

MIGRACIÓN DE REGISTROS CATALOGRÁFICOS EN FORMATO LIBRI (UNSJ) HACIA MARC21

Proyecto: "Búsqueda Estratégica de Conocimiento de los datos de Biblioteca y Alumno de la FCEFYN" Resol. 034/08-CS Código: E824/2007

Autores:

Luis Alberto Olguin, Instituto de Informática, FCEFYN-UNSJ <lolguin@iinfo.unsj.edu.ar>

Nora Maroto, Biblioteca "Emiliano P. Aparicio", FCEFYN-UNSJ <nmaroto@unsj-cuim.edu.ar>

Cristina Coria, Biblioteca Franklin <bibliotecafranklin@speedy.com.ar>

Gabriela Rodríguez, Instituto Superior de bibliotecología, <bibliotecologia@sinectis.com.ar>

Resumen:

El formato catalográfico local LIBRI, diseñado por la Dirección General de Bibliotecas de la Universidad Nacional de San Juan, con la colaboración del Instituto de Informática de la Facultad de Ciencias Exactas de la UNSJ, lleva más de 15 años de uso en las bibliotecas de esta Universidad como así también, a través de la distribución, por parte de la Comisión Nacional de Bibliotecas Populares (CONABIP), del software SIGEBI, en cientos de Bibliotecas Populares de Argentina.

Dentro de las actividades llevadas adelante por el Proyecto "Búsqueda Estratégica de Conocimiento de los datos de Biblioteca y Alumno de la FCEFYN", se conformó un equipo interdisciplinario con bibliotecarios e informáticos de la Facultad de Ciencias Exactas de la UNSJ, docentes del Instituto Superior de Bibliotecología de San Juan y bibliotecarios de la Biblioteca Popular Franklin para concensuar la tabla de migración de datos entre el formato local LIBRI y el formato internacional MARC.

Se expone en el presente trabajo los alcances de la tabla de conversión generada como así también las herramientas informáticas utilizadas en el proceso de migración.

Introducción:

Desde hace ya varios años, el anhelo y las acciones en algunos sectores de la provincia de San Juan han sido numerosas en cuanto a la conformación de un Catálogo Colectivo Provincial de bibliotecas, el cual en un futuro, pueda integrarse a un Sistema Nacional de Información que permita conocer el material que posee cada biblioteca y realizar el procesamiento técnico en forma cooperativa.

Para conseguir este objetivo, se deben generar cambios importantes en las bibliotecas que actualmente poseen sistemas automatizados. Entre estos cambios destaca la selección de un nuevo formato común para la generación de sus catálogos bibliográficos que permita el intercambio de datos con las bibliotecas de la región y el resto del mundo.

A fines de 1980, la Dirección General de Bibliotecas de la UNSJ comenzó a trabajar en el diseño de automatización del catálogo para proveer a estas bibliotecas de un formato común que respondiera a las necesidades y a las características de los distintos tipos de documentos, que facilitara el desarrollo de bases específicas, la cooperación con el catálogo y la integración en redes. Se implementó así, a principios de 1990, el formato LIBRI para el registro de información documental, lo que permitió que las bibliotecas dispusieran de un formato común para el registro de sus documentos y la normalización de la transferencia de información.

Este proceso fue acompañado por la realización de cursos, talleres y asistencia técnica.

LIBRI es el formato oficial de la UNSJ para la carga de datos bibliográficos que cubre todos los tipos de materiales (libros, seriadas, mapas, audiovisuales, etc.). Además es utilizado por una treintena de bibliotecas sanjuaninas (populares y escolares) las que catalogan sus obras siguiendo sus lineamientos.

LIBRI se incluye como formato catalográfico en la distribución del software SIGEBI de la Comisión Nacional de Bibliotecas Populares (CONABIP), organismo dependiente de la Secretaría de Cultura de la Nación.

MARC es hoy en día el formato más utilizado en las bibliotecas de varias partes del mundo y en Argentina es el formato al que muchas bibliotecas (universitarias, públicas, privadas) están adhiriendo.

