

e-Ciencias de la Información

Revista electrónica publicada por la
Escuela de Bibliotecología y Ciencias de la Información,
Universidad de Costa Rica, 2060 San José, Costa Rica

e-Ciencias de la Información

Revista electrónica semestral, ISSN-1659-4142

Volumen 4, número 2, Informe Técnico 1

Julio - Diciembre, 2014

Publicado 1 de julio, 2014

<http://revistaebci.ucr.ac.cr/>

**Metodología para la migración de datos bibliográficos
entre programas de *software* de automatización: de
CEPAL a MARC**

José Iván Saborío-Acuña

Ricardo Chinchilla-Arley



Protegido bajo licencia Creative Commons
Universidad de Costa Rica

Metodología para la migración de datos bibliográficos entre programas de *software* de automatización: de CEPAL a MARC

Methodology for the migration of bibliographic data between automation software: from ECLAC to MARC

José Iván Saborío-Acuña¹, Ricardo Chinchilla-Arley²

RESUMEN

En la actualidad existen aún muchas bases de datos bibliográficas diseñadas en WinISIS sobre formato de intercambio de información CEPAL; sin embargo, la tendencia de los nuevos *software* de gestión de unidades de información es la utilización del formato MARC21, con el inconveniente de que ambos formatos son incompatibles entre sí. Por eso, el presente documento ofrece una guía para migrar del formato CEPAL al formato MARC21 utilizando una tabla de equivalencias de campos en una FST de exportación. Se construyó, además, una base de datos intermedia en WinISIS donde se vaciaron los datos convertidos. Posteriormente, se describe el proceso de conversión de un archivo de salida ISO2709 a un archivo MARC puro, que puede ser cargado desde cualquier sistema de gestión de unidades de información de última generación. También, se describe el proceso de depuración del mismo utilizando el programa MarcEdit y se realiza un ejemplo de la carga final en la aplicación de automatización EspaBiblio 3.2. Se concluye que todo proceso de migración debe responder a un proyecto previamente definido y la importancia de realizar la depuración para garantizar la pureza de los datos.

Palabras clave: automatización de bibliotecas, migración de datos, bases de datos, bases de datos textuales

ABSTRACT

At present there are still many bibliographic databases designed in WinISIS on ECLAC exchange format information, however, the trend of new software management information units is the use of MARC21 format, with the drawback that both formats are incompatible with each other. This document provides guidance for migrating from ECLAC format to MARC21 format using a conversion table of fields in a FST export. It also built an intermediate database in which the converted data WinISIS emptied. Subsequently, the process of converting an output file ISO2709 to a file pure MARC, which can be loaded from any system of information management units described generation. It also describes the process of debugging the same program using the MARCEdit and an example of the final charge is made in the application of automation EspaBiblio 3.2. It is concluded that the process of migration to respond to a previously defined project and the importance of debugging to ensure the purity of the data.

Keywords: library automation, data migration, databases, textual databases

Fecha de recibido: 31 de marzo del 2014

Fecha de aprobado: 12 de mayo del 2014

Fecha de corregido: 03 de junio 2014

¹Universidad de Costa Rica, Escuela de Bibliotecología y C.I. JOSE.SABORIOACUNA@ucr.ac.cr

²Universidad de Costa Rica, Escuela de Bibliotecología y C.I. RICARDO.CHINCHILLA@ucr.ac.cr

1. INTRODUCCIÓN

Hasta hace muy poco tiempo, el *software* de automatización de catálogos WinIsis se ha utilizado como estándar para el manejo de la descripción bibliográfica en muchas unidades de información. Por su parte y para el caso concreto de América Latina, la Comisión Económica para América Latina y el Caribe (CEPAL) tuvo en la década de los 80 y 90 del siglo pasado una participación sumamente activa en el ámbito de la distribución del WinIsis y el intercambio de información bibliográfica.

En primera instancia, la CEPAL desarrolló su formato de intercambio de información (conocido como formato CEPAL) para uso interno. Sin embargo, se detectó la problemática surgida por el diseño libre y desordenado de bases de datos bibliográficas sin formato alguno, lo cual generaba incompatibilidades en el intercambio de información. Por tal motivo y como apoyo a las bibliotecas de América Latina, se distribuyó el formato CEPAL junto con el paquete WinIsis, por lo cual se utilizó de manera extensiva (Ugobono, 2011).

No obstante, debido a los cambios y demandas actuales de los usuarios, ahora se requieren sistemas de automatización modernos, ágiles, modulares y con posibilidades de crecimiento bajo estándares internacionales. Por lo tanto, el WinIsis ya ha dejado de ser una opción. De hecho, hay que recordar que WinIsis es una sistema de primer nivel de automatización (Chinchilla, 2011), no es multiplataforma y su desarrollo se detuvo en la versión 1.5 (solamente corre sobre las antiguas versiones de Windows NT/2000/XP). Esto hace necesario el evaluar otras alternativas para los nuevos proyectos de automatización y para el desarrollo tecnológico de las unidades de información ya automatizadas.

Sin bien es cierto que el formato CEPAL es aún de uso extensivo, los nuevos sistemas de automatización integral de unidades de información han apostado por el formato MARC (MACHINE-Readable Cataloging), desarrollado por la Biblioteca del Congreso de los Estados Unidos de América y que ya va por la versión 21. Con este panorama nos enfrentamos al hecho de que si deseamos migrar a una nueva plataforma informática, es necesario comenzar por migrar de formato de intercambio de información.

Aunque, algunos sistemas como SIABUC© o aplicaciones basadas en WinIsis como ABCD³ pueden cargar archivos tipo ISO 2709, los archivos MRC pueden ser cargados por cualquier aplicación de automatización de unidades de información. Para ofrecer un ejemplo de este último punto, se ejemplifica dicha carga utilizando el programa EspaBiblio 3.2⁴.

Entonces, el objetivo de este documento es atender la problemática descrita, proponiendo una metodología de migración. En primera instancia, se explica cómo realizar la migración de datos del formato CEPAL al MARC utilizando una tabla de equivalencias. En segunda, se explica cómo utilizar la aplicación MarcEdit 5.9⁵ para transformar el

³<http://reddes.bvsaude.org/projects/abcd>

⁴<http://sourceforge.net/projects/espabiblio/>

⁵<http://marcedit.reeset.net/>

archivo ISO 2709⁶ generado por el proceso de exportación de WinIsis en un archivo tipo MRC (MARC puro) y depurarlo. Una vez limpio, este archivo puede ser cargado por el sistema de automatización de nuestra elección.

