

01053 4 24



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Facultad de Filosofía y Letras

Control de autoridad y catálogos de autoridad de materia con sistemas automatizados: aplicación a la *Biblioteca Daniel Cosío Villegas* de EL COLEGIO DE MÉXICO

T E S I S

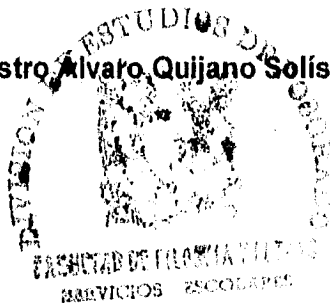
QUE COMO REQUISITO PARCIAL PRESENTA PARA OBTENER EL GRADO DE MAESTRA EN BIBLIOTECOLOGÍA

PILAR MARÍA MORENO JIMÉNEZ



Asesor: Maestro Alvaro Quijano Solís

México, D.F.



1996

TESIS CON FALLA DE ORIGEN

TESIS CON FALLA DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO

Dedicatoria	7
Agradecimientos	9
Advertencia	11
Introducción	13
1.—Control de autoridad y catálogos de autoridad	17
1.1.—Conceptos	17
1.2.—Control de autoridad manual <i>versus</i> automático	19
1.3.—Control de autoridad y calidad de las bases de datos	23
1.4.—Control de autoridad y recuperación de la información ...	24
1.5.—El control de autoridad en el contexto de la automatización de las bibliotecas	28
1.5.1.—Fuente de los registros de autoridad	29
1.5.2.—Almacenamiento de los registros de autoridad	30
1.5.3.—Mantenimiento de la base de datos de autoridad	36
1.5.4.—Relación entre los registros de autoridad y los registros bibliográficos	39
1.5.5.—Estructura de las referencias	40
1.5.6.—Productos y evaluación	42

2.—Antecedentes	43
2.1.—Proyecto de automatización y conversión retrospectiva del catálogo público de la BDCV	43
2.2.—Encabezamientos de materia de la BDCV	45
2.3.—El formato MARCOLMEX para registros bibliográficos ..	46
2.4.—El manejador STAR	46
2.4.1.—Recuperación de información	47
2.4.2.—Índices, pseudoíndices, superíndices	49
2.4.3.—Edición de registros	49
2.4.4.—Emisión de informes	50
2.4.5.—Archivos de autoridad	50
2.5.—La automatización integral con STAR	51
3.—Formulación de hipótesis	55
4.—El sistema de control de autoridad de la BDCV	57
4.1.—Diagnóstico de errores en el catálogo	57
4.2.—La catalogación asistida por STAR	67
4.3.—Aplicación de filtros semi-automáticos: el informe ANALIZA	71
4.4.—Creación del catálogo de autoridad de materia	74
4.4.1.—El formato MARCOLMEX para registros de autoridad ..	74
4.4.2.—Archivos de autoridad con STAR	75
4.4.3.—Creación de registros de autoridad de materia	85
4.5.—Aplicación de un algoritmo de corrección de errores tipográficos en encabezamientos de materia	86
4.6.—Mantenimiento del catálogo de autoridad de materia	93
4.7.—Intercambio de datos de autoridad de materia	97
4.8.—Evaluación de resultados	97
5.—Conclusiones y recomendaciones	101
Notas	105
Bibliografía	111
ANEXO: Formato MARCOLMEX para registros de autoridad	121

A todos
mis compañeros
de la Biblioteca Daniel Cosío Villegas
tantus labor non sit cassus

A mis queridos amigos:
Maribel, Santiago, Ana,
Eduardo, Toñi, Emilio,
Fuensanta, Pedro, Inmaculada,
Heraclio y Lourdes,

soporte precioso
de mis ausencias

A Francisca de Annobón y Francis Soutberg
in memoriam

AGRADECIMIENTOS

Agradezco muy especialmente a Álvaro Quijano Solís el haber aceptado dirigir esta tesis. Su interés, su exigencia, sus críticas y también su estímulo y apoyo institucional han sido muy importantes para llevarla a buen fin.

La colaboración de Francisco Mejía en la parte de programación de cómputo ha significado para mí una oportunidad insustituible para contrastar mis ideas con la realización práctica de las mismas. Le agradezco enormemente su interés por el proyecto, su paciencia, su eficiencia a la hora de buscar nuevas soluciones computacionales a caminos aparentemente sin salida y su buen talante y disposición en todo momento.

Heshmataallah Khorrarnzadeh me hizo importantes precisiones metodológicas, despejando algunas dudas de carácter estadístico, por lo que le quedo también muy reconocida.

Me siento igualmente agradecida a Shirley Ainsworth, quien propuso el procedimiento inicial de diagnóstico de errores en el catálogo y colaboró de forma entusiasta en la recopilación de datos para el mismo.

Brunilda Carretero me asistió en la revisión de la prueba final. Gracias a su criterio, basado en su amplia experiencia, fue posible despejar todas las dudas catalográficas que planteó esta última evaluación.

Le debo también un agradecimiento especial a la Mtra. Gloria Escamilla por todas sus observaciones, especialmente por el cuidado que puso al revisar la definición del formato propuesto.

Agradezco igualmente al Mtro. Ramiro Lafuente su lectura crítica de esta tesis,

FALTA PAGINA

No. 1

A la 9

a partir de la cual me planteé ampliar el primer capítulo, y a la Mtra. Rosalba Cruz por haber dedicado su valioso tiempo a una lectura atenta de la misma.

El haber realizado esta tesis a la vez que desarrollaba mis labores profesionales en El Colegio de México me ha reportado considerables beneficios. Si bien no siempre pude dedicarle todo el tiempo que hubiera requerido para alcanzar a concluirlo en un plazo académicamente razonable y aunque el ritmo de avance se vio marcado en buena medida por las prioridades y plazos de la propia biblioteca, la verdad es que este trabajo no se habría enriquecido tanto si no hubiera tenido la oportunidad de simultanear la reflexión sobre el control de autoridad en las bibliotecas con las necesidades concretas de la Biblioteca Daniel Cosío Villegas. Estoy muy reconocida a la Biblioteca Daniel Cosío Villegas, a mis colegas que fueron aportando críticas sobre los avances que se reflejaban en el trabajo diario y a El Colegio de México por haber colaborado poniendo todos los medios institucionales a mi alcance.

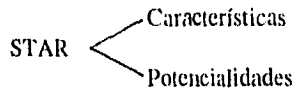
Le debo también un reconocimiento especial a la Dra. Elsa Barberena, quien me animó a superar los sucesivos trámites burocráticos que la presentación del examen de tesis exigía. Estos llegaron a buen fin gracias a la colaboración de eficientes y amables funcionarios de la UNAM como el Lic. Dfáz de la URI y Lic. Erick Camacho de la División de Estudios de Posgrado, a quienes quedo muy reconocida.

ADVERTENCIA

El manejador de bases de datos STAR posee determinadas características de funcionamiento dadas o "cerradas", es decir, sin posibilidad de modificación por parte del usuario/administrador del sistema, y funciones que permiten, mediante programación, desarrollar ciertas potencialidades del mismo.

El sistema de control de autoridad presentado como aportación en este trabajo de tesis se centra en el nivel de diseño de funciones y mecanismos especiales para dicho propósito, basado tanto en las características que ya posee el manejador STAR como en las potencialidades del mismo que deben ser desarrolladas.

Hay pues que deslindar este trabajo creativo, responsabilidad exclusiva de la autora, de las características ya dadas de STAR y del trabajo de programación de cómputo que debió desarrollarse para alcanzar los objetivos propuestos, el cual estuvo enteramente a cargo del Ing. Francisco Mejía, integrante del proyecto de automatización de la Biblioteca Daniel Cosío Villegas.



Funciones programables:

⋮

Control de autoridad

↳ Diseño

↳ Programación

INTRODUCCIÓN

Hoy día son pocas las bibliotecas mexicanas pertenecientes a instituciones de educación superior e investigación que no han iniciado ya el proceso de automatización de sus catálogos. La automatización de las bibliotecas es un paso importante para ofrecer mejores servicios a los usuarios, y la presencia de la computadora y su capacidad para multiplicar las opciones y los puntos de acceso al catálogo, en contraste con las posibilidades que daba el arreglo manual, representan una insustituible oportunidad para conciliar los intereses de usuarios y bibliotecarios. Esta conciliación deberá darse a partir del acercamiento entre los respectivos lenguajes. Es decir, el catálogo automatizado deberá dar al usuario la mayor cantidad posible de asociaciones que permitan a éste la posibilidad de encontrar siempre lo que busca, sin importar si el lenguaje de sus preguntas está o no ortodoxamente formalizado en términos bibliotecarios.

El control de autoridad debe constituir un paso más en la automatización de los catálogos, pues con éste aumentará la capacidad de recuperación eficiente de los sistemas automatizados. Las bibliotecas no deberían conformarse únicamente con automatizar su catálogo sino que deberían también desarrollar sistemas que permitan un mayor aprovechamiento del mismo.

Uno de los retos más importantes de la tecnología para bibliotecas consiste en presentar al usuario la información en la forma más sencilla para él. Pero no debe olvidarse que esto supone un mayor esfuerzo en el diseño y programación de sistemas automatizados de información, que requieren para responder a esta demanda de una mayor complejidad interna. El objetivo final consiste en "emular" la

FALTA PAGINA

No. 13

mente humana: cómo busca información cada usuario, qué le resulta relevante de la información presentada, cómo se comporta para seguir buscando información, etc. son preguntas a las que la inteligencia artificial y los llamados sistemas expertos nos responderán en un futuro quizás no tan lejano.

Al plantearme el tema de esta tesis observé de un lado la importancia que éste revestía para las bibliotecas y de otro la ausencia de literatura actualizada en español que orientara a las mismas por ese camino. Es por ello que un objetivo prioritario para mí fue la recopilación y sistematización del abundante material escrito sobre control de autoridad y catálogos de autoridad automatizados en idioma inglés para presentarlo en forma ordenada y clara a los bibliotecarios de nuestro medio cultural. Es, sin embargo, imprescindible adaptar estos conocimientos y experiencias a nuestro propio ámbito. Con este trabajo he pretendido compilar los conocimientos generalizables imprescindibles que nos pueden servir para buscar soluciones propias a los mismos problemas. A la vez, he realizado un ejercicio de aplicación que me ha servido para darme cuenta de que, si bien es preciso ser muy creativos a la hora de buscar soluciones en nuestras situaciones concretas, las aportaciones teóricas nos permitirán centrarnos adecuadamente en los principios y fines a los que debemos tender, así como las experiencias de otros nos pueden presentar diferentes caminos viables.

A lo largo de este trabajo de tesis abordaré la problemática del control de autoridad y los catálogos de autoridad de materia en el contexto de la automatización de las bibliotecas. Asimismo expondré un plan para, por un lado, depurar y prevenir la entrada de errores de diversos tipos en los encabezamientos de materia del catálogo en línea de la Biblioteca "Daniel Cosío Villegas" de El Colegio de México (BDCV) y, por otro, crear y mantener un catálogo de autoridad de materia automatizado para esta biblioteca. También referiré la puesta en marcha de dicho plan y una evaluación de los resultados obtenidos.

En el primer capítulo realizaré una exposición de las bases teóricas en que se fundamenta el control de autoridad y los elementos que se deben tener en cuenta al llevarlo a cabo con sistemas automatizados según se desprende de la literatura relacionada con el tema producida principalmente en Estados Unidos y en otros países. Los aspectos teóricos se han tratado al nivel básico suficiente para la comprensión de los desarrollos prácticos que se exponen a continuación y no con la profundidad de quien quisiera ahondar mucho más en los aspectos lingüísticos, semánticos o epistemológicos del manejo de encabezamientos de materia. Parece obvio que una tesis sobre control de autoridad supone una toina de postura clara a favor de los lenguajes controlados. Sin abordar específicamente esta discusión, se da cuenta de las razones a favor del control principalmente en los epígrafes 1.3 y

1.4, sobre control de autoridad y calidad de las bases de datos, y control de autoridad y recuperación de la información, respectivamente. Habida cuenta de que lo que manejamos en nuestra biblioteca y sobre lo que nos estamos basando y refiriendo continuamente son las listas de encabezamientos de materia, no introduje tampoco una discusión sobre pre o postcoordinación. Trabajé con lo dado y no entendí la problemática al manejo de descriptores; aunque hay planteamientos que pudieran resultar válidos para ambos sistemas, esto queda al criterio del lector.

En el segundo capítulo explicaré cuáles son los antecedentes del Proyecto de Control de Autoridad de la BDCV. Como veremos, éstos son, administrativamente, el proyecto de automatización y conversión retrospectiva del catálogo público; y, técnicamente, el sistema de encabezamientos de materia de la BDCV, el formato para registros bibliográficos MARCOLMEX y el manejador de bases de datos STAR. Cabe señalar que si bien el manejador STAR posee una serie de características y potencialidades útiles para este propósito la mera instalación del mismo no supone ni mucho menos el contar de forma inmediata con un sistema de control de autoridad automatizado. El sistema de control de autoridad automatizado de la BDCV fue diseñado e implementado a partir del proyecto que constituye una de las partes centrales de esta tesis, tratando de armonizar las cualidades del manejador con los fines perseguidos y optimizando a la vez su uso. Fue un proceso largo y complejo que esperamos pueda ser de utilidad a otros usuarios tanto del sistema STAR como de otros similares. Las bibliotecas que no cuenten con un sistema automatizado complejo pueden beneficiarse también del contenido de esta tesis, que procura además ilustrar sobre aspectos de aplicación general no ligados a ningún sistema en particular.

En el tercer capítulo formularé las hipótesis, tanto de base como a demostrar. Básicamente partiré del supuesto de que con el manejador de bases de datos STAR es posible la creación y mantenimiento de catálogos de autoridad con las principales características señaladas en el capítulo primero. Además sostendré que después de la aplicación de las medidas que se especifican en el capítulo cuarto, la calidad de los registros catalogados en línea con este sistema habrá aumentado.

El método que apliqué aparece detallado en el capítulo cuatro. Allí explicaré cómo se elaboró el diagnóstico de errores del catálogo público en línea de la BDCV y una tipificación de éstos. Asimismo expondré una metodología tanto para la evaluación de algunos aspectos cualitativos de las bases de datos, como para la evaluación personal de los catalogadores. Esta metodología permite medir el nivel de calidad de los registros catalogados por la ausencia de errores en los mismos. Hay, desde luego, otros factores que bien podrían tomarse en cuenta para la evaluación personal de los catalogadores, como son la adecuación de los

temas y clasificaciones asignados, derivados de las capacidades de análisis y síntesis, dominio de los instrumentos catalográficos, "empatía" con el usuario, capacidad de aplicar criterios generales adecuados más allá de las reglas o donde no llegan éstas, etc. Estas últimas cualidades o características del buen catalogador no son tan fácilmente observables ni medibles y no están incluidas en la metodología de evaluación de catalogadores presentada en esta tesis.

En el mismo capítulo, referiré la asistencia que, gracias a ciertas características programables del manejador STAR, puede brindarse en el módulo de catalogación para evitar errores de codificación. Además presentaré el informe ANALIZA, que se programó con STAR con el fin de servir como filtro semiautomático para la detección de errores, fundamentalmente de etiquetado MARC, en los registros catalogados antes del paso definitivo de éstos al inódulo de consulta del catálogo público.

En el apartado cuarto del capítulo cuarto explicaré el procedimiento por el cual creamos el catálogo de autoridad de materia a partir de los envíos en tarjetas que existían en el catálogo manual público. En el siguiente apartado describiré cómo definimos y aplicamos un algoritmo de corrección de errores tipográficos en los encabezamientos de materia. En el apartado sexto expondré los pasos que, a partir de la creación del catálogo de autoridad de materia, debe seguir el catalogador para usarlo y actualizarlo. El apartado séptimo lo dedicaré a exponer las bases sobre las que será posible el intercambio de datos de autoridad de materia con otras instituciones que también posean catálogos disponibles en forma magnética y en formato MARC. Los problemas que nos puedan plantear a los bibliotecarios mexicanos los proyectos cooperativos en este rubro es materia que sería interesante abordar en otra investigación.

Al final del cuarto capítulo se repetirá la prueba que sirvió para elaborar el diagnóstico de errores, comprobando cómo éstos disminuyeron después de haberse aplicado las medidas señaladas.

En el capítulo quinto aparecerán las conclusiones que se pueden derivar de este trabajo, incluyendo algunas reflexiones sobre la utilidad y aplicación de la metodología y mecanismos aquí expuestos no sólo a la Biblioteca Daniel Cosío Villegas, sino a otras semejantes en México.

La bibliografía que utilicé en la elaboración de esta tesis aparece a continuación del capítulo cinco.

Por último, proporciono en un anexo el "Formato MARCOLMEX para registros de autoridad" que diseñé para su uso en la BDCV y que podrá servir a instituciones similares como pauta para el almacenamiento e intercambio de registros de autoridad.

1.—Control de autoridad y catálogos de autoridad

1.1.—Conceptos

El control de autoridad se define como "el proceso por el cual se mantiene la consistencia en las entradas, o puntos de acceso, de un catálogo bibliográfico, tomando como referencia un catálogo de autoridad."¹

Un catálogo de autoridad debe contener todas las entradas posibles para los encabezamientos, ya sea de nombres, títulos uniformes, series o materias, así como las indicaciones que los relacionen entre sí.

Los propósitos que se persiguen con las actividades relacionadas con el control de autoridad son:

- Creación de encabezamientos (puntos de acceso) uniformes que aseguren la colocación de todas las obras relacionadas entre sí bajo los mismos encabezamientos.
- Creación de relaciones (generalmente bajo la forma de referencias cruzadas) que permitan llegar de las variantes de un encabezamiento a la forma establecida para el mismo.
- Provisión de la información adecuada para distinguir un encabezamiento autorizado de otro similar.²

En la literatura norteamericana reciente sobre aspectos teóricos del control de autoridad³ se distingue entre dos conceptos muy relacionados pero diferentes. Se trata del control de autoridad, por una parte, y por otra del llamado "authority work"⁴.

Tillet afirma que el concepto de "authority work" incluye el trabajo de investigación y el esfuerzo intelectual involucrados en la creación y actualización de los registros de autoridad. Las tareas específicas como son: determinar la forma correcta de un encabezamiento según determinadas normas generalmente aceptadas, consignar las variantes no aceptadas del encabezamiento elegido, establecer las relaciones con encabezamientos semejantes, especificar el uso y alcance del encabezamiento, etc. se consideran como parte de este esfuerzo. También están incluidos

el mantenimiento de los archivos de autoridad y el mantenimiento correspondiente de los registros bibliográficos, así como la elaboración de las referencias asociadas a aquéllos.