“Desarrollado hace más de treinta años, el formato MARC ha marcado su impronta en los procesos de automatización de bibliotecas no sólo en los Estados Unidos, sino también en muchos otros países del mundo. MARC es, para muchas comunidades bibliotecarias, sinónimo absoluto de automatización de registros bibliográficos. Para otras comunidades bibliotecarias, entre la que podemos contar la de nuestro país, MARC es todavía un misterio por descifrar. La necesidad de conocer el formato MARC se hace, por otra parte, cada vez más acuciante. Las tendencias mundiales hacia la adopción de formatos MARC es innegable. La Internet ha derribado muchas barreras que impedían o al menos dificultaban la consulta e intercambio de información bibliográfica” [1]

En la actualidad las bibliotecas sanjuaninas alcanzadas por el proyecto trabajan sobre una plataforma similar, CDS/ISIS de UNESCO, sobre la que han implementado para la generación de sus catálogos bibliográficos el Formato LIBRI pero no existe entre ellas una conexión permanente que, por ejemplo, garantice la catalogación unificada a una sola base de datos.

La tabla de migración:

Si bien los integrantes del equipo conocían MARC por distintos cursos realizados, fue necesario el estudio detallado de los dos formatos para, posteriormente lograr la tabla de migración de datos.

Luego de la investigación, lectura y comparación de ambos formatos, LIBRI y MARC, para la implementación de la tabla de migración se decidió adoptar un diseño similar al que se encuentra en <http://www.loc.gov/marc/unimarc21.html> (migración de datos de UNIMARC a MARC).

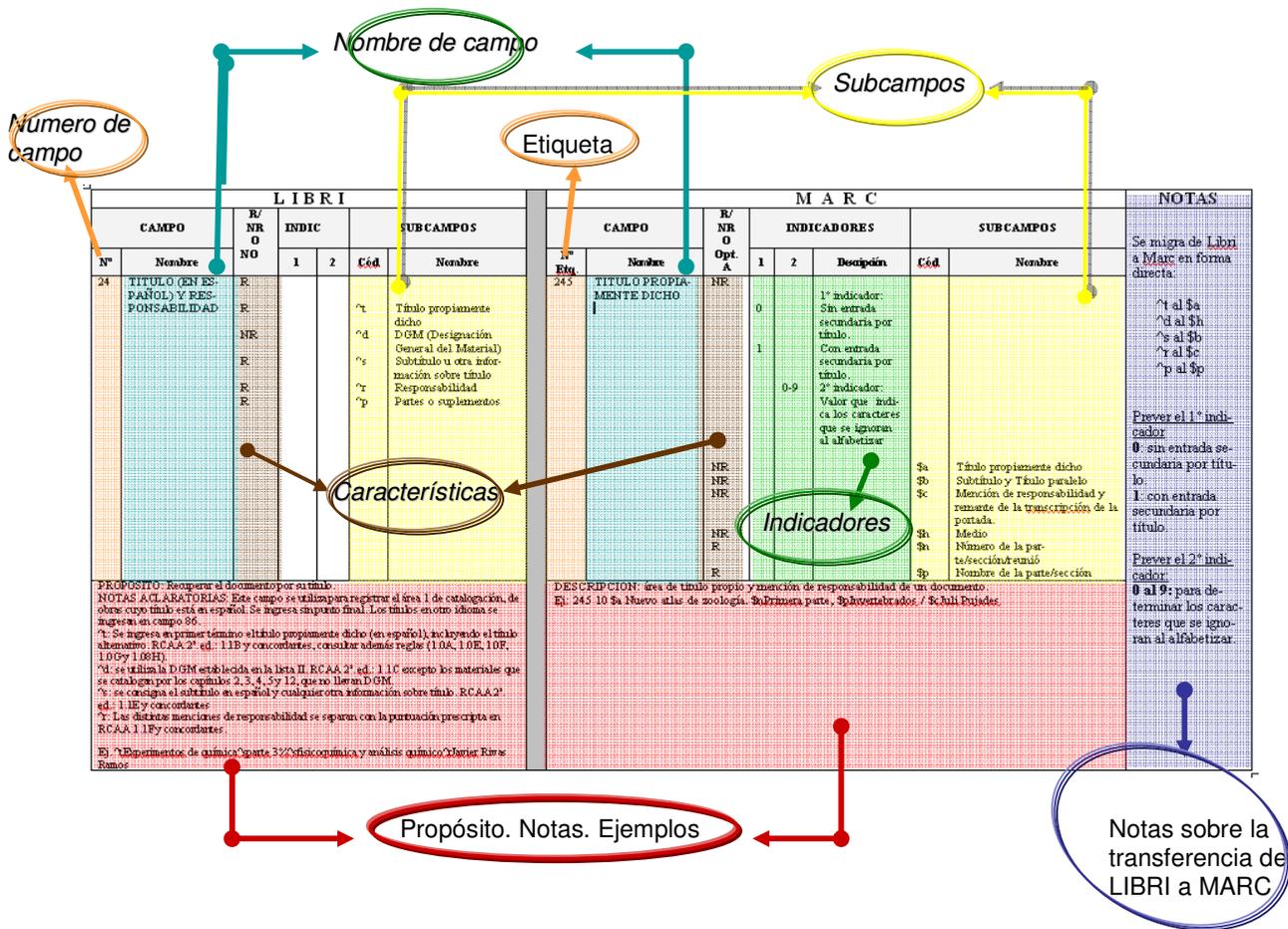
La tabla que presentamos proporciona las especificaciones para la conversión de los registros bibliográficos en formato LIBRI(UNSJ) a formato MARC **abarcando solo la descripción de los campos asociados al material monográfico (LIBROS)**.

