



68th IFLA Council and General Conference

August 18-24, 2002

Code Number: 053-133-S
Division Number: IV
Professional Group: Cataloguing
Joint Meeting with: -
Meeting Number: 133
Simultaneous Interpretation: Yes

Sondeando el MARC para encontrar : FRBR?

Knut Hegna

University of Oslo Library
Oslo, Norway
E-mail: Knut.Hegna@ub.uio.no

Eeva Murtomaa

Helsinki University Library
Helsinki, Finland
E-mails: Eeva.Murtomaa@helsinki.fi

Resumen

*Este artículo resume un proyecto en que se analizaron datos MARC procedentes de dos bibliografías nacionales, bajo el prisma del modelo de datos presentado en el estudio FRBR de la IFLA. Durante el proyecto encontramos que, aunque la información en los registros MARC contiene atributos pertinentes para identificar los tipos de ente de **obra**, **expresión** y **manifestación**, la exactitud y sintaxis formal son demasiado simples para ser tratados adecuadamente mediante programas. Algunos de los resultados pueden servir para presentar mejor las listas de resultados en los OPAC. El proyecto presentó dos sugerencias para un interfaz de usuario OPAC, basadas en las ideas del estudio FRBR y en los resultados del proyecto.*

El informe completo se puede encontrar en:
<http://folk.uio.no/knuthe/dok/frbr/datamining.pdf>

1 - FRBR y MARC

La primera pregunta que nos hicimos fue: ¿Podemos encontrar la estructura FRBR en los registros MARC existentes? Si miramos un solo registro, vemos que hay información sobre la obra, la expresión y la manifestación:

```
*008880325          no          esp
*02000 $a 82-991075-2-0 $b h. $c Nkr 60.00
*04110 $a espnor
*08200 $a 839.822[S]
*10010 $a Ibsen, Henrik $d 1828-1906
*24510 $a Puphejmo (1879) $c Henrik Ibsen ; tradukis:Odd Tangerud ; lingve kontrolita de Esperantista Verkista Asocio (EVA)
*26000 $a Hokksund $b Eldonejo Odd Tangerud $c 1987
*26900 $a [Drammen] : Tangen-trykk
*30000 $a [1], 57 s. $c 24 cm
*50000 $a Originaltittel: Et dukkehjem. - Originalutgave: København : Gyldendal, 1879
*99100 \ $a Tangerud, Odd
```

Éste es un registro procedente de la bibliografía nacional noruega que describe un documento que contiene una versión en esperanto de la famosa obra "Casa de muñecas" de Henrik Ibsen. Encontramos el título de la obra en el campo de nota 500 y la relación con el autor de ésta, en el campo 100. Esta información permite identificar la obra en este registro.

Además, encontramos información sobre la expresión a través de los códigos de lengua en los campos 008 y 041 y una relación con la persona responsable de la traducción en una parte de la mención de responsabilidad y en el campo 991.

Encontramos información sobre la manifestación en aquellos campos que describen el documento que tenemos en la mano: 245, 260, 269, 300 y otros.

Por lo que la respuesta a nuestra pregunta es:

Sí: ¿POR QUÉ? Porque un registro bibliográfico puede:

- describir tanto la obra como la manifestación
- contener rasgos de la expresión
- contener algunas relaciones en las entradas secundarias, notas y descripciones de materia.
¿Elementos del modelo de FRBR están presentes, de alguna manera, en el registro MARC!
Pero, desgraciadamente, la respuesta también es No. ¿POR QUÉ?
- Porque las reglas de catalogación se ajustan bien al catálogo manual y a la bibliografía impresa, pero no al modelo FRBR.
- La parte principal de la información a menudo se graba adecuándose más a la mente y al ojo humano que al ordenador.

No tiene sentido mirar un único registro. Decidimos analizar conjuntos de registros generados a partir de una búsqueda, mirando en las bibliografías nacionales finlandesa y noruega. Nuestro objetivo se limitó a la búsqueda por autor.

Pusimos un signo de interrogación al final del título del proyecto. No sabíamos qué resultados alcanzaríamos. Sentíamos que la investigación debía permanecer abierta. Por lo menos, esperábamos identificar, a este respecto, algunos problemas con las reglas de catalogación y los formatos MARC.

Empezamos mirando en el estudio FRBR las tablas que mostraban qué atributos de los entes son importantes para identificar la obra, la expresión y la manifestación. A continuación intentamos entonces equiparar la información de los registros MARC a estos atributos.