Para esta autora un catálogo de autoridad es un conjunto de registros de autoridad en los que los encabezamientos generalmente son consistentes y únicos. El control de autoridad es el término global que agrupa las operaciones descritas anteriormente, haciendo énfasis en el control sobre las formas diferentes en los puntos de acceso. Es decir, es el control sobre los puntos de acceso.

Se observa en la literatura bibliotecológica, tanto en español como en inglés, cierta inconsistencia en el uso de los vocablos "lista de autoridad/es", "archivo de autoridad/es" y "catálogo de autoridad/es", los cuales generalmente se usan indistintamente. En el "Diccionario técnico de Bibliotecología y Ciencias de la Información" de Ayala, Ayala y Lau se traduce el término inglés "authority file" por los españoles "catálogo de autoridades" y "archivo de autoridades", tomados como sinónimos, y se define este catálogo como "an official catalog of established entries"⁵. El mismo término "authority file", es traducido en el "Diccionario" de Martínez de Sousa como "lista de encabezamientos", a la cual se define, un tanto redundantemente, como "lista de encabezamientos seleccionados para su empleo en un catálogo"⁶. Esta es casi una traducción literal de la definición que de "authority file" proporcionan Wersig y Neveling en su "Terminología de la documentación", y que dice así: "a list of headings selected for use in a catalogue and sources consulted to establish such headings"⁷. Este término es traducido al español en esta misma fuente como "lista oficial de encabezamientos", la cual debe incluir las fuentes consultadas para establecer los encabezamientos. Lo mismo considera la "Encyclopedia of Library and Information Science", quien define "authority file" como "a separate listing of entries and the authority on which they are based"⁸. Los autores de la "Terminología" no incluyen en su obra el término "authority catalog", y, por el subrayado en la definición de "authority file", pareciera que asumen que el término "catálogo" designa únicamente al catálogo bibliográfico. Sí incluyen, en cambio, el término "authority list", traducido como "lista autorizada de términos", y definido como "list of terms that is used as an authority in indexing"⁹.

El catálogo tradicionalmente ha cumplido dos funciones importantes: describir y clasificar. Así, el catálogo bibliográfico describe físicamente las unidades bibliográficas y las clasifica dentro de categorías temáticas definidas por su contenido intelectual. Asimismo, el catálogo de autoridad describe los conceptos expresados por diferentes términos y clasifica esos términos dentro de varias categorías, como son: nombre, materia, conferencia, subdivisión, etc.

Podemos establecer un paralelismo entre las listas de encabezamientos

impresas y los archivos de autoridad automatizados. Una lista, al igual que un archivo, consiste en un conjunto de datos interrelacionados ordenados en forma secuencial y a los que se puede acceder únicamente por el orden establecido, ya sea alfabético o numérico. Una lista de encabezamientos permite el acceso a una serie de términos ordenados en forma alfabética. Un archivo de autoridad reúne un conjunto de registros de autoridad a los cuales la computadora les ha asignado un número consecutivo.

El catálogo de autoridad, a diferencia de la lista, permite llegar a un concepto, expresado por uno o varios términos, en una forma no única. En las bases de datos es aún más notorio este arreglo multidimensional. Así, es posible recuperar un registro de acuerdo a cada uno de los datos o unidades de información en las cuales se ha dividido.

En esta tesis denominaremos "lista de encabezamientos" a las listas de encabezamientos impresas publicadas para su uso en las bibliotecas como fuente autorizada. Así, la Lista de Encabezamientos de Materia de Gloria Escamilla o la Lista de Encabezamientos de Materia de la Biblioteca del Congreso de los Estados Unidos. Nombraremos como "archivo de autoridad" a un conjunto de registros de autoridad y como "base de datos de autoridad" al conjunto de archivos de autoridad que constituyen el catálogo de autoridad, almacenados en un medio magnético, y que permite tener varios puntos de acceso a los registros.

1.2.—Control de autoridad manual *versus* automático

Las bibliotecas que aún no poseen catálogos de autoridad automatizados generalmente emplean una lista de encabezamientos impresa aceptada en la que se van anotando las modificaciones y adiciones pertinentes. Otras bibliotecas usan un catálogo de tarjetas con los encabezamientos autorizados, agregando o reemplazando las tarjetas cuando lo consideran necesario.

Para algunas bibliotecas el catálogo público por materias representa su catálogo de autoridad y se limitan a completarlo con los envíos apropiados de "véase" cada vez que se añade un tema. Para añadir un tema nuevo se consulta una lista aceptada de encabezamientos de materia¹⁰.

Estos sistemas presentan algunas desventajas, como son:

- Al reemplazar una lista de encabezamientos de materia por una edición posterior se deben anotar de nuevo todas las modificaciones en la nueva edición.
- Los catálogos de tarjetas requieren mucho trabajo mecanográfico y deben

ser consultados cada vez que se usa un encabezamiento no incluido en la lista que se toma como base.

- Las modificaciones que han de hacerse en el catálogo implican varias tareas: recuperar los juegos de tarjetas completos, corregirlos y volver a intercalarlos. Cuando el cambio afecta a un número considerable de juegos de tarjetas algunas bibliotecas optan por intercalar envíos de "véase también", dando como válidos dos o más formas equivalentes de un mismo encabezamiento.

Recientemente, se ha considerado que "el control de autoridad automatizado consiste en el uso de una computadora para manejar grandes bloques en el proceso de control de autoridad."¹¹

En 1985 Taylor, Maxwell y Frost¹² describieron y analizaron los métodos automatizados de control de autoridad que estaban disponibles entonces, categorizándolos como sigue:

Procesamiento electrónico de catálogos de autoridad. Consiste en introducir en una computadora un catálogo de autoridad, ya sea con la ayuda de un procesador de palabras o con algún programa de administración de bases de datos.

Una vez introducida la información en la máquina el catalogador puede consultar allí los encabezamientos autorizados. Es decir, la verificación de los encabezamientos la realiza ahora el catalogador en una "lista electrónica", que presenta mayor versatilidad que las listas impresas o catálogos de tarjetas, y que a su vez puede manejarse en forma impresa si se desea.

Como limitaciones de este método señalaremos que no puede usarse interactivamente a la vez que se cataloga en línea y que las modificaciones que se hacen en el catálogo de autoridad deben volcarse posteriormente en los correspondientes registros bibliográficos y viceversa.

Catálogos de autoridad en línea sin conexión con los catálogos bibliográficos. El catálogo de autoridad se encuentra disponible en línea, pero sin conexión con los registros bibliográficos. De modo que puede consultarse rápidamente, pero los cambios que se efectúen en uno de los archivos no se reflejarán automáticamente en el otro.

Sistemas de control de autoridad integrados a un sistema de bases de datos bibliográficos. Este método permite establecer una conexión automática entre el catálogo de autoridad y los registros bibliográficos. Consiste en la creación de un catálogo de autoridad a partir de los encabezamientos tomados del catálogo bibliográfico y en el establecimiento de una relación entre ambos archivos que permita

que cualquier cambio en uno de ellos inicie automáticamente un cambio en el otro. En general, este método presenta las siguientes ventajas¹³:

- Validación automática de los encabezamientos. El sistema puede verificar si un encabezamiento asignado a un registro bibliográfico está autorizado en el catálogo de autoridad.
- Detección de errores. Si un encabezamiento asignado a un registro bibliográfico no aparece admitido en el catálogo de autoridad el sistema puede rechazarlo y alertar al catalogador.
- Despliegue de términos relacionados. El sistema puede desplegar las referencias asociadas a determinado encabezamiento en el momento de su asignación a un registro bibliográfico.
- Reducción del espacio requerido para el almacenamiento de datos en el sistema. Esta reducción de espacio se da en los sistemas que almacenan el encabezamiento únicamente en el registro de autoridad. Este, a su vez, incluye también todos los números de los registros bibliográficos a los que se asignó dicho encabezamiento. Por tanto, los encabezamientos, aunque sí aparecen en el despliegue de los registros correspondientes, no se encuentran de hecho en tales registros.
- Capacidad para efectuar cambios globales. Consiste en la capacidad del sistema para realizar cambios masivos en muchos registros a la vez.

Dalehite¹⁴, por su parte, identificó en 1989 cinco métodos de control de autoridad: manual, proceso en "batch" externo, proceso en "batch" interno, encabezamientos ligados en forma dinámica y encabezamientos no redundantes. El método manual puede aplicarse tanto en catálogos impresos como en automatizados, pero supone menos esfuerzo en éstos últimos. Los otros cuatro métodos requieren de una base de datos legible por computadora.

El método manual consiste en el cambio físico uno por uno tanto de los registros de autoridad como de los registros bibliográficos asociados a ellos. Cuando la base de datos es legible por máquina, determinadas herramientas, como la indización de subdivisiones, búsquedas booleanas o la emisión de informes impresos para identificar los encabezamientos nuevos o eliminados, pueden facilitar este proceso.

El proceso en "batch" externo supone el cotejo en una sola vez de una serie de registros bibliográficos con un archivo de autoridad legible por computadora, para así actualizar los encabezamientos.

El proceso en "batch" interno utiliza un procedimiento de cambio global que permite al usuario identificar las formas anteriores y nuevas de los encabezamientos e iniciar automáticamente el cambio del encabezamiento anterior por el nuevo. El sistema busca la forma no aceptada del encabezamiento y si la encuentra cambia el texto por la nueva forma. Algunos sistemas buscan en los registros uno por uno, otros recorren los índices para seleccionar los encabezamientos sujetos a cambios.

El método de encabezamientos ligados en forma dinámica requiere que los registros de autoridad y los bibliográficos se encuentren ligados interactivamente.

Esta relación permite una identificación rápida de los registros bibliográficos afectados por un cambio, de tal manera que un cambio en un registro de autoridad producirá cambios idénticos en todos los registros bibliográficos asociados a él. Este proceso puede iniciarse inmediatamente o producirse durante las noches corriendo un programa en "batch".

El método de encabezamientos no redundantes, por último, elimina la necesidad de cambiar los registros bibliográficos, ya que el texto de los encabezamientos se encuentra tan sólo en los registros de autoridad. El acceso a los registros bibliográficos se efectúa siempre a través de los registros de autoridad o, dependiendo de la complejidad del sistema, a través de los índices de los encabezamientos.

Como puede observarse, los dos últimos métodos consignados se corresponden con la tercera categoría señalada por los anteriores autores. El establecimiento de una relación entre los registros de autoridad y los registros bibliográficos ha llegado a ser considerado por algunos autores como parte de la definición misma del control de autoridad automatizado¹⁵.

El papel de las computadoras en el control de autoridad incluye principalmente:

- asistencia en el establecimiento de los encabezamientos autorizados, que incluye la verificación y validación automática de los encabezamientos que se introduzcan en el catálogo,
- uso de algoritmos de normalización que minimicen las diferencias en cuanto a espacios, signos de puntuación, mayúsculas y minúsculas, diacríticos, etc. existentes en encabezamientos iguales,
- recuperación de encabezamientos mediante truncamientos, permutaciones, búsquedas por proximidad, por palabra, operadores booleanos, etc. y
- capacidad para efectuar cambios globales que aseguren la consistencia de los encabezamientos.

En el apartado 1.5 discutiremos a fondo estas y otras características de los diferentes sistemas automatizados de control de autoridad, con sus ventajas y desventajas.

1.3.—Control de autoridad y calidad de las bases de datos

La calidad puede definirse como un nivel de desarrollo que tiende a la excelencia de un producto o servicio. Cuando se habla del concepto de calidad total se quiere significar que la calidad del todo es igual a la suma de la calidad de cada una de las partes. La calidad de las bases de datos vendría determinada, pues, por la suma de calidades en cada uno de los pasos del proceso de recopilación, manufactura y difusión de la información contenida en ellas. Más que un control sobre el producto acabado, como ocurre en la industria, se busca aquí controlar la calidad del proceso.

La mayor demanda en la calidad de las bases de datos, según O'Neill y Vizine-Goetz¹⁶, viene dada por dos factores: el creciente uso de grandes bases de datos, ya sea en línea o en discos compactos, y el nuevo énfasis de la sociedad en general sobre la calidad.

No se puede hablar de patrones absolutos de medición de la calidad de las bases de datos, ya que éstos deben definirse en cada caso particular atendiendo a los elementos en juego, como son las necesidades de los usuarios, las normas nacionales e internacionales, y las políticas particulares de las instituciones, entre otros.

La calidad de las bases de datos se mide, entre otras cosas, por la ausencia de errores de diversos tipos. Muchas de las investigaciones sobre la calidad de las bases de datos bibliográficos están encaminadas a encontrar procedimientos, en lo posible automáticos, para la detección y corrección de dichos errores. Sin embargo, como señala Aversa¹⁷, existen diversas técnicas de control de calidad en la producción de bases de datos que van desde las más altas tecnologías, como el uso de complicados sistemas expertos, hasta rutinas de "baja tecnología", siendo lo más recomendable la combinación de soluciones en diferentes niveles. No siempre disponemos de medios complicados o no siempre se necesitan medios complicados para alcanzar nuestros objetivos.

Graham distingue dos aspectos en la calidad de los registros bibliográficos: extensión y precisión. De acuerdo a ello considera varios elementos, como son: nivel de detalle, precisión mecánica (transcripción tipográfica), precisión intelectual, corrección formal de los encabezamientos, existencias (holdings) y adherencia a normas¹⁸.

Para Clack¹⁹, lo que cada organización realice internamente para aumentar el

control de calidad sobre sus bases de datos contribuirá significativamente a mejorar el sistema de control de autoridad. En efecto, gran parte del éxito en la implementación de un sistema de control de autoridad depende de la calidad de las bases de datos. Y, por oposición, nada puede mermar más la credibilidad de un sistema que una base de datos deficientemente preparada, ya que los datos capturados con escasa calidad afectan negativamente la productividad del sistema, disminuyendo la precisión, integridad y/o consistencia de la base. Desde luego, el éxito en la recuperación de los documentos depende también de las técnicas de recuperación, o sea, de la calidad del sistema mismo.

Además, los errores pueden resultar muy caros. Pensemos, por ejemplo, cómo un mayor tiempo de uso de algunas bases de datos bibliográficos en línea aumenta el costo de las búsquedas. De igual forma, un documento no recuperado puede generar una solicitud de compra innecesaria.

Otro factor importante es el efecto multiplicador de los errores, ya que éstos pueden pasar de unas bases a otras como "virus". Datos aparentemente correctos, una vez perdida la conexión con el documento original, pueden llegar a perpetuarse, pues no será fácil advertir la existencia de errores.

La investigación de procesos automáticos y semi-automáticos encaminados a prevenir y corregir diferentes errores en la información de las bases bibliográficas forma parte importante del proceso mismo de control de autoridad.

1.4.—Control de autoridad y recuperación de la información

Los catálogos automatizados, ya sea en línea o en discos compactos, presentan un mayor número de puntos de acceso que los catálogos tradicionales de tarjetas. Además, los sistemas actuales proveen ayuda al usuario en la búsqueda de información.

Varias encuestas llevadas a cabo en los Estados Unidos muestran que al menos la mitad de las búsquedas en los catálogos en línea son temáticas²⁰. Como auxiliares en la recuperación temática se han venido usando varios procedimientos, a saber²¹:

- Indizado de encabezamientos de materia, ya sea en forma exacta o literal, o por palabras, lo cual permite mayores opciones de búsqueda, ya que no siempre han de utilizarse las frases exactas.
- Uso de operadores booleanos "y" "o" "no" para ampliar o restringir las búsquedas.

- Uso de truncamientos limitados o ilimitados, lo cual generalmente amplía la búsqueda y recupera por términos que pueden variar entre sí en su ortografía.
- Inclusión de términos tomados del título y/o tablas de contenido. Este procedimiento resulta muy útil en las ciencias llamadas "duras", especialmente si se usa en combinación con los anteriores y con catálogos de autoridad que incluyan referencias, pero usado en sustitución de aquéllos puede dar lugar a dispersión e inconsistencia en los términos.
- Búsqueda por números de clasificación. La búsqueda por números de clasificación decimal Dewey ha demostrado ser tan útil como la búsqueda por palabras clave. En cambio, el uso de la clasificación de la Biblioteca del Congreso de los Estados Unidos, debido a su extensión y complejidad, parece menos viable.
- Derivación de palabras. Con este procedimiento el sistema recupera las palabras relacionadas con el término que el usuario introdujo en su búsqueda.
- Uso de catálogos de autoridad que incluyan referencias cruzadas.

Aunque, en principio, pudiera parecer que la potencia en la capacidad de búsqueda que poseen algunos de los actuales sistemas para bibliotecas hace innecesario el esfuerzo de desarrollar estrategias para el control de autoridad, lo cierto es que, como algunos estudios han demostrado, sin una estructura adecuada de referencias cruzadas la recuperación de la información en las bases de datos no es tan eficiente.

En un estudio llevado a cabo por investigadores de la Universidad de Western Ontario en Canadá se intentó averiguar si la capacidad de un sistema para realizar búsquedas booleanas y por palabras clave podría substituir a un sistema que incluyera una estructura de referencias cruzadas. Para ello, se realizaron búsquedas utilizando palabras incluidas en los términos no admitidos de la Lista de Encabezamientos de Materia y del Catálogo de Autoridad de Nombres de la Biblioteca del Congreso de los Estados Unidos, encontrándose que el 53.7% de entradas de autor, el 73.5% de entradas temáticas de autor y el 69.6% de entradas de materia no pudieron ser recuperadas. La conclusión de este estudio fue que "incluso en un catálogo en línea con capacidad para realizar búsquedas por palabras clave, búsquedas booleanas y truncamientos, es necesaria una estructura de referencias"²².

Otro estudio reciente, llevado a cabo por Wilkes y Nelson²³, refuerza estos planteamientos. En este caso, se compararon los resultados de la búsqueda por materias en dos catálogos en línea, uno con un sistema de control de autoridad y el otro no. Se analizaron las bitácoras de transacciones (transaction logs) de los usuarios en el catálogo de la biblioteca sin sistema de control de autoridad,

encontrándose que un 40.6% de las búsquedas por materia fueron infructuosas, es decir, dieron como resultado 0. Posteriormente se realizaron esas mismas búsquedas en el catálogo de la biblioteca con sistema de control de autoridad. Los términos que fueron identificados en este sistema con una referencia de "véase" se introdujeron como nuevas búsquedas en la biblioteca sin control de autoridad. El resultado fue que el 73.9% de las búsquedas que en un principio habían resultado infructuosas arrojaron ahora algún resultado.

Un estudio de Bangalore en el que analiza las búsquedas con resultado cero en un catálogo en línea concluye que "la presencia de registros de autoridad relacionados con nuestro catálogo público en línea ayudó a los usuarios a obtener el encabezamiento adecuado"²⁴.