La descripción de los campos LIBRI se ajustan a lo publicado en el manual “Formato LIBRI”, 2ª ed., 1993.

La descripción de los campos MARC es concordante a lo publicado por Library of Congress en su sitio web con fecha Noviembre 2009 (<http://www.loc.gov/marc>).

Las especificaciones de conversión se presentan en forma tabular, dispuestas en el orden de los datos del formato LIBRI.

La tabla presenta tres bloques: uno para la descripción de cada campo en el Formato LIBRI, otro para la descripción del mismo campo en el Formato MARC y el último incluye Notas aclaratorias respecto a la transferencia de datos.



Para explicar con más detalle y estableciendo en forma permanente comparaciones entre los dos Formatos, observamos lo siguiente:

<p align="center">Primera Columna LIBRI: Número de Campo MARC: Número de Etiqueta</p> <p>El número de campo en LIBRI está compuesto por dos dígitos, excepto los campos de Entradas Principales (100, 110, 111, 130). En MARC, en cambio, los números de etiqueta están constituidos por 3 dígitos.</p>	<p align="center">Segunda Columna Nombre del campo</p> <p>Representa el nombre del campo, coincidiendo en general las denominaciones utilizadas en cada formato.</p>
<p align="center">Tercera Columna Repetibilidad/Obligatoriedad</p>	<p align="center">Cuarta Columna Indicadores (MARC)</p>

<p>Muestra las características de Repetibilidad y obligatoriedad de los campos y subcampos.</p>	<p>Corresponde a los Indicadores. Aquí hay una gran diferencia entre los dos formatos ya que LIBRI no presenta estos datos. Esta situación hace que a la hora de la migración de registros hacia MARC sea necesario incorporarlos de manera automatizada</p>
<p style="text-align: center;">Quinta Columna Subcampos</p> <p>Refleja la información de Subcampos. Ambos formatos poseen delimitador de subcampo, aunque los símbolos utilizados en uno y otro son distintos. Sólo en algunas ocasiones el código de subcampo coincide tanto en LIBRI como en MARC.</p>	<p style="text-align: center;">Sexta Columna Notas</p> <p>Consigna notas explicativas sobre la transferencia del Formato LIBRI a MARC</p>
<p style="text-align: center;">Parte Inferior Ejemplos</p> <p>Muestra Propósitos, Notas aclaratorias y ejemplos extraídos del Manual del Formato LIBRI y la correspondencia de esos ejemplos en MARC.</p>	

El mapeo de migración realizado incluye **solo los campos para descripción monográfica** habiéndose logrado que la totalidad de los datos almacenados en los registros involucrados sean transferidos hacia MARC.