2. PROCESO DE MIGRACIÓN

Se debe recordar que al momento de diseñar una bases de datos en WinIsis, deben definirse cuatro archivos de trabajo:

1. Tabla de definición de campos (FDT): aquí se muestra el número y nombre de cada uno de los campos sobre los cuales se trabajará, así como una serie de variables para cada campo, como la definición de los subcampos, si son repetibles o no y si tienen un patrón definido para el ingreso de la información. Es aquí donde se define el formato de intercambio sobre el que se trabajará.
2. Hoja de entrada de datos (FMT): se construye la hoja sobre la cual se digitarán los datos bibliográficos que serán ingresados a la base de datos.
3. Formato de visualización (PFT): en este archivo se diseña la forma como se verán los datos en pantalla o impresos.
4. Tabla de selección de datos (FST): en este archivo se indica cuales serán los campos que servirán como puntos de entrada a los registros.

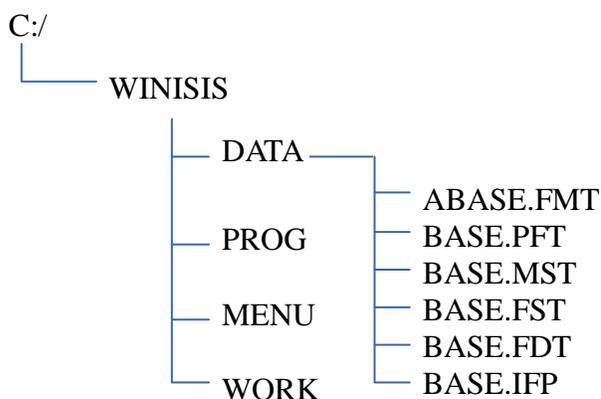
Una vez diseñada la base de datos, se generan de forma automática varios archivos, dentro de los cuales se destacan los siguientes:

- Archivo Maestro (MST): archivo que contiene todos los registros ingresados a la base de datos.
- Archivo Invertido o Diccionario de Datos (IFP): archivo de punteros generado por los campos señalados en el FST previamente definido.

Además, WinIsis controla cada uno de los registros asignando un número consecutivo a cada uno de ellos, denominado Número Maestro (MFN). Es importante especificar que este número no funciona como una llave primaria de cada registro, es simplemente un número consecutivo que se acomodará cada vez que se borren e ingresen nuevos registros.

Todos los archivos mencionados son almacenados en el directorio DATA de WinIsis. La estructura de almacenamiento de todo el programa es la siguiente:

⁶ISO 2709 es el estándar de la International Standar Organization para la descripción bibliográfica. Para mayor información consultar http://es.wikipedia.org/wiki/ISO_2709



En el directorio PROG se encuentran los programas ejecutables, en MENU todos los menús del sistema y en el directorio WORK se almacenan todas las salidas, como impresiones, exportaciones y demás. Este último directorio es muy importante para el proceso que se describirá más adelante.

2.1 Proceso de conversión de CEPAL a MARC en WinIsis

Para este proceso, se utilizan dos bases de datos desarrolladas en WinIsis: la primera con los campos más utilizados del formato CEPAL⁷ y la segunda con los campos equivalentes en el formato MARC. Se construye, además, una FST con las equivalencias de campos y que funcionará como puente entre ambas bases. Si la base de datos a migrar tuviese más campos, es posible ampliar la tabla de equivalencias.

La primera base de datos tiene la FDT tal y como se muestra en la tabla siguiente. Además, se debe tener presente que el formato CEPAL no utiliza subcampos (ver tabla 1).

⁷Para una lista completa de los campos del formato CEPAL se puede consultar el documento Vera, C (2003) Sistema de Información Bibliográfica de la CEPAL: manual de referencia. Santiago de Chile: CEPAL.

Tabla 1
Estructura de campos en formato CEPAL

NOMBRE DEL CAMPO	ETIQUETA
NOMBRE DEL ARCHIVO	5
TIPO DE LITERATURA	10
NIVEL BIBLIOGRAFICO	14
NUMERO DE ACCESO	20
UBICACION FISICA	30
IDIOMA	34
AUTOR PERSONAL	40
AUTOR INSTITUCIONAL	50
TITULO	60
AUTOR PERSONAL ANALITICO	62
AUTOR INSTITUCIONAL ANALITICO	63
SERIE	64
NUMERO DE SERIE	65
NOMBRE Y # DE CONFERENCIA	70
PROYECTOS	75
CIUDAD - CONFERENCIA	90
PAIS – CONFERENCIA	100
FECHA - CONFERENCIA	110
VOLUMEN	130
NUMERO	135
MES	137
ISSN	145
ISBN	147
CIUDAD DE PUBLICACION	160
PAIS DE PUBLICACION	170
EDITORIAL	180
EDICION	190
FECHA DE PUBLICACION	200
DESCRIPTOR GEOGRAFICO	201
# DE PAGINAS	220
NOTAS	230
DESCRIPTOR PRINCIPAL	234
DESCRIPTORES	240
RESUMEN	250

Continúa...

NOMBRE DEL CAMPO	ETIQUETA
DOCUMENTALISTA	260
FECHA	270
DIGITADOR	280
TABLA DE CONTENIDO	290
TITULO A NIVEL ANALITICO	300
TITULO A NIVEL SERIADO	320
ACTIVO	330
TEXTO COMPLETO	400

Fuente: Elaboración propia.

Por su parte, la base de datos sobre la que se descargarán los datos a migrar se diseña utilizando MARC. También, debe tenerse en cuenta que toda migración es propensa a la pérdida de datos si no existe el campo o subcampo equivalente en el formato al que se va a migrar. Por eso, la Tabla 2 muestra la estructura de campos.