Lancaster señala que "aunque el vocabulario que se use en un sistema de recuperación no tiene efectos en la cobertura del sistema, el esfuerzo que el usuario debe realizar para explotarlo aumentará y, consecuentemente, también aumentará el tiempo de respuesta del mismo"²⁵. Algunos usuarios prueban a introducir términos diferentes cuando no obtienen resultados en su búsqueda, pero la mayor parte abandona las terminales sin hacer otro intento. Incluso los usuarios que prueban con otros términos no tienen una garantía de dar con el término admitido. El tiempo que los usuarios deberán pasar frente a las terminales del catálogo automatizado aumentará si cuando obtienen resultados negativos tienen que probar con otros términos de búsqueda. Es importante que las bases de datos contengan mucha información muy relevante pero también que éstas puedan ser explotadas con el menor esfuerzo y el menor tiempo por parte del usuario.

Como hace notar Macafferri "de hecho, se puede decir que el control de autoridad es más importante en un contexto automatizado que en un medio manual, dada la literalidad con la que las computadoras almacenan y recuperan la información"²⁶. Para algunas computadoras la cadena de caracteres: "Educación de adultos" es diferente a la cadena: "Educación de adultos" (en donde entre la palabra "Educación" y la palabra "de" hay dos espacios). Los errores de espacios, puntuación y ortográficos, por lo general, alejan considerablemente los encabezamientos o términos en los índices y, en el caso de los ortográficos, imposibilitan también su recuperación por palabra. Si se busca por la palabra "Cultura" y en un registro esta palabra aparece mal escrita como "Cutura", éste no aparecerá en el resultado de la búsqueda. Hay que destacar, sin embargo, que algunos sistemas almacenan los encabezamientos en los índices con mayúsculas y sin acentos ni diacríticos, con lo cual, aunque existan diferencias de este tipo entre encabezamientos, al quedar indizados juntos, éstas no afectan la recuperación. Por ejemplo, al buscar "BIOLOGIA" el sistema podría recuperar registros en los que aparecen como temas: "Biología", "biología" o "btología".

Johnston asegura que "a la computadora le falta la capacidad intelectual para reconocer asociaciones entre términos que no estén explícitamente ligados"²⁷.

Dalelite señala que el trabajo de autoridad ("authority work") no puede automatizarse, ya que es un proceso intelectual²⁸. Este proceso intelectual es, sin duda, el principal sustento del control de autoridad y la base sobre la que se asentará el funcionamiento automatizado de los catálogos de autoridad.

Por otro lado, en el medio automatizado la importancia de los conceptos de "asiento principal"²⁹ y "arreglo alfabético" ha sido puesta en cuestión. En cuanto al arreglo alfabético, tenemos que decir que sigue desempeñando una función, ya que es frecuente el uso de búsquedas mediante la revisión en los índices, posibilidad que la mayoría de los sistemas ofrece. En el medio automatizado, al igual que en el manual, es importante definir cuáles son las reglas de ordenamiento de palabras y términos, qué palabras o símbolos no se tomarán en cuenta, etc., para facilitar la revisión secuencial. La búsqueda en los índices presenta algunas ventajas, como son: se puede llegar a los términos por aproximación, sin necesidad de conocerlos en forma exacta previamente, y se pueden encontrar términos relacionados próximos en el índice, aunque se pierden los términos relacionados que no están próximos. Por ejemplo, si el usuario no está seguro de si el término adecuado es "Biblioteconomía" o "Bibliotecología" y encuentra en el índice la palabra "Bibliotecología", ahí mismo puede encontrar también "Bibliotecarias", "Bibliotecarios", "Bibliotecas", "Bibliotecas infantiles", "Bibliotecas y estado", pero le quedará más lejos "Bibliología" y aún más lejos "Educación bibliotecológica". Esta forma de búsqueda en índices requiere más tiempo y se recupera más información pero menos relevante. Frente a ésta, la búsqueda por selección, en la que el usuario recupera un conjunto de obras con una o varias características comunes, como las búsquedas en las que se combinan varios términos con operaciones booleanas, toma menos tiempo y suele ser más precisa, recupera menos información pero más relevante y puede recuperar palabras relacionadas aunque dispersas o lejanas en los índices. Como contrapartida, el usuario debe conocer ciertas técnicas y poner especial cuidado al escribir los términos, de los cuales generalmente debe tener un conocimiento previo.

La literatura sobre las ventajas que supone que los sistemas automatizados para bibliotecas posean mecanismos de control de autoridad es abundante, especialmente en los últimos diez años en los Estados Unidos. Prácticamente todos los artículos publicados sobre el tema ofrecen alguna discusión de este tipo. Resumiendo, diremos con McDonald que "a medida que los catálogos en línea proliferan la importancia y relevancia del control de autoridad aumenta en tres formas: primero, es una herramienta que puede usarse para mantener la consistencia e

integridad del catálogo en línea; segundo, establece una estructuración conocida en el catálogo que facilita la recuperación de la información; y tercero, la consistencia, integridad y estructuración permiten que el sistema posea características que de otra forma no tendría."³⁰

1.5.—El control de autoridad en el contexto de la automatización de las bibliotecas

En 1988 Sarah Hager Johnston³¹ llevó a cabo una encuesta entre proveedores de servicios automatizados para bibliotecas en los Estados Unidos. El propósito de esta encuesta fue identificar las tendencias de los sistemas automatizados de control de autoridad y lo que dichos proveedores podían ofrecer. Con los resultados obtenidos la autora sugiere que los bibliotecarios norteamericanos podrían determinar hasta qué punto los proveedores satisfacen sus necesidades y así evaluar y seleccionar algún sistema para su uso local.

Algunos de los sistemas analizados en su artículo ya están disponibles en México³². Sin embargo, lo más interesante es el análisis mismo de esos sistemas, que puede ser también de utilidad en el caso mexicano, ya que permite determinar los elementos que deben tenerse en cuenta al considerar la puesta en marcha de cualquier sistema automatizado de control de autoridad.

Como sugiere Irwin, los elementos señalados por Johnston pueden servir tanto para la selección de un sistema, como para su mayor comprensión si éste se encuentra ya instalado³³.

Los elementos considerados por Johnston fueron: fuente de los registros de autoridad, almacenamiento de los registros de autoridad, mantenimiento de la base de datos de autoridad, relación y consistencia entre los registros de autoridad y los registros bibliográficos, estructura de las referencias, producción de listados y producción de estadísticas.

Partiendo como base de esos elementos, y con apoyo en la literatura reciente sobre control de autoridad, a continuación los analizaremos detalladamente, poniendo especial énfasis en el control de autoridad de materias. Los dos últimos puntos: producción de listados y producción de estadísticas aparecerán aquí bajo el apartado "Productos y evaluación".

1.5.1.—Fuente de los registros de autoridad

La fuente de los registros de autoridad pueden ser las listas de encabezamientos de materia (Escamilla, Rovira, Kidder, etc.), tesauros o catálogos manuales previos a la automatización, capturados en un sistema automatizado.

Otra posibilidad consiste en transferir electrónicamente registros provenientes de otros sistemas, a través de cintas magnéticas, discos flexibles, discos compactos, o por un proceso en línea. La Biblioteca del Congreso de los Estados Unidos, por ejemplo, ha comercializado sus catálogos de autoridad en cintas magnéticas y disco compacto. Actualmente existen escasas opciones en español, ya que pocas son las bibliotecas que cuentan con un catálogo de autoridad de materias en español disponible en forma magnética.

En México únicamente la Biblioteca Central de la UNAM posee un catálogo de autoridad de materia automatizado, que incluye más de 139 000 registros administrados por el programa DBaseIV. No obstante, empieza a haber intentos cada vez más serios de algunas bibliotecas por establecer su propio sistema de control de autoridad automatizado.

En Chile, el Sistema de Bibliotecas de la Pontificia Universidad Católica alimenta una base de datos que ya cuenta con más de 90 000 registros de autoridad de materia y más de 102 000 de autores personales y corporativos, en formato MARC y administrados por el sistema NOTIS versión 5, disponibles a través de la Red Nacional de Información Bibliográfica de ese país. La Biblioteca Nacional de Venezuela, por su parte, posee un catálogo de autoridad de materias, basado en la lista de encabezamientos de materia de la Biblioteca del Congreso de los Estados Unidos, en formato MARC y administrado también por el sistema NOTIS.

Es una práctica común en algunas bibliotecas de los Estados Unidos contratar servicios especiales con empresas como "Library Technologies" o "Blackwell North America", por ejemplo, para efectuar una depuración automática de sus encabezamientos de autores y materias y posteriormente crear los registros de autoridad correspondientes. El costo de estos servicios, así como el tiempo empleado, son relativamente bajos³⁴. Esto se debe a que en aquel país existe un alto grado de normatividad en cuanto a reglas de catalogación y uso de formatos, actuando la Biblioteca del Congreso como rectora y reguladora de los mismos. En México una opción de ese tipo hoy por hoy no tendría viabilidad, ya que:

- no hay uniformidad en el uso de las reglas de catalogación; algunas bibliotecas se guían por la primera edición de las RCAA, mientras que otras ya adoptaron la segunda edición,

- »existen al menos siete formatos tipo MARC en México, no habiendo ninguno que se considere norma nacional³⁵,
- »no existe un agencia con autoridad reconocida a nivel nacional que se responsabilice de la creación y difusión de registros de autoridad.

Algunas bibliotecas, bien por diferentes razones de tipo práctico o porque sus políticas de catalogación han sufrido diversos cambios a lo largo de los años, se han desviado de la lista impresa original de encabezamientos que usan como norma. Por ello, su propio catálogo es lo que constituye realmente su autoridad. De tal manera que resulta más práctico crear los registros de autoridad a partir de los registros bibliográficos. El procedimiento consiste básicamente en transferir electrónicamente los encabezamientos del catálogo bibliográfico a un archivo de autoridad, determinando exactamente la correspondencia de campos, subcampos e indicadores de uno a otro. Como resultado de este proceso se crean una serie de registros de autoridad que solamente contienen los encabezamientos aceptados y que requieren posterior intervención para agregar las referencias de "véase" y "véase también".

Los sistemas automatizados más versátiles admiten registros tanto de creación local como de fuentes externas. En esos casos hay que tener en cuenta la consistencia entre las fuentes, sobre todo si las bases de datos van a participar en proyectos cooperativos o redes de información.

1.5.2.—Almacenamiento de los registros de autoridad

Los registros de autoridad comúnmente se encuentran dentro de formatos definidos. Los formatos de almacenamiento de datos sirven también para facilitar el intercambio de información con otras instituciones.

La IFLA, a través de sus programas de Control Bibliográfico Universal, ha recomendado que "debe existir una infraestructura que posibilite el intercambio eficiente de datos entre las diferentes agencias bibliográficas nacionales", así como "la creación de una infraestructura que posibilite el intercambio eficiente de datos de autoridad a nivel internacional"³⁶.

Actualmente existen varios formatos para datos de autoridad, tipo MARC, definidos como normas nacionales, como el formato USMARC-Authorities (que denominaremos, en adelante, USMARC-A), de los Estados Unidos, el formato australiano AUSMARC-Authorities (basado estrechamente en el anterior) o el formato IBERMARC para autoridades (que se publicará en 1996). Como norma internacional, la IFLA publicó en 1991 el formato UNIMARC-Authorities (que denominaremos, en adelante, UNIMARC-A).³⁷

En el Reino Unido los datos de autoridad están contenidos en el formato UKMARC de intercambio de registros bibliográficos, donde aparecen referencias cruzadas en los campos 9XX. La British Library, por su parte, ha desarrollado un formato propio para registros de autoridad, basado en una versión simplificada del formato USMARC-Authorities, y mantiene una base de datos de autoridad para uso interno³⁸.

El formato USMARC-A³⁹ está diseñado para contener la información relativa a las formas establecidas de nombres, títulos uniformes, materias y subdivisiones de materia usados como puntos de acceso en registros bibliográficos, las formas de esos nombres, materias y subdivisiones de materia que deben usarse como referencias a las formas establecidas, y las interrelaciones entre esas formas. Su estructura es semejante a la del formato USMARC para datos bibliográficos (que denominaremos, en adelante, USMARC-B); así, consta básicamente de cuatro elementos: la guía, el directorio, los campos de longitud fija (o de control) y los campos de longitud variable. Los dos primeros sirven para facilitar el intercambio de los datos, mientras que los dos últimos están destinados a contener los datos para su manejo a nivel local fundamentalmente.

La guía está constituida por los elementos que proporcionan la información para el procesamiento del registro de autoridad. Estos elementos consisten en números o códigos identificados por su posición relativa. De entre ellos cabe destacar la posición del carácter 6, que indica el tipo de registro: cuando el código es "z" se trata de un registro de autoridad.

El directorio consta de una serie de entradas que contienen la etiqueta, longitud y posición inicial de cada campo dentro del registro.

Los campos se identifican por una etiqueta numérica de tres caracteres que se encuentra almacenada en la entrada del directorio para cada campo. Los campos pueden agruparse por bloques según el primer carácter de la etiqueta, que identifica la función de los datos dentro del registro. El resto de la etiqueta sirve para identificar el tipo de información que incluye un campo. Los bloques 1XX, 4XX, 5XX y 7XX guardan además ciertos paralelos en la designación de contenidos, como puede apreciarse por los dos últimos caracteres numéricos de sus etiquetas. Las figuras 1 y 2 ilustran estas características.

En el Reino Unido los datos de autoridad están contenidos en el formato UKMARC de intercambio de registros bibliográficos, donde aparecen referencias cruzadas en los campos 9XX. La British Library, por su parte, ha desarrollado un formato propio para registros de autoridad, basado en una versión simplificada del formato USMARC-Authorities, y mantiene una base de datos de autoridad para uso interno³⁸.

El formato USMARC-A³⁹ está diseñado para contener la información relativa a las formas establecidas de nombres, títulos uniformes, materias y subdivisiones de materia usados como puntos de acceso en registros bibliográficos, las formas de esos nombres, materias y subdivisiones de materia que deben usarse como referencias a las formas establecidas, y las interrelaciones entre esas formas. Su estructura es semejante a la del formato USMARC para datos bibliográficos (que denominaremos, en adelante, USMARC-B); así, consta básicamente de cuatro elementos: la guía, el directorio, los campos de longitud fija (o de control) y los campos de longitud variable. Los dos primeros sirven para facilitar el intercambio de los datos, mientras que los dos últimos están destinados a contener los datos para su manejo a nivel local fundamentalmente.

La guía está constituida por los elementos que proporcionan la información para el procesamiento del registro de autoridad. Estos elementos consisten en números o códigos identificados por su posición relativa. De entre ellos cabe destacar la posición del carácter 6, que indica el tipo de registro: cuando el código es "z" se trata de un registro de autoridad.

El directorio consta de una serie de entradas que contienen la etiqueta, longitud y posición inicial de cada campo dentro del registro.

Los campos se identifican por una etiqueta numérica de tres caracteres que se encuentra almacenada en la entrada del directorio para cada campo. Los campos pueden agruparse por bloques según el primer carácter de la etiqueta, que identifica la función de los datos dentro del registro. El resto de la etiqueta sirve para identificar el tipo de información que incluye un campo. Los bloques 1XX, 4XX, 5XX y 7XX guardan además ciertas paralelas en la designación de contenidos, como puede apreciarse por los dos últimos caracteres numéricos de sus etiquetas. Las figuras 1 y 2 ilustran estas características.

Figura 1

Identificadores de función de los datos

0XX	Información de control, números de clasificación, etc.
1XX	Encabezamientos
2XX	Referencias complejas de véase
3XX	Referencias complejas de véase también
4XX	Trazadas de véase de
5XX	Trazados de véase también de
6XX	Tratamientos específicos, notas, etc.
7XX	Entradas relacionadas
8XX	Grafías alternativas
9XX	Campos reservados para uso local

Figura 2

Identificadores del tipo de información de los datos

X00	Nombres personales
X10	Nombres corporativos
X11	Nombres de reunión
X30	Títulos uniformes
X50	Materias
X51	Nombres geográficos
X80	Subdivisiones generales
X81	Subdivisiones geográficas
X82	Subdivisiones cronológicas

Cabe señalar que a pesar de las semejanzas en la estructura de los campos de ambos formatos, USMARC-A y USMARC-B, las etiquetas iguales no siempre designan lo mismo en los dos. Por ejemplo, los campos 1XX contienen la misma información en uno y otro; sin embargo, el campo 500 es una nota general en USMARC-B mientras que en USMARC-A se trata de un trazado de "véase también de" para nombres personales.

Las figuras 3 y 4 representan un registro de autoridad (parcial) y el despliegue (parcial) correspondiente para la lista de encabezamientos de la Library of Congress.

Figura 3

Registro de encabezamiento establecido para temas

GUIA ***** nzbbb22 ***** nbb4500
 001 exb86004405b
 005 19860808072450.5
 008 860719 inban anba bnbbs bbbbs bbbab anabb bbbu
 040 bb\$e DLC
 150 bb\$a Optical disks
 450 bb\$a Discs, Optical
 450 bb\$a Disks, Optical
 450 bb\$a Optical discs
 550 bb\$wg\$a Optical storage devices

Figura 4

Entrada en la lista de encabezamientos
de materia de la LC

Optical disks (May Subd Geo)
 UF Discs, Optical
 Disks, Optical
 Optical discs
 BT Optical storage devices

Como puede observarse, el campo 550 (véase también de) genera, en este caso, una entrada para un término más general (Broader Term), lo cual viene indicado por el código "g" en la primera posición del subcampo "w". Sin entrar en profundidad en la explicación del uso de este subcampo de control, señalaremos algunas de las aplicaciones de los códigos definidos para el mismo y lo que éstas representan para la generación de la lista de encabezamientos de la Library of Congress. El subcampo "w" puede tener cuatro posiciones con códigos específicos para cada una de ellas. En caso de no ser aplicables alguno de estos códigos, o todos, las posiciones se dejan en blanco. Me interesa resaltar ahora solamente la primera posición del subcampo, representada como \$w/0, cuyos códigos representan lo siguiente⁴⁰:

- a Encabezamiento anterior
- b Encabezamiento posterior
- d Acrónimo
- g Término genérico (Broader Term)
- h Término específico (Narrower Term)

Es decir, que el encabezamiento que aparece en el campo 550 con un código "g" en el subcampo "w", como en el ejemplo anterior, representa un término más amplio que el encabezamiento del campo 150, que representa una entrada de la lista de encabezamientos. En caso de no aplicarse ninguna de los anteriores códigos se asumirá que el encabezamiento que aparece en el campo 550 es un término relacionado (Related Term) con el encabezamiento de la entrada.