Para localizar el título de la obra buscamos títulos uniformes o títulos originales. Primero comprobamos si hay un campo 241 (Título original). La bibliografía nacional noruega no usa este campo, pero la finlandesa si. Si no encontramos el campo 241, comprobamos los campos 500 y 505 en busca del "Título Original" en los dos idiomas: Originaltittel:, Orig.tit.: (abreviado) o Originaltitler: (plural) en registros noruegos y Alkuteos: o Alkuteokset: (plural) en registros finlandeses. Los títulos originales siempre van detrás de este texto. Si esta comprobación no nos da ningún título original en los campos 50X, nos situamos en otros campos del título siguiendo esta secuencia: 248 \$h (Título de parte en registros finlandeses) y el campo 240 \$a. La última posible solución es coger el título o títulos originales del campo 245.

Normalmente los títulos mencionados en el campo 505 se repiten en las entradas secundarias de título (745 \$a en Finlandia, 740 \$t en Noruega), pero no hay ninguna manera de decidir si la información en estos campos corresponde a títulos originales o no. Algunas veces los campos contienen los títulos originales y, otras, otra información sobre el título. Ni el campo 745 ni el campo 740 contienen información que califique al título. En muchos casos, estos campos 74X son entradas secundarias relativas a otra obra más que a la obra en cuestión, pero que de alguna forma se relacionan con ella.

Atributo FRBR	Valor FRBR	NORMARC	FinMARC	Seleccionado
Título de la obra	Alto		241 \$a	SI
		500 \$a	500 \$a	SI
		500 \$a	505 \$a	SI
			248 \$h	SI
		240 \$a	240 \$a	SI
		245 \$a	245 \$a	SI
Relación a la persona responsable	Alto	100 \$a	100 \$a \$h	SI
		70010 \$a	70010 \$a \$h	
Terminación propuesta	Alto	?	?	NO
Forma de trabajo	Moderado	Interpretación de Dewey	008 pos.24-27, 29-30, 33-34	

Tabla 1: Atributos que identifican la entidad de trabajo en formatos MARC.

Autor	Número de registros	Número de líneas de trabajo	Número De líneas de trabajo únicas
Ibsen, Henrik (n)	744	914	220
Wassmo, Herbjørg (n)	149	159	19
Gaarder, Jostein (n)	237	237	14
Solstad, Dag (n)	92	93	35
Kunnas, Mauri (f)	130	133	41
Jansson, Tove (f)	576	595	92
Linna, Väinö (f)	168	168	20

Tabla 2: Resultados del uso del procedimiento de reducción de trabajo en registros de las bibliografías nacionales Noruega (n) y Finlandesa (f).

Además de establecer la relación con el autor de la obra, pensábamos que podríamos identificar la obra u obras en cada registro. Sobre esta base, con estos datos, podríamos reunir las obras idénticas encontradas en varios registros, y diferenciarlas de otras obras. Una vez hecho esto, empleamos otros datos para diferenciar/colocar las diferentes expresiones, identificadas por el código de lengua y de traductor. Se utilizaron otros datos para identificar las manifestaciones.

Éstos son algunos resultados:

Comprobamos que había 744 registros donde Henrik Ibsen aparece como autor o como entrada secundaria. Pudimos extraer 914 identificadores de obra de estos registros, de los cuales identificamos 220 obras únicas. Este resultado parece bueno, pero Ibsen escribió 26 obras y algunas colecciones de poemas. ¿Por qué, entonces, 220 obras?

Había registros con falta de información o con información errónea: las colecciones y selecciones de obras donde los títulos originales o los títulos uniformes no eran identificables aun cuando estaban presentes (12); 33 registros donde el subtítulo estaba incluido en el propio título, dando lugar a 8 obras únicas; 3 registros con erratas en el título original; 14 registros con ortografía moderna que difiere del título original; 10 registros sin título original alguno, que generaban a encabezamientos de la obra con títulos en lenguas extranjeras.

Estos problemas se discuten en nuestro informe.

2 - La interfaz de usuario

Encontramos alguna estructura. La segunda pregunta que nos hicimos entonces fue: ¿Cómo podríamos beneficiarnos de esta estructura para presentar al usuario las listas de resultados de una búsqueda?

Primero buscamos en las tablas FRBR para encontrar qué atributos eran importantes para seleccionar entre entes similares y convertimos esos atributos a etiquetas MARC.

