándar ya que permite la identificación detallada de una entidad usando el mismo conjunto de campos de datos que se usarían si la entidad se describiera utilizando los campos de datos corrientes de un registro. Esta técnica permite una identificación de obras, expresiones y manifestaciones bastante precisa, pero es necesario un mecanismo mejor para insertar las relaciones.

Incluso si los campos de relación y el uso de enlaces integrados parece adecuarse bien para expresar relaciones, hay un desafío mayor en esta técnica cuando se trata de manifestaciones que materializan expresiones múltiples. Si todas las expresiones que contiene se describen mediante campos de datos de relaciones, y no hay uno o más campos de relación que describan obra con obra o expresión con expresión, no hay forma de decir con cuál de las expresiones se relaciona. Podemos asumir que es con todas, pero se considera que esta simplificación produce interpretaciones incorrectas.

Otro tipo importante de relación es la de obras que tienen entidades del grupo 1 o 2 como materias. En UNIMARC hay campos de datos específicos para personas, entidades corporativas o nombres de familia como materias y para título como materia. El campo 604 es para nombre y título y está basado en el método de campo integrado. De nuevo, hay un problema si la expresión que contiene (o mejor, las obras correspondientes) tiene materias distintas, como es el caso de las colecciones de ensayo, etc. No hay una forma específica de indicar qué materia pertenece a qué obra.

El cuadro 3 muestra la agrupación que puede obtenerse del uso de campos de relación y del método de campos integrados. Los diferentes campos “corrientes” describen la manifestación y cualquier información que sea válida para todas las expresiones/obras que contiene. Se usan dos campos de enlace para describir dos expresiones materializadas diferentes y las obras a través de las que se realizan, así como las personas relacionadas con expresiones y obras específicas. Gracias a la etiqueta del campo de relación sabemos que son expresiones que contiene, relacionadas con la manifestación descrita en el registro (usando por ejemplo el 423 para seguir la convención de los ejemplos anteriores). Las entradas de relación adicionales son para una obra relacionada y para una materia. La conexión entre estas entradas y las entradas para las expresiones que contiene, sin embargo, no se describen y no pueden ser interpretadas con exactitud.



Cuadro 3: Ejemplo del uso de campos de relación y del método de campos integrados.

Conclusiones

En este trabajo hemos discutido y explorado algunos de los problemas que existen cuando codificamos FRBR en UNIMARC. El problema de la asignación de identidad y atributo sobre todo se origina en la práctica de la catalogación (y reglas de catalogación) y está causado por una información insuficiente (como la falta de títulos que identifiquen adecuadamente las obras) o campos de datos que son ambiguos en el contexto de FRBR. RDA trata algunos de estos problemas y con suerte surgirán cambios en los campos de datos en los formatos MARC.

El problema de definir y describir un conjunto de entidades con relaciones enlazadas correctamente en UNIMARC es un asunto diferente. El formato actual es suficiente para describir manifestaciones con una sola expresión, pero para estructuras más complejas los ejemplos que hemos mostrado indican que hay varias soluciones actualmente en uso, basadas en notas de contenido, subcampos repetibles en la mención del título y campos de relación. Las dos primeras usan información descriptiva, lo que no es suficiente porque son difíciles de procesar y darán como resultado una pobre identificación de las entidades. Por suerte, el uso de campos de relación parece ser una solución adecuada para resolver muchos de los problemas relacionados con las estructuras más complejas. Puede verse como un truco que añade una nueva profundidad a la estructura básica MARC que parece encajar en el modelo FRBR. Esta técnica imita el uso de múltiples registros relacionados, pero con la ventaja de guardar toda la información en un registro. Con el método de campos integrados podemos

incluir toda la información necesaria para describir e identificar todas las entidades relacionadas con la “rama” de la expresión contenida.

Un problema particular que hemos tratado es la ambigüedad que aparece cuando la información en un campo específico puede relacionarse con múltiples expresiones u obras contenidas, o cuando hay obras/expresiones relacionadas que sólo son válidas para expresiones/obras contenidas específicas. Para los campos de datos relacionados con el contenido esto puede resolverse con reglas sobre qué incluir en los campos integrados de la entrada de relación; pero para las relaciones entre obras en entradas de relación separadas, esta solución es imposible. Un mecanismo de relación de campos comparable al subcampo \$8 en MARC 21 podría resolver el problema, o bien añadiendo un subcampo para este fin (que no entre en conflicto con otros subcampos integrados) o bien usando el desaprovechado primer indicador para este propósito (sólo aplicable para conectar campos de relación).