En 1979 la IFLA constituyó un grupo de trabajo denominado "IFLA Working Group on an International Authority System" con el fin de facilitar el intercambio internacional de información sobre autoridades. En 1984 apareció su primera publicación "Guidelines for authority and reference entries"⁴¹, que marca los elementos que deben estar presentes en entradas de autoridad y de referencia, y el uso que debe hacerse de la puntuación de acuerdo a las convenciones de las ISBD.

Posteriormente se elaboró un formato, basado en los principios de UNIMARC, y bajo los auspicios del "Steering Group on a UNIMARC Format for Authorities". Además se añadió otro principio: los códigos de subcampo serían como los del formato bibliográfico, mientras que las etiquetas diferirían en uno y en otro, debido a las diferentes funciones de los campos en los dos formatos. Por ejemplo, los campos 2-- se dedican al área descriptiva en el formato bibliográfico, mientras que en el de autoridades contienen los encabezamientos. La primera versión del formato se circuló en forma restringida entre 1984 y 1987, publicándose una versión preliminar de circulación más amplia en 1988, y la versión definitiva en 1991. En la figura 5 aparecen los bloques de etiquetas de UNIMARC-A, indicando su función⁴².

Figura 5**Bloques funcionales del formato UNIMARC-A**

- 0-- Bloque de identificación: contiene números que identifican el registro o la autoridad.
- 1-- Bloque de información codificada: contiene elementos de longitud fija (generalmente codificados) que describen diversos aspectos del registro o de los datos.
- 2-- Bloque de encabezamientos: contiene los encabezamientos de autoridad, referencia, o explicativos, para los que se creó el registro.
- 3-- Bloque de notas: contiene notas, para conocimiento del público, que explican la relación entre el encabezamiento de un registro (2--) y otros encabezamientos.
- 4-- Bloque de trazados de "véase": contiene las variantes de los encabezamientos, que deben enviar hacia el encabezamiento para el que se creó el registro.
- 5-- Bloque de trazados de "véase también": contiene encabezamientos relacionados con el encabezamiento para el cual se creó el registro, que deben generar una referencia de "véase también" hacia el mismo.
- 6-- Bloque de números de clasificación: contiene números de clasificación relacionados con el encabezamiento del registro.
- 7-- Bloque de encabezamientos relacionados: contiene una forma del encabezamiento del registro (2--) en otro idioma o escritura, relacionándola con otro registro en el cual aparece esa forma en un campo 2--.⁴³
- 8-- Bloque de fuentes de información: contiene la fuente del registro, reglas bajo las cuales fue creado y notas catalográficas para uso interno.
- 9-- Bloque de uso nacional: contiene datos locales de la institución que crea el registro.

Entre las ventajas que presenta la adopción de un formato MARC para autoridades señalaremos las siguientes:

- Permite establecer una conexión automática entre los registros bibliográficos y los registros de autoridad.
- Posibilita el uso de constantes de despliegue, en forma de términos, frases u otras convenciones como pueden ser el espaciado, puntuación, etc., para ser generadas en forma automática.
- Su estructura interna está diseñada para posibilitar el establecimiento de interdependencias entre los datos codificados para los diferentes tipos de registros,

constituyendo así un instrumento para controlar automáticamente la consistencia de los datos dentro de los registros de autoridad.

- Al obedecer a normas internacionales permite el intercambio de datos de autoridad con otras instituciones, por medios automatizados.

Algunos sistemas funcionan con MARC y pueden además convertir, mediante programas de conversión automática de campos y subcampos, registros no-MARC provenientes de fuentes externas, en registros MARC. Independientemente del formato elegido, un registro de autoridad debe dar cabida a las formas establecidas de los encabezamientos, referencias cruzadas y cualquier otra información relevante, como las notas de alcance y la fuente del encabezamiento. Además es recomendable que el sistema posea mecanismos de validación automática que verifiquen que las etiquetas, códigos, valores, etc. previstos en el formato sean correctos y que todos los campos requeridos se hallen presentes.

En lugar de archivos de autoridad algunos sistemas poseen únicamente "diccionarios" o índices de encabezamientos. Estos índices, al carecer de referencias cruzadas, poseen una utilidad limitada. Con los nuevos desarrollos, algunos sistemas pueden proporcionar una lista de encabezamientos en forma de índice, generada a partir de los registros de autoridad, que incluye las referencias cruzadas.

En cuanto al lugar físico donde se almacena el texto de los encabezamientos ya hemos comentado que algunos sistemas lo hacen exclusivamente en el registro de autoridad⁴⁴. Ello puede representar la desventaja de que el tiempo requerido para la visualización de los registros sea mayor, ya que el sistema tiene que recuperar los encabezamientos del archivo de autoridad previamente al despliegue de los registros bibliográficos.

1.5.3.—Mantenimiento de la base de datos de autoridad

El mantenimiento de la base de datos de autoridad consiste en el proceso de actualización permanente de los registros, que se efectúa mediante la adición, supresión o modificación de los datos contenidos en ellos.

La adición de nuevos registros a la base de datos de autoridad puede realizarse durante las actividades normales de catalogación o por una actualización diaria, semanal o mensual. También se recomienda la emisión de informes periódicos sobre posibles inconsistencias o errores.

Las entradas de los registros bibliográficos pueden estar sujetas a diversas modificaciones, ya sea debido a nuevas reglas o políticas de catalogación, o como consecuencia de cambios en las listas de encabezamientos de materia. La detección y corrección de errores e inconsistencias forma también parte importante de este proceso.

En dos estudios sobre los encabezamientos de materia de la base de datos bibliográficos en línea de OCLC llevados a cabo por O'Neill y Aluri⁴⁵, y O'Neill y Vizine-Goetz⁴⁶ respectivamente, los autores categorizaron los errores que se presentaron en los encabezamientos de la siguiente manera:

1. Inconsistencias en el uso de espacios, signos de puntuación y mayúsculas
2. Errores tipográficos y ortográficos. Estos a su vez divididos en cuatro tipos:
 - omisión de caracteres
 - adición de caracteres
 - sustitución de caracteres
 - trasposición de caracteres
3. Formas no admitidas de encabezamientos
4. Etiquetas de campos o subcampos MARC incorrectas
5. Otros, que por su singularidad, no pueden tipificarse

También descubrieron que los errores que aparecían con mayor frecuencia en la base de datos de OCLC eran los del tipo 1 y 2 e idearon un algoritmo que detecta y corrige automáticamente la mayor parte de éstos.

Dicho algoritmo opera con frecuencias de uso y peso específico (basado en la fuente de la cual proviene el encabezamiento), criterios cuantitativos que ayudan a agrupar y distinguir encabezamientos semejantes. Los errores del tipo 1 pueden reducirse convirtiendo los encabezamientos a mayúsculas compactas, es decir, omitiendo los signos de puntuación, diacríticos y los espacios en blanco, proceso también llamado de "normalización". Después se construye una cadena con los primeros 17 caracteres seguidos de los 17 últimos en orden inverso. La cadena así construida sirve como base de comparación caracter por caracter entre encabezamientos. Como resultado de esta comparación la computadora emite informes que contienen los encabezamientos aparentemente erróneos, permitiendo así la reducción de los errores del tipo 2.

Se han desarrollado también, por otro lado, varios mecanismos para reducir los errores tipificados por los anteriores autores en los puntos 3 y 4.

Uno de ellos es la validación automática de campos interdependientes, generalmente fijos con variables, o de grupos de códigos preestablecidos para determinados tipos de registros o conjuntos de datos. Por ejemplo, el sistema puede verificar que un registro de autoridad contiene todos los elementos indispensables, además de determinados códigos asociados al tipo de registro de que se trate. Si el sistema posee una tabla de los valores definidos para determinados campos éstos pueden validarse automáticamente también.

Otro tipo de procedimiento que asegura la consistencia de las entradas son las llamadas "pruebas de singularidad" (uniqueness checks), con las que el sistema verifica la uniformidad de los encabezamientos aunque aparezcan en campos rentes, o detecta si un mismo encabezamiento aparece en dos registros diferentes (duplicados), o si aparece a la vez como encabezamiento admitido en un registro y como trazado de "véase" en otro.

Una forma de validación del contenido de los encabezamientos son las "pruebas de jerarquía" (hierarchical checks), en las que se verifica si las entidades que anteceden a una entidad subordinada establecida están también representadas en el catálogo de autoridad. Por ejemplo, el nombre de un país bajo el cual se asienta un organismo corporativo, o el nombre de un autor bajo el cual se asienta una de sus obras en una entrada de autor-título.

Una de las herramientas más poderosas para la actualización de la base de datos de autoridad la constituye la posibilidad de efectuar cambios globales. Los cambios globales sirven para reemplazar, modificar o borrar datos en muchos registros a la vez. Estos cambios pueden realizarse en campos enteros, incluyendo los subcampos, o en algún subcampo en particular. También es posible con algunos sistemas reemplazar o modificar cadenas de caracteres específicas independientemente del campo o subcampo en que aparezcan. Algunos sistemas sólo pueden efectuar cambios globales en los campos indizados, otros, en cambio, no requieren que los campos se hallen indizados para efectuar cambios globales.

Rogers⁴⁷, en un artículo sobre sistemas de control de autoridad en las bibliotecas de la Universidad de Ohio, menciona el desarrollo de un programa que permite borrar o reemplazar las subdivisiones de materia especificadas en tres listas. La "lista A" incluye subdivisiones que hay que borrar en todos los encabezamientos en donde aparezcan. La "lista B" incluye más de treinta nombres geográficos que deben cambiarse por formas nuevas, ya sea en subdivisiones o en encabezamientos principales.

Finalmente, la "lista C" incluye más de sesenta formas de subdivisiones abreviadas u obsoletas que deben reemplazarse por otras. Asimismo, Fox y Kanafani⁴⁸, de la Universidad de Washington en St. Louis, refieren el desarrollo de programas de cómputo que realizan cambios globales para convertir encabezamientos geográficos subdivididos en forma directa en encabezamientos con subdivisión indirecta.

La verificación de las subdivisiones es un elemento que añade complejidad al control sobre los encabezamientos de materia de un catálogo en línea.

Según Ludy la mayoría de los encabezamientos de materia del catálogo de cualquier biblioteca grande son encabezamientos con subdivisiones. Además,

según la misma autora, el 75% de los encabezamientos de materia nuevos que aparecen en la Lista de Encabezamientos de Materia de la Biblioteca del Congreso de los Estados Unidos son, de hecho, encabezamientos usados previamente pero con subdivisiones nuevas⁴⁹.

Drabenstott ha dedicado varios artículos⁵⁰ al manejo de subdivisiones generales, geográficas y cronológicas, formulando varias recomendaciones para mejorar el control y la verificación automática de las mismas. Entre estas recomendaciones cabe destacar la creación de registros de autoridad para las subdivisiones flotantes. Para las subdivisiones cronológicas asociadas a nombres geográficos, en cambio, es conveniente utilizar registros de encabezamiento establecido con subdivisión.

En los siguientes epígrafes trataremos también algunos otros aspectos relacionados con el mantenimiento de la base de datos de autoridad.

1.5.4.—Relación entre los registros de autoridad y los registros bibliográficos

Al crear una base de datos de autoridad hay que decidir qué partes del registro bibliográfico estarán bajo control. Las entradas que generalmente se someten a control son las de autores personales y corporativos, conferencias o reuniones, títulos uniformes, series y materias. Además, algunas bibliotecas establecen un control para otros elementos del registro, como son lugares de publicación o editoriales, o para cierto tipo de materiales, como los audiovisuales.

Como sabemos, algunas entradas se componen de más de un elemento, es el caso de los encabezamientos de materia con subdivisiones o de los asientos de autor-título, por ejemplo. Lo idóneo es definir varios tipos de registros para controlar cada una de las partes o éstas juntas según sea necesario.

Tal como hemos visto ya, los registros bibliográficos y los registros de autoridad de un sistema en línea pueden encontrarse separados o ligados. Aunque hay sistemas que funcionan bien a pesar de no tener ligados sus archivos, en general puede decirse que la tendencia a establecer esta conexión ha ganado recientemente terreno. Ello permite también que el usuario efectúe búsquedas en línea en el catálogo de autoridad y después obtenga los registros bibliográficos ligados a los registros de autoridad recuperados.

Además de los beneficios que se obtienen en la recuperación bibliográfica hay que tener en cuenta también la simplificación de las tareas de catalogación que permiten los sistemas automatizados de control de autoridad.

Una ventaja que pueden presentar algunos sistemas es su capacidad para validar automáticamente los encabezamientos cotejándolos con los archivos de autoridad. Si un encabezamiento no existe previamente en el archivo de autoridad,

puede ser rechazado por algunos sistemas, que no permiten su entrada en un registro bibliográfico; otros, en cambio, crean automáticamente un registro de autoridad para ese encabezamiento. Estos registros provisionales de nivel mínimo contienen el encabezamiento nuevo y los datos por "default" correspondientes, así como el esqueleto del resto de los campos, que pueden completarse posteriormente.

1.5.5.—Estructura de las referencias

Las referencias de "véase" y "véase también" sirven para relacionar encabezamientos y recuperar información que puede aparecer en diferentes formas.

Si bien es cierto que siguen siendo necesarias las referencias cruzadas aun con sistemas capaces de recuperar por palabras clave y truncamientos a la derecha, también es cierto que con tales sistemas no es necesario generar tantas referencias como en los catálogos manuales⁵¹. Por ejemplo, tanto "Educación de niños" como "Niños-Educación" pueden recuperarse en la búsqueda por palabras clave: "Niños" y "Educación". Por tanto, si alguna de las dos formas es la establecida en el catálogo en línea, por ejemplo: "Educación de niños", no es necesario elaborar una referencia de la otra forma: "Niños-Educación" véase "Educación de niños". De la misma manera "Mujer" o "Mujeres" puede recuperarse con la palabra truncada a la derecha "Mujer*". Sin embargo, en casos como: "Química biológica" véase "Bioquímica", sí será necesaria una referencia, ya que ni con truncamiento a la derecha ni buscando por palabra podríamos encontrar el encabezamiento correcto sin conocerlo de antemano.

Las referencias pueden generarse automáticamente a partir de la presencia de los campos USMARC o UNIMARC denominados 4XX ó 4-- (trazado de véase) y 5XX ó 5-- (trazado de véase también) respectivamente, en los registros de autoridad. También, a partir de determinados códigos en esos campos, pueden producirse referencias del tipo "nombre anterior:" o "nombre posterior:" para indicar las relaciones entre dos o más organismos corporativos, conferencias o nombres de series.

Asimismo, en los registros de autoridad de nivel mínimo creados automáticamente se pueden incorporar referencias también en forma automática. Según Anderson⁵² éstas se pueden producir de dos maneras: a partir de otros datos incluidos en el registro bibliográfico (cuando éste incluye campos para referencias), o a partir de cierta manipulación del encabezamiento mismo (como, por ejemplo, rotando palabras). Esto último sólo se requiere en sistemas que no cuentan con capacidad para buscar por palabras clave.

Tan importante como generar referencias cuando se necesiten es borrarlas cuando ya no sean necesarias, es decir, cuando los encabezamientos a los que estaban ligadas queden obsoletos. Si permanecen en el catálogo, las llamadas "referencias ciegas" remitirán al usuario a encabezamientos inexistentes. Algunos sistemas poseen la capacidad para detectar y borrar las "referencias ciegas" (deblinding) y pueden evitar que se produzcan mediante "pruebas de reciprocidad" (reciprocity checks). Algunas operan verificando, cada vez que se introduce un campo MARC 5XX (trazado de "véase también"), que existe un registro de autoridad que contiene ese mismo encabezamiento en el correspondiente campo MARC para encabezamiento establecido. Sin embargo, algunos autores⁵³ sostienen que son necesarias algunas referencias ciegas de "véase también" para encabezamientos de materia, con el fin de ampliar el vocabulario de las búsquedas.

Lo podemos ilustrar con el siguiente ejemplo:

150 Efectividad en costos (encabezamiento establecido)

450 Costos, Efectividad en (trazado de "véase")

550 Análisis de costos y beneficios (trazado de "véase también")

Este registro produciría las siguientes referencias: "Costos, Efectividad en" véase "Efectividad en costos", y: "Análisis de costos y beneficios" véase también "Efectividad en costos".

Incluso si la biblioteca no posee ningún documento bajo "Análisis de costos y beneficios", esa referencia puede ser útil. El despliegue de la referencia podría modificarse de la siguiente manera: "Análisis de costos y beneficios" - "la biblioteca no posee ningún documento sobre **Análisis de costos y beneficios**, usted puede buscar también bajo el término relacionado **Efectividad en costos**".

El sistema puede también rechazar una referencia de "véase" en el momento de la edición de un registro de autoridad, si ese encabezamiento ya existe en otro registro como encabezamiento establecido (campos USMARC 1XX ó UNIMARC 2--).

Algunos sistemas pueden cambiar automáticamente en los registros bibliográficos un encabezamiento no autorizado, pero ligado a otro por un "véase" en un registro de autoridad, por el encabezamiento correcto. Asimismo, el sistema puede llevar directamente al usuario que consulta el catálogo en línea hacia los registros bibliográficos correspondientes a un término no establecido, siempre que éste se encuentre ligado al término aceptado por una referencia de "véase". En estos casos, se discute sobre la conveniencia o no de que estos procesos sean "visibles" para el usuario. El que no lo sean conlleva ciertos inconvenientes; por ejemplo,

en los casos en que el usuario desconoce totalmente la relación entre los términos podría pensar que el sistema no está respondiendo adecuadamente. Crawford sostiene que la relación entre la pregunta del usuario y la respuesta del sistema debe estar siempre totalmente clara para el usuario y menciona el ejemplo de un hipotético usuario que quisiera recuperar las obras del reverendo Charles Dogson sobre lógica simbólica y no estuviera interesado en las obras aparecidas bajo el pseudónimo de Lewis Carroll, aunque se trate de la misma persona.⁵⁴

1.5.6.—Productos y evaluación

Las listas e informes impresos constituyen excelentes herramientas para el diagnóstico de la calidad de la base de datos de autoridad, facilitando la corrección y actualización de la misma. La producción regular de estas listas e informes sirve para obtener información periódica acerca de aspectos relevantes como: registros modificados, cambios automáticos efectuados, encabezamientos nuevos añadidos, etcétera.

La elaboración de estadísticas representa otro instrumento útil para la evaluación y administración del sistema. Estas estadísticas incluyen datos acerca de la creación, modificación, tamaño, etc. de los registros de autoridad. Algunas de las medidas más discutidas son: la razón matemática entre registros de autoridad y registros bibliográficos⁵⁵ y, entre referencias y registros de autoridad⁵⁶. Otros datos que pueden tomarse en consideración son:

- número de registros de autoridad de la base
- número de registros añadidos a la base por período de tiempo
- número de registros borrados de la base por período de tiempo
- número de registros modificados por período de tiempo
- número de registros bibliográficos afectados por cambios en registros de autoridad
- tamaño medio de los registros de autoridad
- número de referencias cruzadas generadas
- número de registros sin referencias

Finalmente, la producción de copias de la base de datos de autoridad, ya sea impresas o en otros formatos como cintas magnéticas, discos flexibles o disco compacto, facilita el intercambio y difusión de la misma.