Tabla 2
Estructura de campos y subcampos en formato MARC

NOMBRE DEL CAMPO	ETIQUETA	SUBCAMPO
Nivel bibliográfico	7	^a
Tipo de literatura	8	^a
ISBN	20	^a
Idioma	41	^a
Numero de acceso	90	^a
Autor personal	100	^a
Autor Institucional	110	^a
Nombre y # de conferencia	111	^a^c^d
Título	245	^a
Fecha de publicación	250	^a
Editorial	260	^a^b^c
# de páginas	300	^a
Volumen	362	^a
Título a nivel seriado	440	^a^v
Notas	500	^a
Tabla de contenido	505	^a
Resumen	520	^a
Descriptor general	650	^a
Categoría Geográfica	651	^a
Ubicación física	852	^a
Acceso electrónico	856	^d

Fuente: Elaboración propia.

Se podrá notar que no se incluyen algunos subcampos, como el ^d del campo 100 o los subcampos adicionales del campo 245, esto debido a que no hay un campo equivalente en la Tabla 1. Se insiste en que toda migración es propensa a la pérdida de datos.

Revisadas ambas estructuras, se procede a construir la tabla de equivalencias. Para ello, se desarrolla una FST con la siguiente estructura:

Tabla 3
FST como tabla de equivalencias entre MARC y CEPAL

ETIQUETA	TECNICA	FORMATO
1	0	Mfn
007	0	"00^a"v14
008	0	"00^a"v10
020	0	"##^a"v147
022	0	"##^a"v145
041	0	"##^a"v34
090	0	"##^a"v20
100	0	(1#^a v40)
110	0	"2#^a"v50
111	0	"2#^a"v70,"^c"v90,v100,"^d"v110
245	0	"10^a"v60,v75
250	0	"##^a"v190
260	0	"##^a"v160,v170"##^b"v180,"##^c"v200
300	0	"##^a"v220
362	0	"1#^a"v130,v135,"1#^a"v137
440	0	"1#^a"v64, "^v"v65
500	0	(##^a v230/)
505	0	(0#^a v290/)
520	0	"##^a"v250
650	0	(#4^a v234/), (#4^a v240/)
651	0	(4#^a v210/)
852	0	"4#^a"v30
856	0	"##^d"v400

Fuente: Elaboración propia.

Esta FST presenta las siguientes características:

- En la columna de la izquierda contiene el identificador de campo de entrada MARC donde se almacenarán los datos.
- En la columna central se muestra la técnica de indización. En el caso concreto debe ser tipo 0, que seleccionará el contenido completo del campo.
- En la columna de la derecha se señala el formato de extracción que contiene los indicadores de cada campo, el campo de salida en formato CEPAL que contiene los datos a transferir y los subcampos MARC.

La función de los indicadores la describe la Biblioteca del Congreso de los Estados Unidos:

Indicadores: De las dos posiciones de caracteres que le siguen a cada etiqueta (con excepción de los campos 001 al 009), una o ambas pueden estar ocupadas por indicadores. En algunos campos se utiliza únicamente la primera o la segunda posición; en otros campos se usan las dos, y en algunos como el 020 y el 300 no se usa ninguna. Cuando una posición de indicador no se usa se dice que "no está definida", y dicha posición se deja en blanco. Por regla convencional se representa a los espacios dejados en blanco en los indicadores (no definidos) mediante el símbolo "#" (Biblioteca del Congreso, 2003, párr 31).

Los indicadores son dos dígitos que van del 0 al 9 y no se leen en forma conjunta sino en forma individual, debido a que son dos números individuales. Como se muestra en el ejemplo siguiente (Figura 1), el número 245 corresponde al campo de título, el 13 representa los indicadores (Biblioteca del Congreso, 2003).

Figura 1
Ejemplo de uso de indicadores

245 13 \$a La cordillera andina / \$c Ildefonso Ruiz Calvo; prólogo de Eva Sarasola Ibarra.

Fuente: Elaboración propia a partir de ejemplos de la Biblioteca del Congreso (2003).

El primer indicador es el 1, el cual señala que el título de la obra es un asiento secundario; si fuese 0 significa que el documento no tiene autor o son más de 3 y el título entra como punto de acceso principal.

El segundo indicador es 3, que le señala a la base de datos donde debe iniciar la alfabetización del título, es decir, si es un 3 debe ubicarse en el cuarto campo del título para el acomodo alfabético del mismo, estos 3 campos incluyen espacios en blanco para que la computadora no tome en cuenta dichos caracteres en el proceso de ordenamiento alfabético. En el título *La cordillera andina* el valor del segundo indicador es 3, de manera que los primeros tres caracteres (la "L" la "a," y el espacio) serán ignorados y el título será alfabetizado bajo la letra "c" de "*cordillera andina*".

Es importante recordar que ninguna migración es completamente efectiva y siempre habrá algún porcentaje de pérdida de datos. Esto se debe a que, por su propia naturaleza, no existe uniformidad en la información bibliográfica ingresada y no es posible prever todas las posibles combinaciones que pudiesen presentarse. Esto se explica en el siguiente ejemplo:

Etiqueta	Técnica	Formato
245	0	"10^a"v60,v75

El indicador 1 se utiliza cuando el título tiene entrada secundaria, es decir cuando hay un autor responsable del documento, pero si hay documentos que entran por título, entonces el indicador debería ser 0. El segundo indicador es 0, el cual se utiliza cuando el título no tiene artículos, pero siempre habrán algunos que si los tienen.

Por su parte, el subcampo ^a es para el título propiamente dicho. Sin embargo es necesario contemplar el subcampo ^b para el resto del título, el ^c para la mención de responsabilidad y el ^h para la Designación General del Material (DGM). Si la base de datos original tiene estos subcampos, entonces es necesario agregarlos en la tabla de equivalencias.

Una vez construida la FST de exportación, la que para este ejercicio llamaremos MIGRA.FST, se procede a exportar la base de datos en CEPAL. El archivo de salida, el cual es generado por WinIisis utilizando el estándar ISO2709 como se mencionó con antelación, se denominará MIGRA.ISO (Ver Figura 2). Los parámetros que se deben mostrar en la ventana de diálogo de exportación son los siguientes:

- Nombre del archivo ISO de salida: MIGRA.ISO
- Intervalo desde 1 hasta 999999 (toda la base de datos)
- Longitud de línea de salida (en algunas versiones aparece Caracteres por línea): 0
- Dirección física de la FST de reformateo: C:\WINISIS\DATA\MIGRA.FST

Figura 2

Ventana de exportación en WinIsis utilizando el estándar ISO2709

Fuente: WinIsis 1.5

Una vez ejecutada la exportación, se generará el archivo MIGRA.ISO en el directorio WORK de WinIsis.