Si queremos que los campos de relación soporten plenamente las relaciones de FRBR, hay una necesidad adicional de entradas de campos de relación designados específicamente, o de algún mecanismo que exprese el tipo de la relación añadiendo un identificador de dicho tipo. Lo primero produciría un gran número de campos de relación, lo último requiere algún subcampo designado específicamente para este fin, que es difícil de introducir porque puede entrar en conflicto con los subcampos integrados.

Sólo hemos descrito casos relacionados con expresiones múltiples materializadas en una manifestación, pero somos conscientes de que hay todavía más tipos de estructuras que explorar tales como manifestaciones, expresiones y obras que tienen partes múltiples, enmiendas y otras. Algunos problemas pueden resolverse con campos de relación, pero puede haber casos que requieran otras soluciones.

Finalmente destacaremos algunas recomendaciones más generales para el desarrollo futuro de UNIMARC:

Un cambio en perspectiva desde la presentación al procesamiento

Implantar FRBR en un sistema bibliotecario significa más que cambiar simplemente el formato o el modelo de datos subyacente. FRBR representa un modelo lógico del dominio del discurso, y sus implementaciones dependen de formatos de intercambio y modelos de almacenamiento que reflejen el modelo. Los requisitos reales de los registros bibliográficos en el futuro dependen en gran parte de lo que queramos que los datos sean capaces de hacer. Recientes desarrollos en las interfaces de usuarios de biblioteca han explorado el uso de agrupamientos parecidos a FRBR de los resultados y cada vez somos más conscientes de los beneficios del modelo. Un desafío importante para el futuro es definir los requisitos reales de los datos en términos del procesamiento y consultas que queremos que soporten los datos, no en términos de cómo debería visualizarse el registro.

Revisar la norma

El modelo FRBR ha influido en RDA, y en consecuencia influirá en las normas para codificar registros bibliográficos en el futuro. Se han hecho algunos cambios al formato MARC 21, y es importante que UNIMARC pueda modernizar el formato como corresponde. Un mejor soporte de FRBR no introduce necesariamente incompatibilidad retrospectiva, pero la falta de soporte de FRBR seguramente introduzca incompatibilidad futura.

Definir la mejor práctica para codificar FRBR

Codificar FRBR en MARC no es necesariamente una cuestión de añadir más información al registro e introducir un coste de catalogación adicional. El principal problema es el uso de diversas soluciones en diferentes catálogos así como en el mismo catálogo. Algunas prácticas de catalogación se consideran menos útiles que otras en futuras implantaciones de FRBR. En la norma actual no hay ejemplos que se relacionen con FRBR y hay una falta general de directrices o ejemplos de la mejor práctica para codificar información de un modo compatible con FRBR. Incluso sin cambiar el formato y sin cambiar las reglas y la práctica de la catalogación, es posible sugerir mejores soluciones prácticas para un buen número de casos comunes.

Indicar adecuadamente la versión

La mayoría de los catálogos de bibliotecas contienen registros que se han creado y agregado a lo largo de muchos años o incluso décadas. La práctica y los formatos pueden haber cambiado con el tiempo, lo que causa diferencias menores o mayores. Las migraciones desde los formatos nacionales hacia los formatos reconocidos internacionalmente pueden ajustarse sintácticamente a las normas finales, pero siempre hay restos de una práctica previa que son imposibles de migrar. Cambiar a un uso diferente del formato adaptado al modelo FRBR es un cambio que debería llevarse a cabo usando algún tipo de versión de modo que podamos identificar estos registros. El actual Identificador de la Versión 005 UNIMARC es simplemente la fecha de la última actualización, y no indica nada más. Si un catálogo de biblioteca ha de contener registros que estén adaptados a modos específicos de expresar y procesar FRBR, mientras que otros no, un requisito mínimo es que podamos distinguirlos usando un identificador de la versión.

Referencias Bibliográficas

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, 18-27 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>