2.—Antecedentes

2.1.—Proyecto de automatización y conversión retrospectiva del catálogo público de la BDCV

El proyecto de automatización de la BDCV comenzó en 1982 con el diseño de un sistema para controlar las adquisiciones por compra. En 1984 se inició la definición de un formato tipo MARC que permitió, con la máquina ALTOS de que se disponía, capturar y reproducir etiquetas y tarjetas de las obras catalogadas por la biblioteca a partir de 1986.

En 1990 se da un nuevo impulso al proyecto de automatización integral de la biblioteca con la adquisición de un manejador de bases de datos denominado STAR, que se instaló en una computadora SUN.

A partir de septiembre de 1990 la biblioteca inició la captura de los registros anteriores a 1986⁵⁷. En una primera etapa se ingresaron los registros de obras adquiridas entre 1984 y 1986. Desde abril de 1991 hasta abril de 1992 se codificaron y capturaron las obras que conforman el resto de la colección. A partir de 1992 se comenzó a catalogar en línea.

Se puede decir que hubo tres etapas de captura en el proceso de conversión retrospectiva de nuestro catálogo, con rutinas de trabajo y de revisión diferentes:

- 1) Captura de registros corrientes (obras adquiridas entre 1986 y 1992). La rutina de trabajo consistía en la catalogación sobre hojas de codificación y su posterior captura en diskettes por capturistas. La rutina de revisión incluyó un programa denominado "REVISA" que detectaba una serie de errores de captura en forma automática, tales como la ausencia de separadores de campos y subcampos, y otros errores similares, fundamentalmente de codificación. Sólo cuando este programa marcaba la ausencia de errores podía imprimirse un listado para revisión visual, denominado "ANALIZA", que era revisado por cada catalogador. Después, los capturistas corregían los errores marcados.
- 2) Captura de registros codificados previamente (obras adquiridas y catalogadas entre 1984 y 1986). Las hojas de codificación fueron capturadas

por capturistas experimentados de la biblioteca. Posteriormente a su captura los registros se sometían al programa "REVISA" y se emitía el "ANALIZA" correspondiente, el cual era revisado por un supervisor no profesional bien entrenado.

3) Codificación y captura de las obras adquiridas antes de 1984, por dos sistemas:

---procesador UNIPLEX, en máquinas ALTOS

---editor VI de máquinas SUN

La codificación estuvo a cargo de personal contratado especialmente para ese fin, que contaba con educación bibliotecaria formal y conocimientos del formato MARC, y al cual se le dio capacitación previa. Las tarjetas codificadas fueron revisadas por muestreo por un catalogador/supervisor, quien remitía las correcciones necesarias a los codificadores correspondientes. Después de haberse codificado todas las tarjetas, se comenzó el programa de captura, a cargo del mismo personal contratado para la codificación más algunos capturistas entrenados de la biblioteca. Después de la captura se pasaba el programa "REVISA" y, cuando éste ya no marcaba errores, se emitía un listado que presentaba el mismo formato que la pantalla de captura, es decir, incluyendo codificación. Con este listado se cotejaban las tarjetas correspondientes para corregir posibles errores y posteriormente se sacaba el informe "ANALIZA", que era revisado o bien por el supervisor no profesional, o bien por el supervisor/catalogador.

Una vez capturada toda la información por los procedimientos antes señalados se procedió a volcarla o "vaciarla" en el manejador STAR, que entonces ya se encontraba preparado para recibirla, es decir, poseía la definición de las bases de datos catalográficas en formato MARC.

Es indudable que toda la historia del catálogo tiene implicaciones sobre la calidad de nuestra base de datos actual, y el proceso de conversión retrospectiva no es una excepción. A pesar de todos los filtros, revisiones y controles de calidad hubo huecos por los que se "colaron" errores de todo tipo. Es obvio que una revisión por muestreo no es exhaustiva, como lo es que es materialmente imposible revisar cada tarjeta de un catálogo de trescientas mil una por una. A mi parecer, cuantos más controles se establezcan más errores se podrán evitar y, por oposición, una ausencia de control puede derivar en un auténtico caos.

Aparte las restricciones presupuestales de cada institución, todos los medios dedicados a establecer controles de calidad en la captura de registros redundarán en beneficio del catálogo automatizado. No hay que perder de vista que, aunque dentro del medio automatizado la detección y corrección de errores sea mucho

más sencilla, rápida y económica que en los catálogos manuales, en bastantes ocasiones resulta complicada, consumidora de tiempo y de recursos humanos costosos, por lo que toda labor de prevención constituye una inversión saludable para nuestras bases de datos.

2.2.—Encabezamientos de materia de la BDCV

La lista de encabezamientos de materia que se tomó como base para ir construyendo el catálogo de materias de la BDCV fue la de J. M. Kidder, publicada en 1947 por la Biblioteca Benjamin Franklin. A partir de la lista de Kidder se fueron modificando o agregando encabezamientos nuevos conforme a políticas bien definidas. Así, los pasos prescritos para la asignación de un tema se limitaban a la consulta, en el orden que se mencionan, de las siguientes fuentes:

- 1) Lista de encabezamientos de materia de Kidder
- 2) Catálogo público
- 3) Lista de encabezamientos de materia de Gloria Escamilla
- 4) Lista de encabezamientos de materia de Rovira-Aguayo
- 5) Tesoros y obras de referencia en la especialidad
- 6) Consulta a especialistas

Un auxiliar que se fue convirtiendo cada vez más en una herramienta indispensable fue el catálogo en tarjetas de temas traducidos. Este contenía las traducciones al español, en las formas establecidas por la BDCV, de encabezamientos de materia en inglés, generalmente tomados de las listas de encabezamientos de materia de la Biblioteca del Congreso de los Estados Unidos. Cuando el material que se estaba procesando aparecía ya catalogado en el National Union Catalog (NUC), o bien poseía catalogación en la fuente proveniente bien de la Library of Congress (LC) o de la British Library (BL), los temas comúnmente se consultaban en dicho catálogo y, si no aparecían, se traducían, acudiendo a las fuentes antes señaladas, pasando a formar parte de este catálogo dichas traducciones. Una fuente adicional de consulta para la traducción de temas fue el BILINDEX pero, debido a su reducido tamaño, la mayoría de las veces hubo que hacer nuevas traducciones.

Una limitación importante de nuestro sistema de encabezamientos de materia es la falta de ediciones actualizadas de nuestra lista base, que ha generado una necesidad creciente de añadir nuevos temas. Los temas nuevos no siempre han sido redactados correctamente, a la vez que no siempre se han hecho correctamente las

traducciones de los encabezamientos en inglés provenientes de la lista de encabezamientos de materia de la Biblioteca del Congreso de Washington.

En cuanto a las referencias de "véase" y "véase también" para autores y temas, éstas se reflejaban en el catálogo con tarjetas de envíos. Cuando se detectaban inconsistencias se procedía a unificar los temas y se elaboraban los envíos correspondientes. Previo a la automatización del catálogo de autoridad de materia existían en el catálogo alrededor de cinco mil tarjetas de envíos para autores y temas.

2.3.—El formato MARCOLMEX para registros bibliográficos

Cuando se inició el proceso de automatización de la BDCV en 1984 se vio la necesidad de definir un formato tipo MARC para la catalogación de materiales bibliográficos. Se estudió a fondo el formato LCMARC y se realizó una adaptación del mismo para la biblioteca. Después de dos años de análisis se definió completamente el formato llamado MARCOLMEX, vigente aún con muy pocas modificaciones, debidas principalmente a la evolución misma del formato USMARC.

Se definieron todos los campos aplicables, aunque no todos se usaron. Algunos campos, como el de "Nivel Bibliográfico", fueron definidos pero no se usaron sino hasta la automatización con STAR. El uso del formato ha estado determinado por las necesidades de la biblioteca y por las nuevas herramientas de cómputo disponibles.

2.4.—El manejador STAR

STAR es un manejador de bases de datos con prototipos desarrollados para bibliotecas que requiere de cierto nivel de programación. STAR maneja campos y registros de longitud variable; los campos pueden presentar múltiples ocurrencias, existiendo la posibilidad de añadirlas en el momento de la captura.

El manejador STAR puede instalarse en máquinas SUN y ALFA-MICRO. La BDCV optó por una SUN modelo SPARCserver 470, ahora SPARCserver 670MP., que funciona con el sistema operativo UNIX. Esta computadora pertenece al tipo de las minicomputadoras y es multiusuario, ya que puede atender hasta a 96 usuarios desde terminales simultáneamente, y además conectarse a redes locales y de ese modo ampliar el número de los usuarios que son atendidos.

2.4.1.—Recuperación de información

En cuanto a la capacidad de recuperación de información, STAR cuenta con los siguientes operadores de búsqueda:

Booleanos	
--Intersección	: AND
--Unión	: OR
--Exclusión	: NOT
Otros	
Proximidad	: NEAR (hasta cinco palabras antes y después)
	: W/n (Words/número de palabras antes y después)
	: W/P (palabras en el mismo párrafo)
	: W/O (palabras en la misma ocurrencia)
	: W/F (palabras en el mismo campo)
	: W/S (palabras en el mismo subcampo o frase)
--Adyacencia	: ADJ (palabras juntas en el orden establecido)
	: W/O (palabras juntas sin importar el orden)
---Truncamiento	: * (ilimitado, en cualquier posición)
	? (limitado, en cualquier posición)
--Intervalo	: ":" (permite la unión de varias búsquedas consecutivas)
--Prefijación	: { } (datos presentes o ausentes, en cualquier posición)
--Nulo	: " " (ausencia de datos en un campo o subcampo)
---Numéricos	: = (igual a)
	: < (mayor que)
	: < = (mayor o igual)
	: > (menor que)
	: > = (menor o igual)

Todas las búsquedas se realizan en los índices, donde las palabras aparecen en mayúsculas, sin diacríticos ni acentos.

La figura siguiente muestra un ejemplo de búsqueda en STAR.

Figura 6

Ejemplo de búsqueda en STAR

S1 GUERRILLA[S] URBAN* AND (HISPANOAMERICA OR LATINOAMERICA)

Recupera registros que contienen en cualquier parte las palabras GUERRILLA o GUERRILLAS, además de las palabras URBANA, URBANAS, URBANO, URBANOS, URBAN, etc. y, ya sea HISPANOAMERICA o LATINOAMERICA

S2 S1 AND PAIS NOT=US

publicados fuera de los Estados Unidos

S3 S2 AND LP=SAN ADJ SALVADOR

publicados en San Salvador

S4 S3 AND ED=UNIVERSIDAD NEAR SALVADOR

editados por la Universidad de El Salvador

S5 S2 AND PAIS=NQ

publicados en Nicaragua

S6 S4 OR S5

publicados en San Salvador por la Universidad de El Salvador más los publicados en Nicaragua

S7 S6 AND FPUB >=1990

publicados de 1990 en adelante

PAIS = País de publicación

LP = Lugar de publicación

ED = Editorial

FPUB = Fecha de publicación

2.4.2.—Índices, pseudoíndices, superíndices

Con STAR pueden construirse índices por campos y subcampos así como por las palabras contenidas en ellos.

Los índices de tipo "Field" (F) almacenan la información tal y como aparece en el campo o subcampo correspondiente. La recuperación a través de esos índices requiere de una búsqueda en los términos exactos. Los índices de tipo "Word" (W), en cambio, ordenan todas las cadenas de caracteres separadas por espacios, ya sean palabras, fechas u otras expresiones, que aparecen en el campo o subcampo de que se trate. Estos índices permiten la recuperación por palabras o términos, pudiendo combinarse en la búsqueda varios de ellos.

Además, STAR permite la definición de los llamados "pseudoíndices" para recuperar información presumiblemente contenida en varios campos y/o subcampos. Los principales pseudoíndices definidos en la base de datos bibliográficos de la BDCV, denominada MONOGRAFIAS, son: AUT (autores), TIT (títulos), CLASIF (clasificación), CONF (conferencias), SER (series) y TEM (temas). Estos pseudoíndices contemplan los diferentes campos y subcampos MARC correspondientes. Por ejemplo, en una búsqueda por AUT, el sistema revisará los índices de los campos MARC 100, 110, 700, 710, etcétera.

Los llamados "superíndices" hacen referencia a los índices en que la computadora debe buscar los términos adecuados después de haber rastreado en los archivos de autoridad los términos introducidos en una búsqueda. Esta forma "triangulada" de búsqueda se denomina en STAR "Supersearch".

2.4.3.—Edición de registros

Con STAR es posible crear, editar o borrar registros tanto individualmente como en forma global. Esto facilita las modificaciones, correcciones, o cualquier tipo de cambio que sea necesario en el catálogo.

Mediante la definición de "views" se pueden editar conjuntos de registros del catálogo con determinada información en un orden establecido. Los registros seleccionados aparecen en la pantalla incluyendo sólo los campos que van a ser revisados o modificados.

Los cambios globales pueden efectuarse sobre uno o varios campos o subcampos, o bien sobre todo el registro. Las principales operaciones globales que puede realizar STAR son:

- Añadir una nueva ocurrencia de un campo o subcampo (AO)
- Borrar una ocurrencia de un campo o subcampo duplicado (DD)
- Borrar una ocurrencia de un campo o subcampo con determinada información (DO)
- Modificar una o varias ocurrencias de un campo o subcampo, reemplazando la información que contenían por otra (RO)
- Reemplazar un texto por otro en cualquier campo o subcampo (RT)

2.4.4.—Emisión de informes

Con STAR se pueden producir en forma sencilla una gran variedad de informes. El usuario puede definir diferentes formatos de salida para uso permanente y además producir otros para necesidades concretas o inmediatas.

Los principales formatos de salida definidos por la BDCV son: CORTO, LARGO y MARC. El formato CORTO muestra los resultados de una búsqueda en forma de bibliografía ordenada por fecha de publicación. El formato LARGO muestra los registros bibliográficos en forma de ficha catalográfica convencional. Por último, la forma de salida MARC muestra los registros con todos los códigos MARCOLMEX asignados a los mismos por la biblioteca. Como dijimos, es posible, por ejemplo, seleccionar del formato de salida MARC algunos de los campos y subcampos incluídas y emitir un informe especial con esos datos únicamente.

Otra opción atractiva con STAR es la de emitir informes tomando como base los índices. Cuando se quiere revisar un índice normalmente la pantalla muestra la secuencia de líneas conteniendo los términos indizados junto a un "contador" que indica el número total de registros que incluyen dichos términos. Además de estos datos, en los informes denominados INDEX pueden aparecer también los números de los registros específicos y secuencias tomadas de diferentes índices, o sea, combinaciones de varios índices.

2.4.5.—Archivos de autoridad

Los archivos de autoridad se denominan en STAR archivos de "lookup". En estos se almacena la información concerniente a las entradas, referencias cruzadas, notas, etcétera.

Los archivos de autoridad sirven como referencia para validar o rechazar los encabezamientos y emitir mensajes de ayuda al catalogador en el momento de la captura. También sirven para orientar las búsquedas en el catálogo, dirigiéndolas de términos no aceptados a términos aceptados. Finalmente, mediante operaciones

de programación, pueden determinarse las instrucciones precisas para que se ejecuten algunas acciones correctivas en el catálogo.

En los apartados 4.5 y 4.6 explicaré detalladamente la estructura de autoridad y el funcionamiento con STAR de los archivos de la BDCV.

2.5.—La automatización integral con STAR

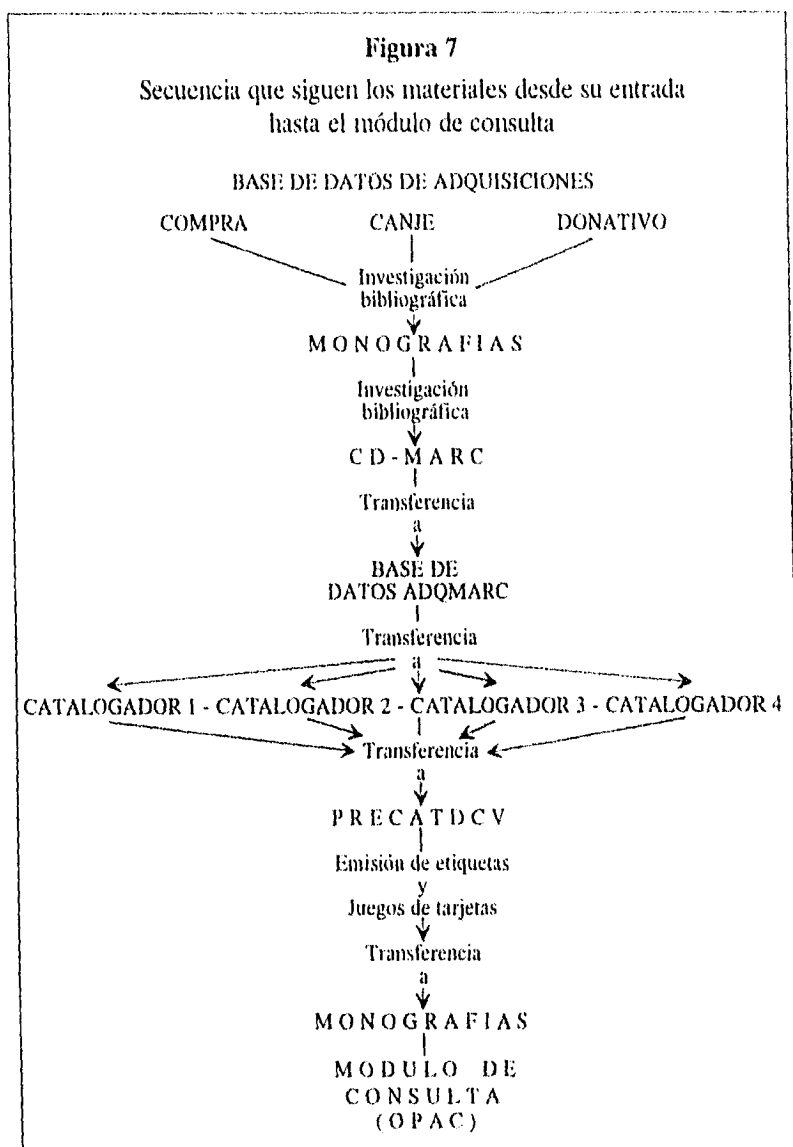
La BDCV se ha automatizado a base de "módulos" interconectados, consistentes en diversas bases de datos que interactúan entre sí. En la actualidad, se encuentran funcionando de manera integrada los módulos de adquisiciones, catalogación y consulta. La secuencia que siguen los materiales desde su entrada hasta el módulo de consulta se muestra en la figura 7.