Ahora es necesario salir de la base de datos de CEPAL e ingresar a la base de datos diseñada sobre MARC señalada en la Tabla 3. Una vez dentro se procede a importar el archivo ISO generado con antelación. En la ventana de diálogo de importación es necesario demostrar lo siguiente:

- Archivo ISO de entrada: MIGRA.ISO
- Primer MFN a ser asignado: 1
- En las opciones de importación que se ofrecen, se debe seleccionar CARGAR. Con esto se asegura que se eliminará cualquier registro previo existente.
- Longitud de línea de entrada (o caracteres por línea): 0

Figura 3
Ventana de importación en WinIisis utilizando ISO2709

Fuente: WinIisis 1.5

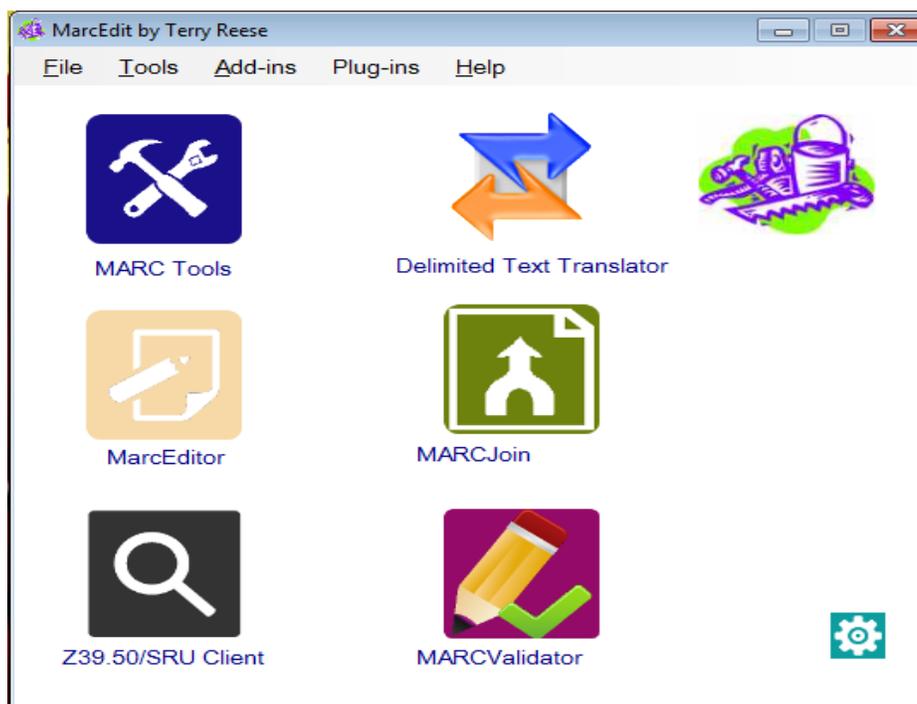
Una vez realizado el proceso, se pueden verificar los registros ingresados a la base y posteriormente se procede a exportar la base de datos completa una vez más. En esta oportunidad en la ventana de diálogo de exportación se indica el nuevo nombre del archivo de salida (para este ejercicio será MARC.ISO) y los caracteres por línea nuevamente serán 0 y ya no es necesario señalar la FST de reformato. Este proceso generará el archivo indicado y lo almacenará nuevamente en el directorio WORK.

2.2 Depuración de datos utilizando MarcEdit

Los datos almacenados en el archivo ISO generado en el proceso anterior siempre tendrán algunas inconsistencias y errores, sobre todo con el manejo de tildes y demás signos propios, en este ejemplo concreto, del idioma español. Por tal motivo, es necesario realizar alguna limpieza directamente sobre cada uno de los registros contenidos en el archivo de salida. Para depurar dichos datos se utiliza MarcEdit⁸, un programa multiplataforma de acceso libre para editar y corregir registros MARC.

⁸La aplicación puede ser descargada de forma libre en <http://marcedit.reeset.net/downloads>

Figura 4
Ventana principal de MarcEdit 5.9



Fuente: MarcEdit 5.9

Como primer paso, debemos transformar el archivo ISO en un archivo MRC. Para ello, una vez abierto el MarcEdit, se ingresa al menú **Add-ins** y se selecciona la opción **CDS-Isis.iso – Marc Translation**. Se selecciona el archivo de entrada (marc.iso) que se encuentra en la carpeta WORK de WinIsis y se muestra el nombre del archivo de salida. Este ejercicio se llamará openbiblio.mrc y se guardará en c:\winisis\work\ Una vez señalados los archivos correspondientes, se presiona el botón **OK**, la pantalla “parpadea” una vez y se procede a presionar el botón **Close** (Figura 5).

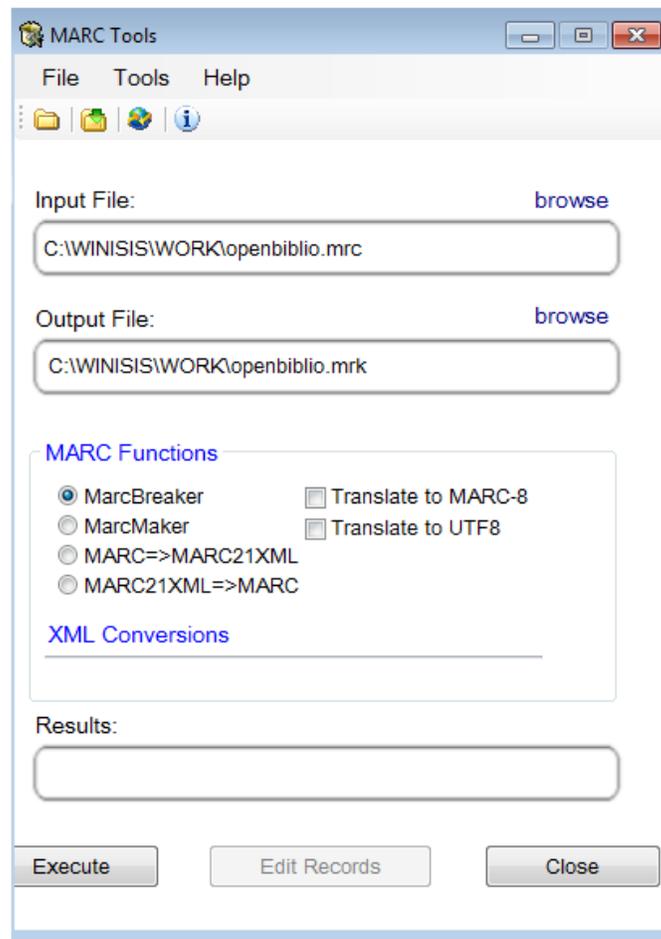
Figura 5
Ventana de MARC Translation



Fuente: MARC Translation.