A continuación se muestra un cuadro comparativo de los campos analizados.

L I B R I		M A R C	
00	Masterfile Number	001	Número de control de registro
02	Unidad Catalogadora	003	Identificador del número de control
03	Fecha de Alta o Modificación	005	Fecha de última modificación
		008 (00-05)	Fecha de ingreso
05 Y 06	Nivel Catalográfico Nivel Catalográfico Superior	000 (07)	Nivel Bibliográfico
10	ISBN	020	ISBN
13	Código UNSJ	913	Campo Local
24	Título (en español) y responsabilidad	245	Título Propiamente Dicho
26	Otros Títulos	246	Variante del Título
27	Título Uniforme	240	Título Uniforme
28	Entrada Secundaria por Nombres Personales	700	Entrada Secundaria por Nombres Personales
29	Entrada Secundaria por Entidad	710	Entrada Secundaria por Entidad
30	Entrada Secundaria por Reunión	711	Entrada Secundaria -Nombre de Congreso
40	Idioma	008 (35-37)	Idioma del Item (Determina el Idioma en que está escrito el Ítem).
		041	Idiomas (Para más de un idioma). Se usa cuando el código de idioma del campo 008 (35-37) no es suficiente.
42	Edición	250	Mención de Edición
43	Lugar y Editor	260	Publicación, Distribución, etc.
45	Fecha de Publicación		
50	Descripción Física	300	Descripción Física
54	Notas Generales	500	Nota General
53	Nota de Tesis	502	

			Nota de Tesis
52	Nota de Referencia Bibliográfica	504	Nota de Bibliografía
60	Clasif.ón Decimal de Dewey	082	Clasificación Decimal de Dewey
61	Clasif. Decimal Universal	080	Clasificación Dec. Universal
63	Delimitador Geográfico	052	Clasificación Geográfica
64	Delimitador Temporal	045	Periodo cronológico de contenido
65	Descriptor (Tema)	650	Entrada secundaria de materia - Término temático
66	Descriptor (Nombres de personas)	600	Entrada secundaria de materia - Nombre de persona
67	Descriptor (Entidad)	610	Entrada secundaria de materia – Nombre de entidad
68	Descriptor (Reunión, Congreso, etc.)	611	Entrada Secundaria de Materia -Nombre de Congreso
70	Número de Inventario	942	Número de Inventario (Campo Local)
75	Signatura Topográfica	952	Signatura Topográfica (Campo Local)
86	Título (en otros idiomas) y responsabilidad	245	Título propiamente dicho
87	Área de Serie	490	Mención de serie sin entrada secundaria o con entrada secundaria en forma diferente
89	Descriptor (Título Uniforme)	630	Encabezamiento de materia: Título Uniforme
100	Entrada principal por autores Personales	100	Entrada principal por autores Personales
110	Entrada principal por Entidad	110	Entrada principal por Entidad
111	Entrada principal por reunión	111	Entrada principal por nombre de Congreso
130	Entrada principal por Título Uniforme	130	Entrada principal por Título Uniforme

Como complemento del estudio realizado se gestionó ante la Embajada de Estados Unidos la realización de un taller sobre MARC en nuestra provincia, que fue dictado por la Lic. Graciela Spedalieri en el mes de septiembre de 2009 y versó sobre MARC Bibliográfico y de Autoridades. La buena predisposición, apoyo y amplios conocimientos de la Lic. Spedalieri sobre esta temática impulsó al equipo a continuar con más fuerza y seguridad las actividades que se venían desarrollando en el proyecto.

Proceso de migración

El objetivo perseguido en este proceso es el de permitir que cada biblioteca usuaria de LIBRI sea capaz de realizar la migración hacia MARC garantizando que toda la información almacenada pueda ser mapeada entre los dos formatos siguiendo la tabla de conversión propuesta.

Al comenzar con el estudio de la técnica de migración a aplicar, se recurrió a la consulta web y lectura de documentación existente sobre la migración de base Isis hacia MARC.