El Departamento de Desarrollo de Colecciones acopia y turna las solicitudes de obras nuevas y recibe los materiales que ingresan a la biblioteca ya sea por compra, canje o donativo. La información relativa a esas obras se captura en las bases de datos del módulo de adquisiciones.

Antes de transferir esta información a las bases de datos de los catalogadores se hace una investigación en la base de datos MONOGRAFIAS para verificar si existen en el catálogo público otras ediciones o traducciones de la misma obra, u obras del mismo o diferente autor con temática semejante, con el fin de auxiliar al catalogador en el futuro proceso del material. En caso de existir alguna información relevante se anota el correspondiente número de registro en un papelito y se inserta en el libro o material de que se trate. Asimismo se anota a lápiz el asiento correcto para el/los autor/es. Una vez investigados los materiales en el catálogo público se procede a hacer una nueva investigación en CD-MARC. Si el material aparece ya catalogado en CD-MARC se copia el registro correspondiente en un archivo provisional, desde donde se transferirá a STAR.

Por medio de una operación global se hace el traspaso de la información correspondiente a los materiales por catalogar a las bases de datos de cada uno de los catalogadores. A estas bases llegarán, por un lado, los registros capturados en el Departamento de Desarrollo de Colecciones, con los datos de autor, título, pie de imprenta y colección; y por otro, los registros "cargados" de CD-MARC que, gracias a un programa especial de transferencia, pasarán la información codificada según el formato MARCOLMEX-B.

Hay que señalar que tuvieron que hacerse algunos ajustes al formato de edición de datos para la catalogación con STAR, debido al funcionamiento de este manejador. Así, las modificaciones principales consistieron: primero, en la colocación de los indicadores de los campos al final de los mismos, precedidos del se-



parador |a (como si se tratara del subcampo "a"), ya que STAR toma siempre la primera cadena de información de un campo como la correspondiente al subcampo "a"; y segundo, en la distinción como campos "diferentes" de los temas principales, es decir, los que aparecen en primer lugar en la ficha catalográfica convencional, con la designación 690P, 691P, 692P, etc. para éstos, a diferencia del resto de los temas, designados sin importar su orden como 690, 691, 692, etc. Esta distinción era importante, ya que los temas principales deben coincidir o concordar con la clasificación asignada al material, y además la recuperación por materias puede realizarse en forma "ponderada", es decir, pueden recuperarse solo los temas considerados principales en los documentos de una búsqueda determinada dejando afuera los que contienen dicho tema en forma secundaria. En los ejemplos que presentemos en adelante tómense en cuenta las características formales que acabamos de mencionar.

3.—*Formulación de hipótesis*

En el presente trabajo nos planteamos fundamentalmente dos objetivos: auxiliar al catalogador en el proceso de catalogación y asignación de temas en línea, y elevar la calidad de la base de datos mediante la reducción de las posibilidades de error en la misma. El objetivo final es, desde luego, facilitar al usuario la búsqueda en la base de datos y poner a su disposición un producto de mayor calidad.

El trabajo parte de una hipótesis de trabajo (A) sobre la que se basa la hipótesis formal (B) que pretendemos demostrar, a saber:

A.—Con el manejador de bases de datos STAR es posible la creación de un sistema de control de autoridad que cumpla con las siguientes funciones:

- Verificación y validación automática de encabezamientos
- Detección de errores
- Corrección de errores
- Despliegue de términos relacionados y notas de alcance
- Asistencia en el uso de subdivisiones
- Almacenamiento y transferencia de registros de autoridad en formato MARC
- Creación de registros de autoridad para autores personales, corporativos, conferencias, temas, títulos uniformes y series
- Recuperación de registros de autoridad mediante consulta a las bases de datos de autoridad
- Despliegue de registros de autoridad en diferentes formatos
- Recuperación de información temática en el catálogo bibliográfico auxiliada por los archivos de autoridad

La hipótesis que se pretende demostrar es que:

B.—Después de la aplicación de las medidas preventivas y correctivas que se especifican en el capítulo cuarto, la calidad de los registros catalogados, medida por el número promedio de errores por registro, habrá aumentado.

Para medir la calidad de los registros catalogados se propone una metodología que incluye una evaluación personal de los catalogadores.

Asimismo, basados en la literatura sobre el tema, se proponen una serie de mecanismos preventivos y correctivos que harán aumentar la calidad de los registros catalogados.

4.—El sistema de control de autoridad de la BDCV

4.1.—Diagnóstico de errores en el catálogo

El siguiente estudio de diagnóstico fue realizado dentro del marco del Proyecto de Control de Autoridad de la Biblioteca Daniel Cosío Villegas, por la que suscribe, coordinadora del proyecto, y Shirley Ainsworth, referencista de la biblioteca. El diagnóstico fue presentado en un documento interno en marzo de 1993³⁸ con el fin de proporcionar los elementos necesarios en la toma de decisiones conducentes a la limpieza del catálogo automatizado de la biblioteca para su difusión pública.

Los objetivos principales de esta depuración eran, por un lado, permitir la consulta del catálogo a través de un "OPAC" (Online Public Access Catalog) a partir de julio del mismo año y, por otro, preparar la información para la edición de un disco compacto con los registros de monografías de la biblioteca a finales del año.

El diagnóstico de errores se realizó en:

1. Registros completos
2. Índices por campos
3. Problemas específicos

Para cada uno se siguió un procedimiento diferente:

1. Procedimiento con registros completos

Supusimos que el universo de registros (aproximadamente 250 mil) era relativamente homogéneo y extrajimos dos muestras del catálogo, a partir de los registros cien mil y doscientos mil, de 80 y 73 registros respectivamente. De éstos obtuvimos las correspondientes impresiones para su revisión. Estas impresiones contenían todos los campos codificados de cada registro tal como fueron capturados.

A continuación revisamos cada uno de los registros seleccionados y anotamos en ellos todos los errores que fuimos encontrando. Decidimos agrupar los errores por tipos, de acuerdo tanto a trabajos anteriores³⁹, como al tipo de acción que se debería tomar para abatirlos. Estas acciones serán especificadas más adelante en este mismo trabajo.

FALTA PAGINA

No. 56

457

Así, consideramos errores del:

Tipo 1: Espacios, puntuación, mayúsculas, errores tipográficos y ortográficos.

Tipo 2: Codificación y etiquetado MARC.

Tipo 3: Forma y uso de las entradas.

Tipo 4: Otros.

Sin embargo, observamos que este procedimiento presentaba las siguientes limitaciones:

a) Al considerar los registros aisladamente no se podían detectar totalmente las inconsistencias entre datos que resultan de la comparación con registros semejantes o cercanos. Por ejemplo, las diferentes formas de entrada para autores, materias, etcétera.

b) Por la misma razón, no se podía diagnosticar la frecuencia total de algunos errores específicos no encuadrados en las tres primeras categorías anteriormente definidas (errores del tipo 4).

Por otro lado, no todos los errores se consideran igualmente graves, dependiendo del campo en el que se presenten. Los errores que aparecen en los campos indizados revisten mayor importancia que aquellos encontrados en los campos no indizados, o descriptivos. Así pues, vimos la necesidad de un análisis posterior con otros procedimientos.

2. Procedimiento con índices por campos

En este caso se sacaron impresiones de partes secuenciales, con longitudes arbitrarias, tomadas al azar de los índices de: autor personal (incluyendo todas las etiquetas MARCOLMEX para autor personal), autor corporativo (etiqueta 110a), conferencias (en cualquier campo MARCOLMEX para conferencias), editorial (260b), series (440a), temas generales principales (690ap)⁶⁰, temas geográficos principales (691ap), temas principales de autor personal (692ap), temas principales de autor corporativo (693ap), temas principales de conferencias (694ap) y temas principales de títulos uniformes (695ap). Estas impresiones incluyeron la entrada, seguida de la etiqueta de campo correspondiente, el número total de registros con dicha entrada y el número de cada registro involucrado. La figura 8 muestra una parte de un índice analizado.

Figura 8

Parte de un índice de conferencias analizado

PMJ on 03/08/93 at 17:15:24 Index Report

Page: 1

INTERNATIONAL CONFERENCE FOR PEACE, SECURITY AND COOPERATION (111):
 Count = 1. Record list: 125521.
 INTERNATIONAL CONFERENCE HUMAN AND SOCIAL DEVELOPMENT IN EUROPE (111):
 Count = 1. Record list: 88545.
 INTERNATIONAL CONFERENCE OF AGRICULTURAL ECONOMISTS (111):
 Count = 2. Record list: 15146, 88710.
 INTERNATIONAL CONFERENCE OF DIRECTORS OF NATIONAL LIBRARIES (111):
 Count = 1. Record list: 25141.
 INTERNATIONAL CONFERENCE OF ECONOMIC HISTORY (111): Count = 3.
 Record list: 92992, 125768, 149334.

También aquí se agruparon los errores en los tipos ya descritos. Cuando una misma entrada presentó varias formas distintas se tomó una de ellas como buena y el resto de las variantes como errores.

Cabe señalar que en la mayoría de los campos sólo se tomó el primer segmento de información (subcampo a), puesto que, por el momento y por razones de espacio, así están definidos los índices en nuestra base de datos. Sin embargo, creemos que sería más útil en el futuro tomar partes de índices con campos completos, ya que la información adicional que contienen los otros subcampos puede constituir la diferencia entre una y otra entrada.

3. Problemas específicos

Hay problemas en el catálogo que escapan a la categorización antes descrita en los tres primeros rubros y que, aunque tal vez poco frecuentes, pueden revestir importancia por afectar significativamente la recuperación ya sea de la información o del material mismo.

Este tipo de problemas aparecen a diario en la interacción continua con el catálogo y nos fueron señalados por los bibliotecarios profesionales de la BDCV al ser encuestados informalmente. Sin embargo, resultan generalmente más difíciles de "imaginar" a priori y no siempre son fácilmente cuantificables.

En general, estos problemas pueden resolverse con relativa facilidad e implican básicamente tiempo de los bibliotecarios, sin necesidad de colaboración de personal de cómputo.

Algunos de estos problemas que consideramos importantes serán detallados más adelante.

A continuación se muestran los resultados obtenidos por cada uno de los procedimientos antes descritos.

Análisis de registros completos:

Como muestra el cuadro siguiente, la distribución de errores de los tres primeros tipos fue muy similar, decreciendo ligeramente de los errores de tipo 1 (24.9 %) a los de tipo 2 (24.2 %) y los del tipo 3 (23.5 %). Los errores del tipo 4 presentaron una ocurrencia mucho menor (4.6 %). El total de los registros afectados por errores de los cuatro tipos representó el 77.1 % de la muestra. O, expresado de otra manera, la frecuencia promedio de errores por registro fue de 0.915.

--Resultados del análisis de registros completos--

	Muestra 1	Muestra 2	Muestra Total
Tamaño	73	80	153
ET1	15	36	51
RA	12	26	38
%	16.4	32.5	24.9
ET2	29	14	43
RA	24	13	37
%	32.9	16.2	24.2
ET3	20	19	39
RA	20	16	36
%	27.4	20.0	23.5
ET4	3	4	7
RA	3	4	7
%	4.1	5.0	4.6
ET	67	73	140
RA	59	59	118
%	80.8	73.8	77.1

$$FE = 140/153 = 0.915$$

FE = Frecuencia promedio de error: núm. total de errores/núm. total de registros analizados

ET1 = Núm. de errores del tipo 1 ET4 = Núm. de errores del tipo 4

ET2 = Núm. de errores del tipo 2 ET = Núm. total de errores de los cuatro tipos

ET3 = Núm. de errores del tipo 3 RA = Núm. de registros afectados

% = Porcentaje de registros afectados por error sobre la muestra total

Análisis de índices:1) **Búsqueda: autor personal**

Tamaño de la muestra	:	726	entradas
Registros relacionados	:	1635	registros
Errores	:	66	entradas
Porcentaje de error	:	9.1%	entradas
Errores por tipo:			
T1	:	35	entradas (4.8 %)
T2	:	7	entradas (1.0 %)
T3	:	21	entradas (2.9 %)
T4	:	4	entradas (0.6 %)

Nota: Hubo entradas con dos tipos de errores a la vez.

2) **Búsqueda: autor corporativo (dos muestras)**

Tamaño total de las muestras	:	223	entradas
Registros relacionados	:	1270	registros
Errores	:	78	entradas
Porcentaje de error	:	35 %	entradas
Errores por tipo:			
T1	:	32	entradas (14.3 %)
T2	:	20	entradas (9.0 %)
T3	:	26	entradas (11.6 %)

Nota: Hubo entradas con dos tipos de errores a la vez. No hubo entradas con errores del tipo 4.

3) **Búsqueda: conferencias**

Tamaño de la muestra	:	360	entradas
Registros relacionados	:	1902	registros
Errores	:	49	entradas
Porcentaje de error	:	13.6%	entradas
Errores por tipo:			
T1	:	28	entradas (7.8 %)
T2	:	15	entradas (4.2 %)
T3	:	7	entradas (1.9 %)

Nota: Hubo entradas con dos tipos de errores a la vez. No hubo entradas con errores del tipo 4.

4) Búsqueda: editorial (dos muestras)

Tamaño total de las muestras	:	272	entradas
Registros relacionados	:	1597	registros
Errores	:	57	entradas
Porcentaje de error	:	20.9%	entradas

Errores por tipo:

T1	:	29	entradas (10.7 %)
T2	:	7	entradas (2.6 %)
T3	:	22	entradas (8.1 %)

Nota: Aquí es notorio cómo las diferencias formales nos dan una enorme variedad de entradas distintas para una misma editorial (ej.: Maissonneuve, 19; Putnam's, 15; Allen and Unwin, 13). Además estas diferencias hacen que se alejen unas de otras, siendo más difíciles de detectar. Por lo mismo, el conteo de errores lo estimamos aproximado. Tómese en cuenta que este campo, en principio, no está sujeto a reglas estrictas de uniformidad en las entradas. Sin embargo, las dificultades en la recuperación podrían hacer recomendar algún tipo de control.

5) Búsqueda: serie

Tamaño de la muestra	:	315	entradas
Registros relacionados	:	1071	registros
Errores	:	82	entradas
Porcentaje de error	:	26.0%	entradas

Errores por tipo:

T1	:	46	entradas (14.6 %)
T2	:	26	entradas (8.2 %)
T3	:	10	entradas (3.2 %)

Nota: El mayor porcentaje de errores tipográficos se debe en parte a la existencia de series en alemán dentro de la muestra.

6) Búsqueda: tema general principal (dos muestras)

Tamaño total de las muestras	:	514	entradas
Registros relacionados	:	4118	registros
Errores	:	154	entradas
Porcentaje de error	:	30.0%	entradas

Errores por tipo:

T1	:	46	entradas (8.9 %)
T2	:	103	entradas (20.0 %)
T3	:	24	entradas (4.7 %)

Nota: Hubo entradas afectadas por más de un error a la vez.

- 7) Búsqueda: tema geográfico principal (dos muestras)
- | | | | |
|------------------------------|---|-------|-----------|
| Tamaño total de las muestras | : | 509 | entradas |
| Registros relacionados | : | 1337 | registros |
| Errores | : | 151 | entradas |
| Porcentaje de error | : | 29.7% | entradas |
- Errores por tipo:
- | | | | |
|----|---|----|-------------------|
| T1 | : | 34 | entradas (6.7 %) |
| T2 | : | 36 | entradas (7.1 %) |
| T3 | : | 88 | entradas (17.3 %) |
| T4 | : | 3 | entradas (0.6 %) |

Nota: Hubo entradas afectadas por más de un error a la vez.

- 8) Búsqueda: tema principal autor personal
- | | | | |
|------------------------|---|-------|-----------|
| Tamaño de la muestra | : | 102 | entradas |
| Registros relacionados | : | 256 | registros |
| Errores | : | 26 | entradas |
| Porcentaje de error | : | 25.5% | entradas |
- Errores por tipo:
- | | | | |
|----|---|----|-------------------|
| T1 | : | 10 | entradas (9.8 %) |
| T2 | : | 9 | entradas (8.8 %) |
| T3 | : | 4 | entradas (3.9 %) |
| T4 | : | 4 | entradas (3.9 %) |

Nota: Hubo entradas afectadas por más de un error a la vez.

- 9) Búsqueda: tema principal autor corporativo
- | | | | |
|------------------------|---|-------|-----------|
| Tamaño de la muestra | : | 130 | entradas |
| Registros relacionados | : | 711 | registros |
| Errores | : | 51 | entradas |
| Porcentaje de error | : | 39.2% | entradas |
- Errores por tipo:
- | | | | |
|----|---|----|-------------------|
| T1 | : | 6 | entradas (4.6 %) |
| T2 | : | 26 | entradas (20.0 %) |
| T3 | : | 19 | entradas (14.6 %) |

10) Búsqueda: tema principal conferencia

Tamaño de la muestra	: 20	entradas
Registros relacionados	: 24	registros
Errores	: 10	entradas
Porcentaje de error	: 50.0%	entradas
Errores por tipo:		
T2:	9	entradas (45 %)
T3:	1	entradas (5 %)

Nota: El tamaño de la muestra fue menor porque el tamaño del índice también lo es. No hubo errores del tipo 1 ni del tipo 4.

11) Búsqueda: tema principal título uniforme

Tamaño de la muestra	: 23	entradas
Registros relacionados	: 70	registros
Errores	: 9	entradas
Porcentaje de error	: 39.1%	entradas
Errores por tipo:		
T1	: 1	entradas (4.3 %)
T2	: 8	entradas (34.8 %)

Nota: El tamaño de la muestra fue menor porque el tamaño del índice también lo es. No hubo errores del tipo 3 ni del tipo 4.

Análisis de problemas específicos:

Los problemas que aparecieron como atípicos en los anteriores procedimientos, junto con los señalados por los bibliotecarios profesionales de la BDCV en la encuesta informal que realizamos fueron:

- Registros sin título: 19
- Registros sin número de adquisición: 37
- Signos de arroba (@) inadecuados (por problemas de transferencia de información capturada previamente en diskettes): 1983 registros.
- Temas “pegados” y revueltos (por problemas de transferencia de información capturada previamente en diskettes): aproximadamente mil registros.
- Analíticas de revistas del COLMEX con distinta signatura topográfica: detectado, pero no cuantificado.

- Registros en los que la información codificada de los campos fijos no coincidía con la información correspondiente de los campos variables: detectado, pero sólo parcialmente cuantificado.
- Duplicados (registros idénticos capturados dos veces): detectado, pero no cuantificado.