La acción anterior ha generado un archivo tipo MRC que es necesario depurar. Para ello, se abre el Marc Tools para convertirlo en a la extensión MRK con la finalidad de editarlo y depurarlo como se muestra en la Figura 6.

Figura 6
Ventana de MarcTools



Fuente: MarcEditor 5.9

Una vez convertido el archivo con la extensión mrk, se abre con MarcEditor, que consiste en un editor de texto plano. En dicho editor se selecciona el menú **File** y la opción **Open**. Se busca el archivo openbiblio.mrk generado en el paso anterior y se abre para proceder a su limpieza (Figura 7).

Figura 7
Archivo MRK abierto en MarcEditor

```

=LDR 0076800000000002170004500
=001 000001
=007 00$aM
=041 ##$aEs
=090 ##$a01
=110 2#$aInstituto Costarricense de Turismo
=245 10$aProyecto turistico en bahia culebra: Estudio de factibilidad, econ{Ostrok}micos y
finaniero, plan maestro y dise{THORN}os preliminares
=260 ##$aSan Jos{82}Costa Rica##$bChecchi & Co##$c1978
=300 ##$a327 p.
=505 0#$aAnálisis de mercado y estrategia de mercadeo, introducci{Ostrok}n orientadora y
supuestos basicos, La base del turismo en Costa Rica, Tendencia del turismo en areas
competitivas
=650 #4$aBAHIA CULEBRA
=650 #4$aPROYECTO TURISTICO
=650 #4$aPAPAGAYO
=650 #4$aINSTITUTO NACIONAL DE TURISMO
=651 4#$aCOSTA RICA
=651 4#$aPNUMA
=852 4#$aParques

=LDR 0066800000000002290004500
=001 000002
=007 00$aM
=008 00$aM
=041 ##$aEs
=090 ##$a03
=100 1#$aChaves, Ana Cecilia
=110 2#$aMinisterio de Vivienda y Asentamientos Humanos
=245 10$aCondiciones previas al desembolso del Banco Centroamericano para la fase
C:\Documents and Settings\Openbiblio\Configuración local\Temp\tmp11.tmp loading in 0,00000(

```

Fuente: MarcEditor 5.9

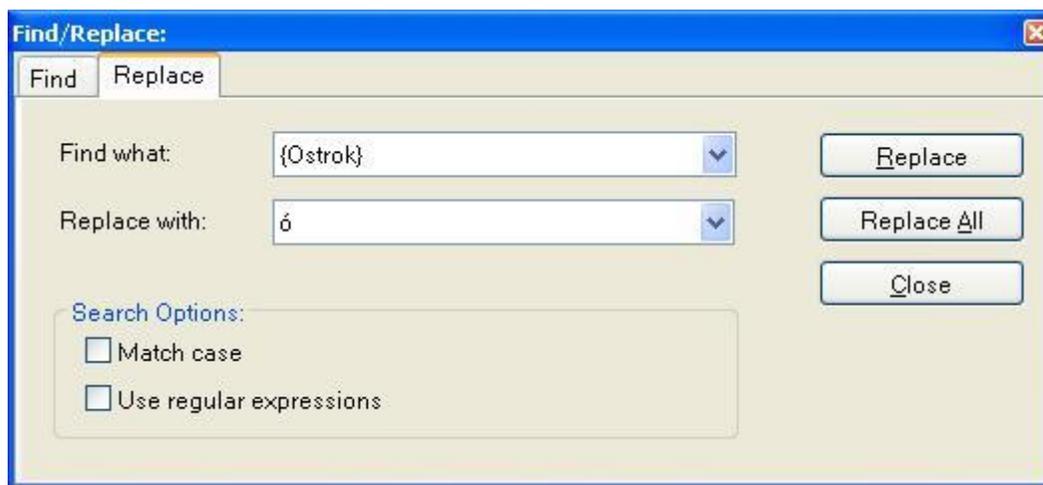
Este proceso requiere de mucho cuidado y dedicación. Por lo tanto, es importante tener presentes los caracteres básicos que deberán ser tomados en cuenta para su respectiva recuperación, que no fueron reconocidos en el proceso de conversión y que se sustituyeron por grupos de caracteres encerrados en llaves {}. A manera de ejemplo tendríamos los siguientes: Á, á, É, é, Í, í, Ó, ó, Ú, ú, ü, Ñ, ñ.

Es importante indicar que, si se tuviesen registros en otros idiomas diferentes del inglés, se deben incluir los caracteres propios de dicho idioma. Después de una rápida revisión por los primeros registros, se puede ir formando una tabla de conversión de la siguiente forma:

á {A0}	Á {aelig}	ñ {THORN}
é {82}	É {90}	Ñ {AElig}
í {Lstrok}	Í {D6}	ı {middot}
ó {ostrok}	Ó {hooka}	ı̇ {Uhorn}
ú {Dstrok}	Ú {caron}	ü {81}

Una vez identificadas las series de caracteres a convertir, se selecciona en el menú **Edit** la opción **Replace**, se indica la serie de caracteres a convertir en la primera línea de diálogo, el caracter que se desea recuperar en la segunda línea de diálogo y se selecciona **Replace All**. Se presiona **Accept** y el editor se encargará de realizar las conversiones necesarias a lo largo del documento.