Entre las múltiples opciones existentes, nos pareció apropiado aplicar aquella que propone la generación de un archivo plano producto de la “impresión a archivo” de registros Isis utilizando un formato de presentación “.pft” con marcado especial para su posterior tratamiento por un software que permita la generación del iso2709-MARC. [2]

El procedimiento para la migración de registros LIBRI hacia Iso-2709-MARC que se presenta requiere de las siguientes etapas:

1. Control global de la base de datos a migrar
2. Pre-procesamiento de campos de LIBRI que así lo requieran
3. Volcado de registros LIBRI en un archivo plano, utilizando el utilitario MX y un formato de presentación (li2marc.pft) con marcado especial.
4. Procesamiento del archivo plano utilizando el software MARCEdit para obtener el Iso2709-MARC.

Control global de la base de datos a migrar

En esta etapa, cada biblioteca debe controlar la existencia de los campos obligatorios de LIBRI, detección de errores de carga en los campos codificados (idioma, tablas, etc.).

Esta tarea es importante planearla adecuadamente para resolver posibles errores u omisiones que se puedan haber cometido al momento de generar/modificar los registros LIBRI.

Pre-procesamiento de campos de LIBRI

Del análisis detallado de los campos de LIBRI vs. MARC, en los que se encontró la necesidad de efectuar un pre-procesamiento de la base de datos, antes de la migración, fue en los campos en donde se han almacenado datos con “marcado para técnica 2 de FST en Isis”, es decir campos en los que se han incluido marcas con paréntesis angulares (<>) a modo de “puntos de acceso”. En nuestro formato esto ocurre en los campos 65 (Descriptor Tema) y 54 (Notas).

En el caso del campo 65, se trata de un campo no repetible, que almacena información temática de la obra y cada descriptor es tratado como una ocurrencia en MARC (campo 650).

En el caso del campo 54, las notas de contenido o los puntos de acceso no controlados son cargados realizando el marcado con paréntesis angulares. Lo que se busca con el pre-procesamiento de este campo es discriminar en un campo auxiliar la información de punto de acceso marcada para luego ser migrada al campo 653 de MARC.

Para este pre-procesamiento se han utilizado los utilitarios MX y MXCP de Cisis [3].

```
mx libri
  "proc='d24 d86 d65 a59#',v54,
  '# a65#','^t',v65^t,'^d',mhl,v65^d,'#',
  ('a24#',mpl,v24,v86,'#)"
  create=temp
mxcp temp create=temp1 "repeat=,65"
mxcp temp1 create=li2marc "repeat=<,59"
```

Detalle del proceso batch que genera la base li2marc, con lo campos 54 y 65 de LIBRI procesados.

Volcado de registros LIBRI en un archivo plano

Para esta etapa del proceso, se utiliza el programa MX el cual permite la manipulación de una base de datos Isis desde la “línea de comando” y permite la utilización del Lenguaje de Formateo Cisis [4].

Como el objeto de este archivo plano es que sea leído por el software MARCEdit [5], la estructura de cada renglón de este archivo se ajusta a lo requerido por el software para archivos con extensión “.MRK” [6]

Para el armado del archivo de salida se diseñó un “formato de salida de Isis” (li2marc.pft) y sobre este formato se realizan las conversiones necesarias en los campos. Esto es posible gracias a la potencia del utilitario MX de Bireme y al lenguaje de formateo extendido que utiliza.

Dentro del script encontramos campos que migran sin dificultad, como es el caso del campo 10 de LIBRI que se corresponde al campo 020 de MARC.

```
/*
010 -> 020
*/
if p(v10) then
  '=020 ##', ('$a',replace(v10^n,'-','),l (lv10^c)l)/
,fi,

/* fin 010 -> 020 */
```

Se migra el ISBN almacenado en el campo 10 de LIBRI al campo 20 de MARC

En otros casos la migración requiere el uso de lista de selección, como es el caso del idioma de la obra, para ubicar su código en el campo 008 de MARC. Para estas situaciones se utiliza la estructura Select Case de Cisis.

```

select v40[1]
case 'ES': 'spa'
case 'EN': 'eng'
case 'FR': 'fre'
....
elsecase s(v40, ' ').3
endsel,

```

De acuerdo al contenido de la primer ocurrencia del campo idioma de LIBRI, se convierte al código de idioma de MARC.