Conclusiones:

Aunque no hay un estándar específicamente definido sobre el nivel de calidad aceptable para los datos incluidos en las bases de datos bibliográficos, parece claro que éste debería marcarse por debajo de 1. Es decir, una frecuencia promedio mayor de un error por registro indicaría un nivel deficiente de calidad de la base, y cuanto más se acercara a cero dicha frecuencia la calidad de la base aumentaría.

En mi opinión, cada biblioteca debe estimar, de acuerdo a criterios de costo-beneficio, qué nivel de calidad puede alcanzar con los recursos materiales y humanos de que dispone.

En un artículo de Reeb⁶¹ sobre criterios cuantitativos para la evaluación de la calidad de la catalogación éste fija un nivel máximo deseable de error de 0.8 en promedio para los registros recién catalogados. Este nivel presumiblemente disminuye al incorporarse los registros a la base, ya que se supone que pasaron una revisión y se corrigieron los errores detectados.

En el caso de la Biblioteca Daniel Cosío Villegas cabe señalar que a los errores de proceso, que se infiltraron durante el programa de conversión retrospectiva del catálogo, se unieron los de transferencia de la información que había sido anteriormente capturada con otro sistema. De cualquier manera, consideramos que antes de ofrecer al público nuestra base de datos debíamos disminuir el nivel de error de los registros y, a partir de que presentamos nuestro informe de diagnóstico, se dedicaron recursos a este fin.

Dado que, como hemos señalado anteriormente, los errores que se presentan en los campos indizados revisten mayor importancia, la mayoría de las acciones que se pusieron en marcha estuvieron encaminadas a abatir los diferentes errores en esos campos, especialmente en los de temas, dada la alta frecuencia con la que hemos visto que se presentan las búsquedas temáticas.

Las medidas propuestas para abatir los errores que encontramos fueron:

Para los errores del tipo 1:

➤ Acciones preventivas:

Establecimiento de mecanismos de validación de datos a través de catálogos de autoridad automatizados.

➤ Acciones correctivas:

Implementación de un algoritmo de corrección de errores tipográficos en encabezamientos de materia.

Para los errores del tipo 2:

➤ Acciones preventivas:

Establecimiento de mecanismos automáticos de catalogación asistida consistentes en pantallas de ayuda que especifiquen códigos y subcampos MARC para cada campo, así como el uso de éstos de acuerdo a políticas de la biblioteca.

Mejoramiento del sistema de revisión ("ANALIZA") de registros catalogados antes de su ingreso definitivo al catálogo.

➤ Acciones correctivas:

Combinación de procedimientos automáticos (como la emisión de informes "ad hoc" resultado de búsquedas específicas, cambios globales automáticos, etc.) y manuales (revisión uno a uno de registros para seleccionarlos según tipo de problema, corrección uno a uno de registros, etcétera).

Para los errores del tipo 3:

➤ Acciones preventivas:

Creación y mantenimiento de catálogos de autoridad automatizados.

➤ Acciones correctivas:

Combinación de procedimientos automáticos y manuales.

Para los errores del tipo 4:

➤ Acciones preventivas:

Todas las ya mencionadas.

➤ Acciones correctivas:

Para los problemas de duplicados, actualizar resultado del inventario (mediante etiquetado de códigos de barras) con el sistema.

Combinación de procedimientos automáticos y manuales.

Una recomendación general de este estudio fue la de repetir este tipo de pruebas después de haber puesto en marcha las acciones propuestas, con el fin de medir su eficacia y valorar cuantitativamente la disminución del nivel promedio de errores en el catálogo.

4.2.—La catalogación asistida por STAR

No cabe duda de que la prevención de errores es tanto o más importante que la corrección de éstos. Los mecanismos de prevención de errores además de evitar un trabajo posterior de corrección ayudan en el trabajo mismo de captura de datos durante la catalogación.

STAR provee la posibilidad de programar tres niveles de ayuda para cada campo en el módulo de edición de datos. El primer nivel consiste en la aparición de una línea con información en el margen inferior de la pantalla. En el segundo nivel de ayuda se accede a una o varias pantallas que pueden redactarse libremente. El tercer nivel de ayuda consta de una pantalla que indica varios parámetros asociados con el campo, como son: nombre de la etiqueta del campo (Field label), nombre del campo (Field name), ocurrencias (Occurrences), líneas (Lines), número máximo de caracteres por línea (Max per line), tipo de datos (Data type) y regla de validación (Validation rule). La figura 9 ilustra con unos ejemplos este tercer nivel de ayuda.

Figura 9

Ejemplos de parámetros asociados con cada campo
en el tercer nivel de ayuda con STAR

Field Label	: 092	Field Label	: PAIS
Field Name	: 092	Field Name	: Pais
Occurrences	: 1	Occurrences	: 1
Lines	: 1	Lines	: 1
Max per Line	: 60 characters	Max per Line	: 2 caracteres
Data Type	: Text	Data Type	: Text
Validation rule	: Subfields ab	Validation rule	: Lookup CODIGOPAIS

El nombre del campo es el que aparece en la pantalla de edición de datos, mientras que la etiqueta es la denominación interna que se le asigna. Un campo puede tener una o más ocurrencias dependiendo de si se define éste como repetible o no. También hay que definir si el texto del campo puede ocupar una o varias líneas y el máximo de caracteres en cada línea. Asimismo se debe especificar para cada campo si los datos de entrada son de tipo textual (alfabético o alfanumérico) o numérico. La regla de validación aplicable le indicará al sistema los subcampos válidos para el campo, o con qué datos comparar los que se están introduciendo.

Esta definición de parámetros se hace al definir la estructura misma de la base datos y sirve, como dijimos, para asistir al catalogador en el momento de la captura de datos. Así, no será posible abrir una ocurrencia extra en un campo no repetible, ni escribir cuatro caracteres si el campo está definido, por ejemplo, como de tres; en un campo numérico no se permitirá introducir caracteres textuales, ni se admitirán códigos no definidos en determinados campos. En todos los casos el sistema impedirá avanzar si no se rectifican los datos erróneos y aparecerá un mensaje indicando el error, como: "El campo está lleno", "Ese código no existe. Teclee ayuda" o "Los subcampos admitidos son abc", por ejemplo.

Para el primer nivel de ayuda pusimos el nombre en extenso del campo, o una explicación breve del mismo, resaltado en forma luminosa. En el segundo nivel de ayuda redactamos textos de ayuda que especifican el uso del campo según políticas de la biblioteca, los códigos admitidos y una explicación de los mismos, todos los valores de indicadores y subcampos MARCOI.MEX, y cualquier otra ayuda o recordatorio sobre el campo. Las figuras 10 y 11 muestran ejemplos de pantallas en el segundo nivel de ayuda.

Figura 10

Ejemplos de pantallas en el segundo nivel de ayuda con STAR

ILUS Código de ilustraciones

Pueden asignarse los siguientes códigos:

- a ilustraciones
- b mapas
- h facsímiles
- o fotografías

Si el material que se está catalogando carece de ilustraciones, el campo se deja en blanco.

IDIOM Código de idioma

Los códigos de idioma se toman de las listas de la Biblioteca del Congreso de los EE.UU. Los códigos de lenguas indígenas de México se encuentran en una lista aparte. Ambos pueden consultarse también en línea en la base denominada CODIGOIDIOM.

Por ejemplo: japonés: jpn mixteco: mio

En este campo no hay valor por "default". El español se codifica: spa

Cuando el texto del material se encuentra en varios idiomas se codificará el principal de ellos (el del título o el del primer título cuando hay títulos en varios idiomas).

Figura 11
Ejemplo de pantalla en el segundo nivel de ayuda con STAR

260	Pie de imprenta	
Valor de indicadores	Código de subcampo	Descripción
0		El editor no es el asiento principal
1		El editor es el asiento principal
b		Blanco
	a	Lugar de publicación
	b	Nombre del editor
	c	Fecha de publicación
Ejemplo:		
260 México b Fondo de Cultura Económica c 1982 a 0b		

Entre las reglas de validación se incluye la definición e los archivos de "look-up" (véase epígrafe 2.4.5) que contienen todos los valores admitidos para determinados campos. Así, se capturaron los códigos de países, idiomas y temas geográficos (campo MARC 043) de las listas autorizadas de la Biblioteca del Congreso de los Estados Unidos, denominándose respectivamente CODIGOPAIS, CODIGOIDIOM y CODIGO043 a estos archivos (Véase figura 12). A éstos se añadieron también los códigos de Escamilla⁶² para lugares e idiomas de México. Dado que las listas primero mencionadas están en idioma inglés se hizo la correspondiente traducción al español y al francés, y se incluyeron algunos sinónimos conocidos en cualquiera de esas lenguas. Además, en el archivo CODIGO043, se añadió la determinante de lugar correspondiente de la Clasificación Decimal Dewey.

Figura 12

Registros de archivos de lookup

Base: CODIGOPAIS		Base: CODIGOIDIOM	
Valor	: sp	Valor	: eng
Acción a tomar	: Accept	Acción	: Accept
Mensaje de STAR	: España	Mensaje	: Inglés
Traducción	: sp	Traducción	: eng
Pais :	España	Idioma :	Inglés
Pais :	Spain	Idioma :	English
Pais :	Espagne	Idioma :	Anglais
Base: CODIGO043			
Valor	: n-mx-oa	Es lo que debe escribirse.	
Acción	: Accept	El sistema lo da como bueno.	
Mensaje	: Oaxaca	Resaltado en la parte inferior del registro editado, para confirmar.	
Traducción	: n-mx-oa a bb	Es lo que escribe el sistema en el registro bibliográfico	
Lugar	: Oaxaca	Campo que se indiza para recuperar	
CDD	: 972.72	Clasificación decimal Dewey correspondiente	

STAR permite moverse desde la pantalla de edición de datos a pantallas de búsqueda en otras bases, regresando después de la búsqueda al lugar preciso en el que estábamos. De esta manera, es posible consultar rápidamente los archivos antes mencionados en el momento de la catalogación. Las búsquedas pueden efectuarse por palabras tanto en inglés como en español y francés. Cabe mencionar también que cada vez que se introduce un código admitido aparece una línea en el margen inferior de la pantalla que confirma el significado de ese código en términos normalizados en español.

Otra opción interesante con STAR consiste en la posibilidad de que al introducir uno o varios caracteres predeterminados se escriban automáticamente los restantes (Véase figura 13).

Esto ayuda a evitar errores de "dedo" y a uniformar los datos.

Figura 13

Mensajes de ayuda que especifican caracteres que hay que introducir para obtener información completa

FCAP Fecha de entrada

Es la fecha en que se está catalogando el material. Para marcar la fecha vigeme sólo escriba // y **ENTER**. Las dos primeras cifras corresponden al día, las dos siguientes al mes y las dos últimas al año.

040 Fuente de catalogación

Se usa para indicar si la catalogación es original o se apoya en otras fuentes. Existen cuatro opciones:

- | | | |
|---|-----------------|---|
| 1 | CM c CM | Catalogación original |
| 2 | CM c CM d CM | Topográfico (nuevas ediciones o traducciones) |
| 3 | NUC c CM d CM | Copia LC (incluye CD-MARC) |
| 4 | BL c CM d CM | Copia BL |

Aquí teclee sólo el número de la izquierda. Esto es, si se quiere indicar que la catalogación es original ponga únicamente 1 seguido de **ENTER**.

043 Código de área geográfica

Este código consta de siete posiciones. Las posiciones intermedias sin códigos se llenan con guión.

Ejemplos: Afganistán: a-fg
 Africa : f

Aquí sólo anote el código correspondiente y **ENTER**, los indicadores y líneas faltantes aparecerán automáticamente.

4.3.—Aplicación de filtros semi-automáticos: el informe ANALIZA

Entre las rutinas de catalogación de la biblioteca, como ya mencionamos en el apartado 2.1, se incluía la revisión de un informe impreso denominado "ANALIZA" conteniendo los materiales catalogados diariamente por cada catalogador. Dicho informe presentaba las obras catalogadas en forma de ficha catalográfica y, por tanto, la información correspondiente a campos fijos, indicadores u otra información codificada no aparecía. Sin embargo, como ya dijimos, previo a la emisión de este informe se corría un programa llamado "REVISA", que detectaba en forma automática la ausencia de campos, indicadores, etc. Con el cambio al

actual sistema de cómputo estas herramientas quedaron obsoletas, pues ya no funcionaban eficientemente, y hubo que sustituirlas. Las funciones que cumplía el programa "REVISA" fueron sustituidas en su mayoría por los mecanismos ya explicados en el epígrafe anterior, con la ventaja adicional de realizarse desde STAR, sin necesidad de recurrir a un programa por fuera. Además, se vio la necesidad de modificar el informe "ANALIZA" adaptándolo a las nuevas características de la catalogación en línea. Dicho informe se produciría ahora también desde STAR.

El nuevo informe "ANALIZA" que diseñamos incluye toda la información codificada, en una forma legible y con una estructura que facilita la revisión.

Se definieron los campos obligatorios para los cuales, en caso de estar ausentes, aparece el siguiente mensaje:

******* ESTE CAMPO DEBE CONTENER INFORMACION *******

o bien: **FALTA INFORMACION**, dependiendo de si el campo ocupa una línea entera o parte de una línea. De igual modo, si los indicadores de los campos 1XX's, 24X's y 260 se omiten aparece el mensaje: **FALTAN INDICADORES**. Para hacer más legible este informe se separó la información en bloques divididos por líneas en blanco. A su vez, para evitar que éste fuera demasiado extenso, algunos campos de poca longitud se agruparon en una sola línea. Así, la línea CLAS (signatura topográfica) incluye los campos 049 (colección) y 092 (clasificación); y los campos fijos PAIS, Tipo de fecha, F1 (Fecha 1) y F2 (Fecha 2) aparecen también en una sola línea.

Cada registro que aparece en el ANALIZA va precedido por su número en la base de datos de cada catalogador. De esta forma es posible acceder rápidamente al registro que se debe modificar. El número de control de la biblioteca (campo 035) figura en el margen superior derecho de cada registro en orden ascendente, permitiendo una revisión rápida de la secuencia.

Algunos datos de este informe aparecen resaltados en negritas, con el fin de distinguirlos de otra información, como el 092 (clasificación Dewey asignada por la biblioteca) para distinguirlo del 082 (clasificación Dewey asignada por otras agencias), o para evitar duplicaciones, como en los campos 69X's de temas principales. A la vez, al estar resaltados CLAS y 69X's, es más sencillo comprobar si ambos concuerdan. También van resaltados en negritas los indicadores que pueden presentar variaciones relevantes (campos 1XX's, 24X's y 260), así como el campo fijo de Nivel bibliográfico y los subcampos de los campos 901 (número de adquisición) y 903 (número de código de barras).

Otra característica importante del informe es que los campos fijos asociados a determinados campos variables aparecen emparejados con éstos. Así, van juntos los campos: 24X's e ID (idioma) y 041 (obras traducidas); 260 y PAIS, Tipo de fecha, F1 y F2; 300 e ILUS; 69X's y 043 (código de área geográfica).

Además, es posible visualizar este informe en la pantalla antes de mandarlo imprimir.

Para ilustrar lo anterior reproducimos parte del informe ANALIZA (figura 14).

Figura 14

Informe ANALIZA

BEC on 01/03/94 at 18:51:07

Registro: 1968

BIBN m 009786 |a bb
 040 * * * * ESTE CAMPO DEBE CONTENER INFORMACION * * * *
 020 84-333-0021-0 |a bb
 082 709 |b C8 |a bb
 CLAS f |a bb 709 |b H673/no.14 |a bb
 100 Ballesteros Arranz, Ernesto |e dir. |a 20
 245 Cultura y arte islámicos hasta el siglo XIII |a FALTAN INDICADORES
 ID spa
 250 3 ed. |a bb
 260 Madrid |b FALTA INFORMACION |c 1981 |a 0b
 PAIS sp Tipo de fecha s Fecha 1 1981
 300 19 p. |b ilus. |c 23 cm. |a bb
 ILUS a
 440 Historia universal del arte y la cultura |v 14 |a b0
 500 Para diapositivas véase: T/190 |a bb
 690P Arte islámico |x Historia |y Hasta 1800 |a b1
 691 Países árabes |x Arte |x Historia |y Hasta 1800 |a b1
 043 ma----- |a bb
 903 39050514624P |a bb
 903 39094592984X |e ej. 2 |a bb

4.4.—Creación del catálogo de autoridad de materia

4.4.1.—El formato MARCOLMEX para registros de autoridad

El formato "Autoridades MARCOLMEX" (MARCOLMEX-A) está diseñado para contener la información relativa a las formas establecidas de nombres, materias y subdivisiones de materia usadas como puntos de acceso en registros bibliográficos, las formas de esos nombres, materias y subdivisiones de materia que deben usarse como referencias a las formas establecidas, y las interrelaciones entre esas formas.

El término "nombre" se refiere a: nombres personales, nombres corporativos, nombres de conferencias o reuniones, nombres de jurisdicciones geográficas, combinaciones nombre-título, y títulos uniformes. Un nombre puede ser la entrada principal, secundaria, de materia o de serie en los registros bibliográficos.

El término "materia" se refiere a: tópicos o temas, nombres geográficos, nombres con subdivisiones de materia, y temas y nombres geográficos con subdivisiones de materia.

El término "subdivisión de materia" se refiere a las expresiones que sirven como subdivisión general, subdivisión geográfica o subdivisión cronológica de otros temas.

En un registro de autoridad MARCOLMEX un encabezamiento es el contenido de un campo 1XX, 4XX ó 5XX, que documenta la forma del mismo para fines de indización y recuperación u organización en un archivo. La formulación de un encabezamiento en cada tipo de registro de autoridad se basa en las reglas de catalogación y de construcción de encabezamientos de materia generalmente aceptadas (p.e.: RCAA, Kidder). El resto de los contenidos en un registro de autoridad sigue las prácticas propias definidas por la BDCV.

Al igual que el formato MARCOLMEX para registros bibliográficos (MARCOLMEX-B), el formato MARCOLMEX-A está fuertemente basado en su correspondiente USMARC (versión de 1993). Esto se debe, en parte, al afán de guardar cierta consistencia entre los diferentes formatos MARCOLMEX, pero también al nivel de complejidad y flexibilidad de uso del formato USMARC-A, además de la ventaja que supone aprovechar instrumentos muy elaborados procedentes de esta frontera de donde se recibe un gran flujo de información catalográfica. Otro elemento que justifica la elección de un formato MARC para autoridades lo constituye la posibilidad de exportar registros.