Figura 8
Ventana de diálogo de la opción Replace en MarcEdit



Fuente: MarcEdit 5.9

Hay que recordar que es necesario salvar el documento cada vez que se realiza una conversión, con el fin de evitar pérdidas de trabajo por posibles imprevistos. Al final de la labor de limpieza tendremos un archivo MARC listo para ser cargado por cualquier sistema de automatización de unidades de información.

2.3 Carga de datos en un sistema de automatización de unidades de información

Para ejemplificar la carga de un archivo MARC en un sistema de automatización de unidades de información, se utilizará EspaBiblio 3.2⁹. Se escogió dicho programa por ser un *software* de tercer nivel de automatización (Chinchilla y Fernandez, 2012), pero de complejidad mediana (Fernández, 2013), lo cual facilita la ejecución y comprensión del ejemplo.

⁹EspaBiblio es *software* libre (GNU General Public License) basada en OpenBiblio y puede descargar la aplicación en <http://sourceforge.net/projects/espabiblio/>

Como primer paso, es necesario ingresar al sistema y autenticarse como usuario administrador, a fin de tener los permisos necesarios para poder realizar los cambios pertinentes. Posteriormente se debe revisar que los campos contenidos en la hoja de trabajo *book* sean los mismos que contiene visibles la tabla 3, a fin de verificar la correcta descarga de los datos. Para ello basta con ingresar al menú **Catalogación** y seleccionar **Nuevo material**. Por defecto se encuentra seleccionada la hoja de trabajo *book*, correspondiente a libros.

Figura 9
Ventana principal de EspaBiblio 3.2

The screenshot shows the main interface of EspaBiblio 3.2. At the top, there is a header with a logo of books and the text 'Your Library Name'. To the right, there is a box containing library information: 'Fecha actual: Jan 08, 2014', 'Horario de servicio: M-F 8am-9pm, Sa noon-5pm, Su 1-5pm', 'Dirección de la Biblioteca: stree 15', and 'Teléfono de la Biblioteca: 111-222-3333'. Below the header is a navigation menu with tabs: 'Inicio', 'Prestamos', 'Catalogación', 'Administración', 'Reportes y estadísticas', and 'Catálogo Público en Linea'. The 'Catalogación' tab is active. On the left side, there is a sidebar menu with options like 'Salir', 'Buscar material', 'Nuevo material', 'Eliminar Bibliografías por lotes', 'Búsqueda en línea Z39.50', 'Búsqueda masiva de datos via ISBN', 'Importar desde un archivo CSV', 'Subir datos Marc', 'Consultar Biblioteca del Congreso USA', 'Tablas CDU', 'Tablas CDD', and 'Tablas IBIC'. The main content area shows a form titled 'Añadir nuevo Bibliografía:'. The form has a header 'Los campos marcados con * son requeridos.' and several fields: '* Tipo de material:' (dropdown menu with 'book' selected), '* Colección:' (dropdown menu with 'Adult Nonfiction' selected), '* Número de entrada:' (three empty text input boxes), 'Mostrar en OPAC:' (checkbox checked), 'Campos USMarc:' section with '* 245 Title:' (text input), '245 Remainder of title:' (text input), '245 Statement of responsibility, etc.:' (text input), '* 100 Personal name:' (text input), and '650 Topical term or geographic name as entry element:' (text input).

Fuente: EspaBiblio 3.2

Si falta algún campo, es necesario agregarlo. Para hacerlo se siguen los siguientes pasos:

1. Dentro del menú **Administración** se selecciona **Tipos de Material**, donde aparecerán las distintas hojas de trabajo.
2. Se selecciona **Campos Marc** en la línea *book*, como se muestra en la siguiente figura:

Figura 10
Ventana de Administración de EspaBiblio. Opciones de tipo de material

Administración

Tipos de material

Añadir nuevo tipo de material

Tipos de material:

*Función	Descripción	Archivo Imagen	Número libros
Editar Eliminar Campos MARC	audio tapes	tape.gif	0
Editar Eliminar Campos MARC	book	book.gif	0
Editar Eliminar Campos MARC	cd audio	cd.gif	0
Editar Eliminar Campos MARC	cd computer	cd.gif	0
Editar Eliminar Campos MARC	equipment	case.gif	0
Editar Eliminar Campos MARC	magazines	mag.gif	0
Editar Eliminar Campos MARC	maps	map.gif	0
Editar Eliminar Campos MARC	video/dvd	camera.gif	0

Fuente: EspaBiblio 3.2

3. y, posteriormente, da *click* sobre la única opción que aparece en la parte superior: **Agregar un campo MARC para este tipo de materiales (book)**. Como se muestra en la siguiente figura:

Figura 11
Ventana de Administración de EspaBiblio. Agregar campos MARC a un tipo de material

Administración

Tipos de material

Agregar un campo MARC para este tipo de materiales (book)

Función	etiqueta [Tag]	Código de subcampo	Descripción	Requerido?
Editar Eliminar	007	a	007 Nivel bibliografico	FALSO
Editar Eliminar	008	a	008 Tipo de literatura	FALSO
Editar Eliminar	090	a	090 Numero de acceso	FALSO
Editar Eliminar	110	a	610 Autor Institucional	FALSO
Editar Eliminar	111	a	111a Nombre de la Conferencia	FALSO
Editar Eliminar	111	c	111c Lugar de la Conferencia	FALSO
Editar Eliminar	111	d	111d Fecha de la Conferencia	FALSO
Editar Eliminar	111	n	111n Numero de Conferencia	FALSO
Editar Eliminar	22	a	022 International Standard Serial Number (ISSN)	FALSO
Editar Eliminar	250	a	250 Edición	FALSO
Editar Eliminar	362	a	362 Volumen	FALSO
Editar Eliminar	41	a	041 Idioma del texto	FALSO
Editar Eliminar	440	a	440 Mención de Serie	FALSO
Editar Eliminar	505	a	505 Nota de Contenido	FALSO

Fuente: EspaBiblio 3.2

4. Esta opción permite el poder agregar los campos faltantes, esquematizados con antelación en la Tabla 3. Por ejemplo, si faltase el campo con la etiqueta 505, simplemente lo agregamos, al igual que el subcampo correspondiente y su descripción. Es posible indicar, también, si dicho campo es requerido o no. Esto es importante, ya que si se indica que sí es requerido (**true**), el sistema mostrará que es un campo obligatorio. En este caso particular, dicho campo no es obligatorio, por lo que se señala **FALSE**. Por último, el tipo de control, que consta del campo de texto para unas cuantas palabras o la otra opción llamada área de texto, que se utiliza para escribir mucha más palabras que la anterior opción (Figura 12).