Para hacer la selección más sencilla es importante también decidir qué tipo de clasificación es más apropiado. La clasificación debe ser fácilmente identificable por el usuario y debe variar según el nivel de ente que se presenta.

Creemos que la lista de resultados de una búsqueda debe aparecer según la búsqueda desarrollada y según los propios resultados. Cuando buscas a una persona concreta la lista de resultados de sus obras debe constar siguiendo un orden, quizás alfabético o cronológico, o bien puede ser una lista de las distintas funciones que él realiza, relacionada con los entes (autor, ilustrador, traductor).

En la primera interfaz, la lista de resultados se presenta mediante tarjetas superpuestas. La parte superior de las tarjetas se visualiza desplegándose de forma horizontal, el frente de las tarjetas se

despliega de forma vertical, diferenciando la información gracias al encabezamiento existente en la parte superior de cada tarjeta.

Los dos ejes (horizontal y vertical) podrían corresponder a autores-obras, obras-expresiones, expresiones-manifestaciones, materias-obras. Esto dependería de la búsqueda realizada. Si el usuario busca a un autor concreto y hay varios resultados, el eje horizontal mostrará la lista de los nombres de persona y el eje vertical las obras de cada persona. Si, por otro lado, la búsqueda proporciona un único resultado, el eje horizontal contendrá las obras del autor y el eje vertical las diversas expresiones.

Esto se puede ver con las obras de Jostein Gaarder sobre el eje horizontal y las expresiones sobre el eje vertical bajo cada obra. El número de expresiones identificado por cada obra se proporciona entre paréntesis al final del encabezamiento de cada obra. El número de manifestaciones por cada expresión se encuentra al final de la información sobre la expresión.

El usuario podría seleccionar una expresión específica iniciando, por consiguiente, una ventana independiente que contiene todas las manifestaciones bajo esta expresión. Las manifestaciones de la expresión escogida se ordenan cronológicamente y el año de la publicación se presenta al comienzo cada registro. La ventana contiene botones con varias funciones como imprimir, salvar y ordenar, y la ventana no desaparece a menos que el usuario la cierre explícitamente. Esto significa que los usuarios pueden mantener tantos conjuntos de manifestaciones como deseen, todos en ventanas independientes.

La segunda interfaz de usuario presenta las obras en una estructura de árbol, como nodos o ramas de una lista de resultados. Las obras aparecen ordenadas alfabéticamente según el título original y el número al final del título indica cuántas expresiones se identifican bajo esta obra.

Los nodos de cada obra se expanden y las hojas son las expresiones. Las expresiones se ordenan de la misma forma que la interfaz previamente descrita, la lengua original en la parte superior y después las lenguas por orden alfabético.

Estas hojas son activas, una pulsación del ratón abre una ventana de manifestaciones igual que en el caso anterior.

3 - Resumen

Una de las razones principales para el ruido que hemos observado a través de nuestros experimentos deriva de un uso poco experto del campo 700 y también de la manera en que los sistemas indexan esta información junto con el campo 100, sin tener en cuenta la función real que la persona tiene.

Consideramos que la función en el campo 700 debe ser obligatoria en las reglas de catalogación y que la lista de funciones debería además ampliarse. Los sistemas deberían utilizar la función para presentar al usuario final unas listas de resultados más estructuradas.

Esto permitiría que los sistemas presentaran todas las funciones que un autor pueda tener en una base de datos, haciendo posible que los usuarios finales pudieran elegir todas o algunas de las funciones en sus navegaciones bibliográficas.

Nuestra investigación también mostró que sería una ventaja para el análisis si pudieran introducirse los títulos originales de una manera más consistente, en campos repetibles e independientes.

Los códigos de lengua son uno de los atributos más importantes para identificar las diferentes expresiones de una obra específica. También son importantes para identificar si una manifestación es una traducción o no.

La normalización de la información es en sí misma una forma de enlace. El uso de ficheros de autoridades en el entorno bibliográfico ayuda a establecer la estructura de navegación, tanto al unir entidades como al diferenciar entre ellas. Los ficheros de autoridad más comúnmente utilizados son los de nombres de persona y de entidad. Nuestro trabajo muestra que también se necesita un gran trabajo de autoridades para que sea posible poner la misma obra bajo un solo encabezamiento.

Gracias.

Translation by Mar Hernandez