La misma estructura Select Case es de utilidad para determinar el “número de caracteres no significativos” en los campos en que este dato se debe almacenar en el indicador 2 de MARC. Lo que se implementó es una lista de artículos en varios idiomas, agrupados por su longitud y de esta manera se logra discriminar el contenido del indicador del campo.

<pre> select f(instr(s(mhu,v24^t,mhl), '),1,0) case '2': select s(mhu,v24^t,mhl).2 case 'A ': '2' case 'O ': '2' elsecase '0' endsel, case '3': select s(mhu,v24^t,mhl).3 case 'LA ': '3' case 'EL ': '3' case 'UN ': '3' case 'LE ': '3' case 'IL ': '3' case 'LO ': '3' case 'AN ': '3' case 'DI ': '3' case 'AS ': '3' case 'OS ': '3' case 'UM ': '3' case 'DE ': '3' elsecase '0' endsel, </pre>	<pre> case '4': select s(mhu,v24^t,mhl).4 case 'LOS ': '4' case 'LAS ': '4' case 'LES ': '4' case 'DES ': '4' case 'GLI ': '4' case 'THE ': '4' case 'DER ': '4' case 'DAS ': '4' case 'DES ': '4' case 'DEM ': '4' case 'EIN ': '4' case 'UMA ': '4' case 'UNS ': '4' elsecase '0' endsel, case '5': select s(mhu,v24^t,mhl).5 case 'UNAS ': '5' case 'UNOS ': '5' elsecase '0' endsel, elsecase '0' endsel, </pre>
---	--

```
MX li2marc
pft=@li2marc.pft
convert=ansi
lw=500
-all now >registros.mrk
```

La línea de comando utilizada para el volcado de los registros LIBRI al archivo plano con extensión .MKR

El resultado de la ejecución del proceso de volcado es un archivo plano, para nuestro ejemplo denominado **registros.mrk** que contiene la información de los registros LIBRI que se desean migrar.

```
10 «^n0-471-82173-X»
100 «^aGhezzi^ncarlo»
28 «^aJazayeri^nmehdi»
60 «^n005.13^e18 EN^bJUFEXA»
65 «^d<PROGRAMACION-LENGUAJES>^tLEMB.1»
86 «^tProgramming language concepts 2-E^rCarlo Ghezzi, Mehdi Jazayeri»
40 «EN»
43 «^INew York^eJ. Wiley»
45 «1987»
50 «^axvii, 428 p.^c24 cm.»
2 «^xJUFEXA^iJG»
3 «960509»
5 «M»
54 «^nBibliografía: p. 375-405»
54 «^nGlosario: p. 349-405»
54 «^nLa Biblioteca posee la versión en español»
70 «^i02301^p1^e1»
70 «^i02302^p1^e2»
75 «^bJUFEXA^n005.13^IG421pI»
```

Registro originalmente cargado en la base con formato LIBRI

```

=LDR 01324nama 22003255a 4500
=001 \000004
=003 \JUFEXA
=903 \JG
=005 20091102101002.1
=040 \aJUFEXA$bspa
=008 960509s1987 III 000 0 eng d
=020 \a047182173X
=245 10$aProgramming language concepts 2-E / $cCarlo Ghezzi, Mehdi Jazayeri.
=260 \aNew York : $bJ. Wiley, $c1987
=500 \aBibliografía: p. 375-405.
=500 \aGlosario: p. 349-405.
=500 \aLa Biblioteca posee la versión en español.
=082 0$a005.13$218 EN
=650 \7$aPROGRAMACION-LENGUAJES$2LEMB.1
=952 \p02301$e1$q1
=952 \p02302$e2$q1
=975 \n005.13$I421pI
=100 1$aGhezzi, Carlo.
=700 1$aJazayeri, Mehdi.
=300 \axvii, 428 p.$c24 cm.

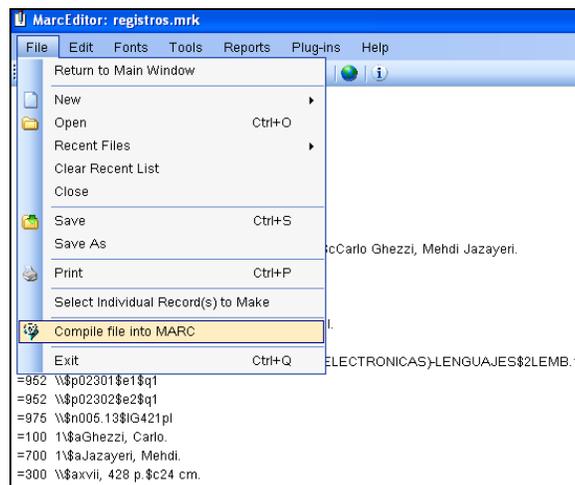
```