Figura 12
Ventana de Administración de EspaBiblio32. Diálogo de agregación de campos

Inicio	Prestamos	Catalogación	Administración	Reportes y estadísticas																
<div style="display: flex; justify-content: space-between;"> Salir </div>																				
<ul style="list-style-type: none"> ▶ Administración ▶ Administración de bibliotecarios ▶ Configuración de la biblioteca ▶ Grupos de usuarios ▶ Grados de usuario ▶ Personalización de Campos de datos para las copias 																				
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 50%;">Campo</th> <th style="width: 50%;">Valor</th> </tr> </thead> <tbody> <tr> <td>Tipn de material</td> <td>hnnk</td> </tr> <tr> <td>etiqueta [Tag]</td> <td>505 Seleccionar</td> </tr> <tr> <td>Código de subcampo</td> <td>a</td> </tr> <tr> <td>Descripción</td> <td>Tabla de contenido</td> </tr> <tr> <td>Requerido?</td> <td>FALSO ▼</td> </tr> <tr> <td>Tipo de control</td> <td>Campo de texto ▼</td> </tr> <tr> <td colspan="2" style="text-align: center;"> Enviar Cancelar </td> </tr> </tbody> </table>					Campo	Valor	Tipn de material	hnnk	etiqueta [Tag]	505 Seleccionar	Código de subcampo	a	Descripción	Tabla de contenido	Requerido?	FALSO ▼	Tipo de control	Campo de texto ▼	Enviar Cancelar	
Campo	Valor																			
Tipn de material	hnnk																			
etiqueta [Tag]	505 Seleccionar																			
Código de subcampo	a																			
Descripción	Tabla de contenido																			
Requerido?	FALSO ▼																			
Tipo de control	Campo de texto ▼																			
Enviar Cancelar																				

Fuente: EspaBiblio 3.2

Para realizar la carga de datos, se selecciona Subir datos Marc dentro del menú Catalogación. Con el fin de verificar los registros, se puede realizar una carga de prueba, para lo cual se deberá marcar la casilla **Verdadero**. Si se desea realizar la carga final, se debe marcar la casilla **Falso**. Para la entrada de datos, se debe seleccionar el archivo openbiblio.mrc generado en el punto anterior por MarEdit y que se encuentra en la carpeta WORK de WinIisis (c:\winisis\work\openbiblio.mrc). Se selecciona la codificación de caracteres (UTF-8 es la codificación para español), el tipo de la colección donde se ubicará el material, el tipo de material (**book**) y si se desea que los registros puedan ser visualizados en el catálogo en línea del sistema (**Opac**). Al finalizar esta opción se da *click* sobre el botón **Cargar Archivo Almacenado** (Figura 13).

Figura 13
Ventana de Catalogación de Espabiblio. Subir datos MARC

Inicio	Prestamos	Catalogación	Administración	Reportes y estadísticas	Catálogo Público en Línea
<p>Salir</p> <ul style="list-style-type: none"> » Buscar material » Nuevo material » Eliminar Bibliografías por lotes » Búsqueda en línea Z39.50 » Búsqueda masiva de datos vía ISBN » Importar desde un archivo CSV » Subir datos Marc » Consultar Biblioteca del Congreso USA 					
<p>Nota: Recomendamos el tamaño del archivo menor de 10 MB. Archivos mas grandes de registros MARC puede provocar errores al analizar los datos .</p> <p>Prueba de importación: Verdadero <input checked="" type="radio"/> Falso <input type="radio"/></p> <p>Entrada de archivo USMarc: C:\WINISIS\WORK\openbiblio.mrc <input type="button" value="Examinar..."/></p> <p>codificación de caracteres: UTF-8 <input type="button" value="v"/></p> <hr/> <p>por defecto:</p> <p>Colección: Adult Nonfiction <input type="button" value="v"/></p> <p>Tipo de material: book <input type="button" value="v"/></p> <p>Mostrar en OPAC: Si <input type="button" value="v"/></p> <p style="text-align: center;">Cargar Archivo Almacenado</p>					

Fuente: Espabiblio 3.2

Si se realizó la carga de prueba, se visualizarán los registros para realizar una verificación del archivo MRC, tal y como lo muestra la siguiente figura. De esta forma es posible realizar cambios sobre el mismo antes de efectuar la carga final (Figura 14).

Figura 14
Ejemplo de salida de verificación

Registros MARC		
Etiqueta	Sub	Datos
041	a	Es
090	a	01
110	a	Instituto Costarricense de Turismo
245	a	Proyecto turistico en bahia culebra: Estudio de factibilidad, económicos y financiero, plan maestro y diseños preliminares
260	a	San JoséCosta Rica###
260	b	Checchi & Co###
260	c	1978
300	a	327 p.
505	a	Análisis de mercado y estrategia de mercadeo, introducción orientadora y supuestos básicos, La base del turismo en Costa Rica, Tendencia del turismo en áreas competitivas
650	a	BAHIA CULEBRA
650	a	PROYECTO TURISTICO
650	a	PAPAGAYO
650	a	INSTITUTO NACIONAL DE TURISMO
651	a	COSTA RICA
651	a	PNUMA
852	a	Parques

Fuente: EspaBiblio 3.2

Una vez realizado el procedimiento nuevamente marcando la casilla **Falso**, se cargarán los datos en la base de datos del sistema de forma definitiva. Finalizado el proceso, el sistema enviará un mensaje con los datos técnicos del archivo almacenado similar al siguiente:

This data is not uniconed string & would be converted.

Technical information:

Detected encoding: SJIS

Conversion: UTF-8 > UTF-8

Registros transferidos: 34

Ahora es posible realizar búsquedas, ediciones, controles de copias y demás sobre los registros almacenados desde el EspaBiblio.