Archivo plano “registros.mrk” preprocesado y con etiquetas para MARCEdit

Procesamiento del archivo plano utilizando el software MARCEdit

Para obtener el ISO2709-MARC, se debe ingresar a MARCEdit y desde el editor del mismo cargar el archivo plano generado (**registros.mrk**).

Simplemente elegimos la opción “Compile File into MARC” del menú File, indicamos en nombre del archivo de salida (por ej. registros.mrc) y confirmamos la generación del mismo.



Pantalla de Editor de MARCEdit

Conclusión

Dado que en nuestro país no hay aún una estructura, un sistema de información, un sistema bibliotecario que permita que se tome una decisión común para todos, sino que estas surgen a nivel de las instituciones individualmente, consideramos oportuno usar el formato MARC, al que están adhiriendo gran cantidad de bibliotecas en el país.

Si bien al momento de la presentación de este informe, los trabajos propuestos para la migración se encuentran en la etapa de análisis y testeado por parte de las bibliotecas participantes del proyecto, podemos concluir que luego del estudio de los formatos involucrados, la migración de registros en formato LIBRI hacia MARC es posible, logrando que la totalidad de los datos almacenados en el formato local sean transferidos al internacional, respetando las estructuras de ambos y muy especialmente los requerimientos de puntos de acceso propuestos por LIBRI.

El código generado para la migración será publicado en la página del Instituto de Informática, www.idei.unsj-cuim.edu.ar, junto con la versión preliminar de la tabla de migración construida, para la descarga libre y gratuita.

Pasos futuros

El próximo objetivo a alcanzar es completar la tabla de migración con la totalidad de los campos que conforman LIBRI (descripción de analíticas, mapas, etc.).

La capacitación de recursos humanos en el uso de MARC es sin dudas otra de las tareas a futuro que serán encaradas por el equipo del proyecto..

Complementariamente se avanzará sobre el estudio de software de gestión bibliotecaria open source a fin de proponer un nuevo sistema para implementar en las bibliotecas afectadas al proyecto.

Referencias

- [1] http://www.grebyd.com.ar/r_marc.html
- [2] Migración de bases Isis a Koha. http://wiki.koha.org/doku.php?id=isis_a_koha
- [3] Utilitarios Cisis. <http://bvsmodelo.bvsalud.org/php/level.php?lang=es&component=28>
- [4] Lenguaje de formateo Cisis. http://productos.bvsalud.org/reference/pt/format34_es.htm
- [5] Software MarcEdit. <http://people.oregonstate.edu/~reaset/marcedit/html/index.php>
- [6] Archivo .MRK. <http://wiki.koha.org/doku.php?id=isisakoha#isis-text-marc-koha>

Bibliografía

- *Manual de MARC : cómo interpretar y usar registros MARC* / Déborah J. Byrne ; traducido y adaptado de la 2ª ed. en inglés en GREBYD por Nicolás Rucks. Buenos Aires : GREBYD, 2001. xxv, 270 p. ISBN 987-98052-2-4
- Universidad Nacional de San Juan. Facultad de Ciencias Exactas, Físicas y Naturales (1993). *Formato Libri* (2a. ed.). San Juan: la Facultad
- *Marc 21 bibliográfico: monografías impresas curso de educación a distancia* (2a. ed. 2003). Buenos Aires: Siu