Figura 15

Ventana de Catalogación de EspoBiblio. Visualización de la información bibliográfica de los registros cargados

Inicio	Prestamos	Catalogación	Administración	Reportes y estadísticas	Catálogo Público en Línea												
<p>Salir</p> <ul style="list-style-type: none"> ▶ Buscar material ▶ Información bibliográfica ▶ Edición básica ▶ Editar-MARC ▶ History ▶ Nueva copia ▶ Reservas ▶ Borrar ▶ ¿Usar datos para un nuevo registro? ▶ Nuevo material ▶ Eliminar Bibliografías por lotes ▶ Búsqueda en línea Z39.50 ▶ Búsqueda masiva de datos vía ISBN ▶ Importar desde un archivo CSV ▶ Subir datos Marc 																	
<p>Información bibliográfica:</p> <p>Tipo de material: book</p> <p>Colección: Adult Nonfiction</p> <p>Número de entrada:</p> <p>245 Title: Proyecto turistico en bahia culebra: Estudio de factibilidad, económicos y financiero, plan maestro y diseños preliminares</p> <p>245 Remainder of title:</p> <p>100 Personal name:</p> <p>245 Statement of responsibility, etc.:</p> <p>Mostrar en OPAC: Si</p> <p>Añadir nueva copia</p> <p>Información de la copia bibliográfica:</p> <table border="1"> <thead> <tr> <th>Función</th> <th>Código de barras #</th> <th>Descripción</th> <th>Estado</th> <th>Estado Dt</th> <th>Fecha de devolución</th> </tr> </thead> <tbody> <tr> <td colspan="6">No se han creado copias.</td> </tr> </tbody> </table> <p>Información bibliográfica adicional:</p> <p>650 Subject added entry--topical term - Topical term or geographic name as entry element: BAHIA CULEBRA</p> <p>650 Subject added entry--topical term - Topical term or geographic name as entry element: PROYECTO TURISTICO</p> <p>650 Subject added entry--topical term - Topical term or geographic name as entry element: PAPAGAYO</p> <p>650 Subject added entry--topical term - Topical term or geographic name as entry element: INSTITUTO NACIONAL DE TURISMO</p>						Función	Código de barras #	Descripción	Estado	Estado Dt	Fecha de devolución	No se han creado copias.					
Función	Código de barras #	Descripción	Estado	Estado Dt	Fecha de devolución												
No se han creado copias.																	

Fuente: EspoBiblio 3.2

Es importante indicar que siempre será necesario realizar una revisión minuciosa de cada uno de los registros transferidos, a fin de limpiar cualquier otra inconsistencia que se haya producido durante el proceso de migración.

3. CONCLUSIÓN

Las unidades de información deben ir a la vanguardia en la tecnología de almacenamiento y recuperación de la información. El uso de formatos de intercambio de información ha permitido la transferencia de datos entre ellas y la migración a nuevas y más modernas plataformas. Asimismo, contar con un usuario mucho más alfabetizado tecnológicamente implica que se deben ofrecer servicios de alta calidad sobre sistemas de automatización cada vez más sofisticados.

La migración hacia dichos sistemas es imperante. Sin embargo, es importante tener en cuenta que siempre existirá pérdida de información, sobre todo, si es necesario el migrar no solamente de sistema, sino también de formato. Como se mencionó, el estándar que actualmente se utiliza de manera extendida por los nuevos sistemas de automatización es el MARC, por lo que es necesario establecer un plan de migración procurando transferir la mayor cantidad de campos y con la menor pérdida posible.

Por otra parte, el proceso de limpieza de la información posterior debe realizarse con sumo cuidado y esmero, a fin de garantizar hasta donde sea posible, la pureza de los datos iniciales. Es importante realizar un estudio previo, con base en la integridad de la base de datos fuente, que determine el posible nivel de inconsistencias que podrían surgir en el proceso. Debido a la cantidad de inconsistencias y pérdida de información, cabe la posibilidad de que dicho estudio muestre que el proceso de depuración sería más costoso que ingresar de forma manual todos los datos en el nuevo sistema.

Por lo tanto, este documento pretende ser una guía con los elementos básicos para realizar una migración de datos de manera exitosa. Sin embargo, todo el proceso de migración debe responder a un proyecto de automatización o de migración de datos previamente diseñado.

4. REFERENCIAS

- Bermúdez, Y. (2013). *Migración de datos de WinIsis a OpenBiblio*. San José, Costa Rica: EBCI.
- Biblioteca del Congreso de los Estados Unidos. (2003). *¿Qué es un registro MARC, y por qué es importante?* Recuperado de <http://www.loc.gov/marc/umbspa/um01a06.html>
- Chinchilla, R. (2011). El software libre: una opción para la automatización de unidades de información. *Bibliotecas*, 29(2). Recuperado de <http://eprints.rclis.org/19389/1/1557-3836-1-SM.pdf>
- Chinchilla, R. y Fernández, M. (2012). Bibliotecas automatizadas con software libre: establecimiento de niveles de automatización y clasificación de las aplicaciones. *Bibliotecas*, 30(2). Recuperado de <http://eprints.rclis.org/19390/1/4912-10390-1-SM.pdf>
- Fernández, M. (2013). Clasificación del software libre orientado a la automatización integral de bibliotecas según el nivel de complejidad de la biblioteca: bibliotecas simples, bibliotecas de mediana complejidad y bibliotecas de alta complejidad. *E-Ciencias de la Información*, 3(1). Recuperado de <http://revistas.ucr.ac.cr/index.php/eciencias/article/view/8491/0>
- Furrie, B. (2001). *Conociendo MARC bibliográfico: catalogación legible por máquina*. Washington: Biblioteca del Congreso de los Estados Unidos.
- Ugonono, M. (2011). Cronología del CDS/ISIS. *Boletín electrónico ABGRA*, 3(3).
- United Nations Educational, Scientific and Cultural Organization [Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, UNESCO]. (2003). *WinIsis: manual de referencia* [Versión 1.5]. Madrid: CINDOC-CSIC.
- United Nations Educational, Scientific and Cultural Organization [Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, UNESCO]. (2010). *CDS/ISIS database software: UNESCO and Information processing tools*. Recuperado de http://portal.unesco.org/ci/en/ev.php-URL_ID=5330&URL_DO=DO_TOPIC&URL_SECTION=201.html
- Vera, C. (2003). *Sistema de Información Bibliográfica de la CEPAL: manual de referencia*. Santiago de Chile: CEPAL.