Las principales diferencias entre los formatos USMARC-A y MARCOLMEX-A consisten en la omisión de algunos campos en éste último. No se añadió ningún campo numérico que no estuviera contemplado en el USMARC-A. En algunos campos se restringió el uso de determinados códigos no aplicables al país o a la BDCV. En ningún campo se añadieron códigos extra que no aparecieran en el USMARC-A. El contenido de algunos campos MARCOLMEX-A en los registros de autoridad de la BDCV puede variar de los contenidos previstos en el USMARC-A.

De igual modo que con el formato MARCOLMEX-B, en principio, se definieron el mayor número de campos y códigos aplicables, y se espera que las necesidades de la biblioteca vayan determinando el uso o no de los mismos.

En el Anexo se presenta una versión sucinta del formato MARCOLMEX-A.

4.4.2.—Archivos de autoridad con STAR

Como dijimos, los archivos de autoridad se denominan en STAR archivos de "lookup". Los archivos de "lookup" de STAR deben contener necesariamente los siguientes campos:

VALOR — Es cada uno de los textos que puede introducirse en el momento de la captura. Es un campo repetible.

ACCION — Aquí se indica al sistema si debe aceptar, rechazar, enviar un mensaje o confirmar el contenido del texto introducido.

MENSAJE — En este campo se escribe el mensaje que aparece al introducir el texto correspondiente a un campo VALOR.

TRADUCCION — Es el texto correcto que corresponde a los diferentes campos VALOR de un registro de autoridad. Este texto se copiará automáticamente en el campo correspondiente del registro bibliográfico cuando en éste se escriban cualquiera de los textos de los campos VALOR.

A continuación, voy a ilustrar brevemente este funcionamiento. Si un catalogador introduce en un registro bibliográfico, en el campo 100, regido por una regla de validación de archivo de "lookup", por ejemplo: "Cárdenas, Magolo", el sistema enviará un mensaje que diga "Encabezamiento validado", a la vez que sustituirá dicho texto por "Cárdenas, Magdalena Sofía", la forma correcta del asiento para esa autora. El registro de autoridad correspondiente se verá de la siguiente manera:

VALOR: Cárdenas, Magdalena Sofía
VALOR: Cárdenas, Magolo
ACCION: Aceptar
MENSAJE: Encabezamiento validado
TRADUCCION: Cárdenas, Magdalena Sofía

Esto quiere decir que tanto si el catalogador introduce la forma no admitida como si captura la correcta, el sistema escribirá la forma admitida y validará el encabezamiento.

Cuando tratamos de adaptar el formato MARCOLMEX-A a los archivos de "lookup" de STAR nos dimos cuenta de que existían algunos inconvenientes que había que superar. A saber:

- 1) En el campo VALOR deberían incluirse tanto los valores admitidos como los no admitidos. Esto implicaba que un campo VALOR podría equivaler tanto a un campo IXX como a un campo 4XX del formato MARCOLMEX de autoridad.
- 2) Al tener que incluir también en el campo TRADUCCION el valor correcto o admitido, se duplicaría el campo MARCOLMEX IXX que, por definición, no es repetible.
- 3) Al regirse cada campo por una sola regla de validación no es posible reunir en un único archivo todos los registros de autoridad, ya sean éstos de autor personal, corporativo, reunión o temas. STAR sólo puede verificar si el texto que se está introduciendo corresponde al contenido en un campo VALOR, pero no podría distinguir dentro de los campos VALOR si se trata de autores personales, corporativos, etcétera.
- 4) El campo MENSAJE no puede ocupar más de una línea, ni es repetible, por lo tanto no podría incluir las notas de alcance.
- 5) Si se incluían en el campo VALOR y/o TRADUCCION los indicadores correspondientes, éstos deberían ser capturados en el registro bibliográfico en forma idéntica, lo cual no es posible dado que los indicadores de los campos en el registro bibliográfico son diferentes a los indicadores que deben codificarse en el registro de autoridad.
- 6) Con la actual versión de STAR sólo es posible validar el primer subcampo de un campo. Por tanto, los archivos de "lookup", con esta versión, no admitirían registros de subdivisión ni registros con subdivisiones. Esta fue considerada la limitante más importante, pero dado que se prevé que la

próxima versión de STAR contemplará la validación de subcampos, se vio la conveniencia de no dedicar demasiado tiempo a buscar caminos alternativos sino más bien a preparar el camino futuro.

- 7) Los archivos de "lookup" de STAR, aunque interactúan con las bases de datos bibliográficos, no pueden considerarse del tipo de archivos ligados (en el sentido en que se definieron en el capítulo I de esta tesis), es decir, cualquier cambio en los archivos de autoridad no genera automáticamente un cambio en los registros del catálogo que contienen el encabezamiento modificado, sino que las modificaciones deben hacerse en forma programada y diferida.

Después de considerar estas limitaciones llegamos a la conclusión que los archivos de "lookup" podrían funcionar correctamente como archivos de autoridad con algunos ajustes y que el resultado sería un formato "cuasi-MARC", también denominado "MARCoide", pero convertible a MARC para propósitos de intercambio.

Así, se definieron los siguientes archivos de autoridad en archivos de "lookup":

- + AUTORINPER: autoridad de nombre personal. Sirve como regla de validación a los campos MARCOLMEX 100, 700, 692P (autor-tema principal) y 692.
- + AUTORINCOR: autoridad de nombre corporativo. Sirve como regla de validación a los campos MARCOLMEX 110, 710, 693P y 693.
- + AUTORINREU: autoridad de nombre de conferencia o reunión. Sirve como regla de validación a los campos MARCOLMEX 111, 711, 694P y 694.
- + AUTORIMGEN: autoridad de materia general o tópico. Sirve como regla de validación a los campos MARCOLMEX 690P y 690.
- + AUTORIMGEO: autoridad de materia de lugar geográfico. Sirve como regla de validación a los campos MARCOLMEX 691P y 691.
- + AUTORISER: autoridad de títulos uniformes y series. Sirve como regla de validación a los campos MARCOLMEX 130, 695P, 695 y 440 (serie). El funcionamiento de este archivo como catálogo de autoridad no se encuentra aún suficientemente estudiado, previéndose su implementación para una fase posterior del proyecto de control de autoridad.

En la figura 15 se presenta la hoja de captura correspondiente a la base de datos AUTORIMGEN, con una explicación de la misma, válida también para las otras bases de datos de autoridad.

Figura 15	
AUTORIMGEN : --Hoja de captura--	
RCAT : n	ENCAB : 040: c Enter
TSER : n	ENTM : 043:
SNUM : n	EVREF:
ENTP : b	USO :
ENTS : b	STATU :
FUENT : d	ROM
TIPR :	TSUBD :
GEO :	ORG :
450 :	[Texto del valor admitido]
450 :	[Texto de los sinónimos]
450 :	
450 :	
450 :	
450 :	
Acción :	a Enter (=aceptar)
Mensaje :	v Enter (=Validado. Vea posibles subd.)
150 :	Texto igual al primer 450
150ip/r :	b Enter (=b1) Indicadores para el registro bibliográfico
150ind. :	Indicadores para el registro de autoridad
Subc. :	Posibles subdivisiones. Ayudas para subdividir el tema.
550 :	
550 :	
550 :	
260 :	
360 :	
667 :	
670 :	
675 :	
678 :	
680 :	
681 :	
682 :	

Los valores que aparecen en negritas son fijos, es decir, serán siempre los mismos para encabezamientos de tópicos

En los campos con un valor seguido de "Enter", bastará escribir esto para que aparezca todo el texto, que también es fijo para este tipo de registro

Los campos 450 son los campos VALOR. El campo 150 es el campo TRADUCCION. Convencionalmente, la primera ocurrencia del campo 450 contendrá el encabezamiento admitido y será igual al campo 150. La segunda contendrá, cuando éste se conozca, el equivalente en inglés del encabezamiento. Los indicadores para el registro bibliográfico se "pegarán" en la próxima versión de STAR. Los indicadores para el registro de autoridad se "pegarán" en el programa de conversión a MARC.

En cada hoja de captura sólo se consignan los campos de notas que se aplican al tipo de registro correspondiente.

Nota: Las abreviaturas que aparecen al principio corresponden a nombres mnemotécnicos asignados a los campos fijos.

Para lograr que los mensajes de notas se puedan visualizar en el momento de la asignación de los encabezamientos, se incluyó en el formato de captura de los catalogadores algunos campos adicionales, en los cuales se copia la información de los campos correspondientes del registro de autoridad. Cabe aclarar que estos campos no se copiarán al pasar el registro catalogado al catálogo público. Así, los campos 360 (referencias complejas de véase también) y 5XX del registro de autoridad aparecerán en el registro catalográfico en un campo llamado "Véase también". Las notas contenidas en los campos 66X, 67X y 68X se verterán en el campo "Notas" del registro catalográfico. El campo "Subc." será igual en ambos registros, de autoridad y catalográfico. Ilustramos este funcionamiento en el ejemplo que sigue:

Figura 16

Registro de autoridad (resumido)

450: Cocaína

450: Cocaine

150: Cocaína

Subc.: Industria

550: Coca

667: Se aplica a la droga elaborada a partir de la coca. Para referirse a la planta use: -Coca-

Registro catalográfico (resumido)

690: Cocaína --Al escribir este texto y presionar "Enter" el sistema muestra un mensaje iluminado, en el margen inferior derecho de la pantalla, que dice: "Validado. Ven posibles subd." A la vez, se copian los mensajes siguientes:

Véase también: Coca

Notas : Se aplica a la droga elaborada a partir de la coca. Para referirse a la planta use: -Coca-

Subc. : Industria

A continuación voy a presentar un cuadro que resume los aspectos que cubre el funcionamiento de las bases de datos de autoridad que diseñamos, con ejemplos reales que ilustran las ventajas que proporcionan, tanto en el proceso de asignación de temas como en la recuperación temática.

Figura 17
 Ejemplos para la demostración del funcionamiento
 de archivos de autoridad durante la catalogación y recuperación
 con STAR

Aspecto	Ejemplo	Escribe/Encuentra
Sinónimos	Planificación familiar	Natalidad, Limitación de la
Términos en inglés	Real property	Bienes raíces Subc: Compra, venta, etc. Impuestos Valoración
Subdiv.	Biografía	Subc: No subdivida geográf. Use también como subd. No se use bajo autores
	Investigación agrícola	Agricultura Subc: Investigación
	Capuchinos	Subc: Para subd. geog. use: Capuchinos en...
Refs. de V.A.	Medicina folklórica	V.t.: Botánica médica Medicina mágica, mística y espagórica
Refs. complejas de V.A.	Bebidas	V.A.: tipos de bebidas como --Bebidas alcohólicas-- y nombres de bebidas como --Pulque--
Notas de alcance	Trabajadores agrícolas	Nota: Aquí se asientan las obras sobre los peones y asalariados del campo. Para los pequeños propietarios de tierra use --Campesinos--

Aspecto	Ejemplo	Escribe/Encuentra
Homónimos	Papas	V.t.: los nombres de Papas bajo: ---Iglesia Católica--- Nota: Para referirse a los tubérculos use: ---Patatas---
Mayúsculas, acentos y diacríticos	educacion cívica	Educación cívica
	bilenguismo	Bilingüismo
Lugares geográf.	República de China	Taiwan
Lugares	Malawi	V.t.: Rodesia y Nyasalandia Nyasalandia (Protectorado) Nota: Nombre oficial del antiguo protectorado británico de Nyasaland a partir de 1964. Para obras sobre este territorio entre 1907 y 1952 use ---Nyasalandia (Protectorado)--- Para obras sobre el período entre 1953 y 1963 use ---Rodesia y Niasalandia---
	Países nórdicos	V.t.: Suecia, Islandia, Finlandia, Dinamarca, Noruega
Seudónimos	Concolorcorvo	Carrió de la Vandera, Alonso Subc: 1715-1783 El lazarillo de ciegos caminantes
	Lepot, François	Oliva, Enrique Subc: 1923-

Aspecto	Ejemplo	Escribe/Encuentra
Autores con dificultad para determinar el apellido	José Pascual Buxó	Pascual Buxó, José Sube: 1931-
	José Agustín	Ramírez, José Agustín Sube: 1944-
Autores con nombres iguales o muy parecidos	Cardenal Richelieu	Richelieu, Armand Jean du Plessis Sube: cardenal, duque de, 1582-1642 Nota: No confundir con el mariscal Richelieu, asentado bajo: --Richelieu, Louis François Armand de Vignerot du Plessis, duque de, mariscal de Francia, 1696-1788--
	Dumas, Alejandro	Dumas, Alexandre Sube: 1802-1870. Los tres mosqueteros V.t.: Dumas, Thomas Alexandre, 1762-1806 Dumas, Alexandre, hijo, 1824-1895. La dama de las camelias
Nombres iguales o muy parecidos	Unamuno, Miguel	Sube: 1932- Nota: Autor uruguayo. Diferente al filósofo español asentado bajo: --Unamuno y Jugo, Miguel de--
Nombres difíciles de escribir y/o muy largos	Chimalpahin	Chimalpahin Quauhtlehuanitzin, Domingo Francisco de San Antón Muñón Sube: 1579-1629

Aspecto	Ejemplo	Escribe/Encuentra
Siglas de organismos	IFLA	International Federation of Library Associations and Institutions
Autores corporativos en otros idiomas	Partido Comunista Chino	Chung-Kuo Kung Ch'an Tang

Se definieron tres formatos de salida para despliegue en línea e impresión de registros de autoridad en los archivos AUTORIMGEN, AUTORIMGEO y AUTORINPER. Estos tres formatos son:

- + **REVISA.** Es el formato que sirve para la revisión de captura de los registros de autoridad. Muestra los códigos, campos y subcampos MARC.
- + **FICHAS.** Muestra los registros de autoridad con los símbolos convencionales de "x" (véase de) "vt" (véase también), etc., propios de las listas de encabezamientos impresos del tipo de Escamilla.
- + **INDICE.** Muestra en secuencia alfabética los encabezamientos admitidos. Ejemplos de registros con cada uno de los formatos se presentan a continuación.

Figura 18
Ejemplo de formato REVISA

REGISTRO: 1526

RCAT = n ENTP = b TIPR = a ENTM = a STATU = a ORG = b
 TSER = n ENTS = b GEO = d EVREF = a ROM = n 040 = CM |c CM
 SNUM = n FUENT = d EN CAB = z USO = a TSUBD = b

450 : Sociología
 450 : Sociology
 ACCION : Accept MENSAJE : Validado. Vea posibles subd.
 150 : Sociología
 150IND P/R : b1
 150INDICAD : b0
 550 : Sociología de la educación
 550 : Sociología política
 550 : Sociología rural
 550 : Sociología urbana

Nota: Las abreviaturas que aparecen al principio corresponden a nombres mnemotécnicos asignados a los campos fijos.

Figura 19

Ejemplo de formato FICHAS para autores y temas

Record: 22

Autor: Mahfuz, Naguib, 1912-
 x Mahfuz, Nayib
 Mahfuz, Najib
 Naguib Mahfuz
 Nayib Mahfuz
 Najib Mahfuz

Nota: Se incluyeron los nombres en orden directo para permitir esta posibilidad de búsqueda a los usuarios.

Record: 9

Encabezamiento: Mapas
 vt la subdivisión: --Mapas-- bajo nombres de áreas geográficas
 Cartografía
 x Maps
 -- Conservación y restauración
 Empleo
 Proyección

Figura 20

Ejemplo de formato INDICE para autores

Mahfuz, Naguib, 1912-
Maira, Antonio Augusto de Oliveira
Martín de Poechos
Oppianus
Osorio de Oliveira, José, 1900-
Pérez Opheim, Alicia, 1921-
Petrus Alfonsi, 1062-1110?
Pila, Raúl
Piskorskii, Vladimir
Quevedo y Villegas, Francisco Gómez de, 1580-1645
Quirce Balma, Carlos Manuel
Rabasa, Emilio, 1856-
Reclus, Elisée

4.4.3.—Creación de registros de autoridad de materia

Para evitar que los archivos de autoridad ocuparan un espacio excesivo en la computadora, decidimos incluir en estos archivos únicamente los registros de encabezamientos con referencias y/o notas. O sea, aquellos que sobrepasan el "nivel mínimo" de descripción. Esto implicó algunas consecuencias para el mantenimiento del sistema de control de autoridad que explicaremos más adelante.

Así pues, el catálogo de autoridad quedó constituido en primera instancia por los registros de autoridad con referencias y/o notas. Consideramos que los encabezamientos de materia cuyo uso se refleja en el catálogo público quedan validados de por sí, aunque no posean un registro de autoridad. A la hora de "exportar" registros de autoridad con fines de intercambio, todos estos encabezamientos usados en el catálogo pueden, mediante un programa de conversión, pasar a registros de autoridad de "nivel mínimo". Estos encabezamientos quedarían automáticamente codificados en virtud de la equivalencia exacta que existe entre campos del registro bibliográfico y de autoridad (en este caso, los campos a los que cada archivo de autoridad sirve como regla de validación "alimentarían" la misma).

También se copiarían los valores fijos correspondientes a cada tipo de registro. El resultado sería una cantidad de registros con algunos campos fijos codificados y el campo variable 1XX, que correspondería al encabezamiento en cuestión.

Esquemáticamente, podríamos describir el sistema de autoridad de la biblioteca en la siguiente forma:

- A) Lista impresa de encabezamientos de materia:
 - a) algunos usados por la biblioteca
 - b) otros no usados por la biblioteca
 - c) la biblioteca usa algunos que no están en la lista
- B) Catálogo bibliográfico automatizado. Encabezamientos usados:
 - a) de la lista de encabezamientos
 - b) de otras listas de encabezamientos
 - c) creados por la biblioteca
 - d) puede generar registros de autoridad de nivel mínimo en formato MARC intercambiable
- C) Archivos de autoridad automatizados. Encabezamientos de materia:
 - a) todos usados por la biblioteca
 - b) con referencias y/o notas
 - c) puede generar registros de autoridad con referencias y/o notas, en formato MARC intercambiable
- D) Catálogo de autoridad automatizado:
 - a) compuesto por registros de nivel mínimo generados automáticamente a partir del catálogo bibliográfico más los registros de autoridad con referencias y/o notas
 - b) puede generar un archivo electrónico con fines de intercambio
 - c) puede generar una lista impresa con fines de uso interno o externo

4.5.—Aplicación de un algoritmo de corrección de errores tipográficos en encabezamientos de materia

El algoritmo de corrección de errores tipográficos en encabezamientos de materia que creamos para su uso en la BDCV se basó en el ya mencionado de O'Neill y colaboradores (véase epígrafe 1.5.3).

El archivo de autoridad que se tomó como referencia fue la base de datos AUTORIMGEN, que contenía entonces 1840 encabezamientos admitidos diferentes (tomados de los campos 150 y 550, es decir, los encabezamientos admitidos como

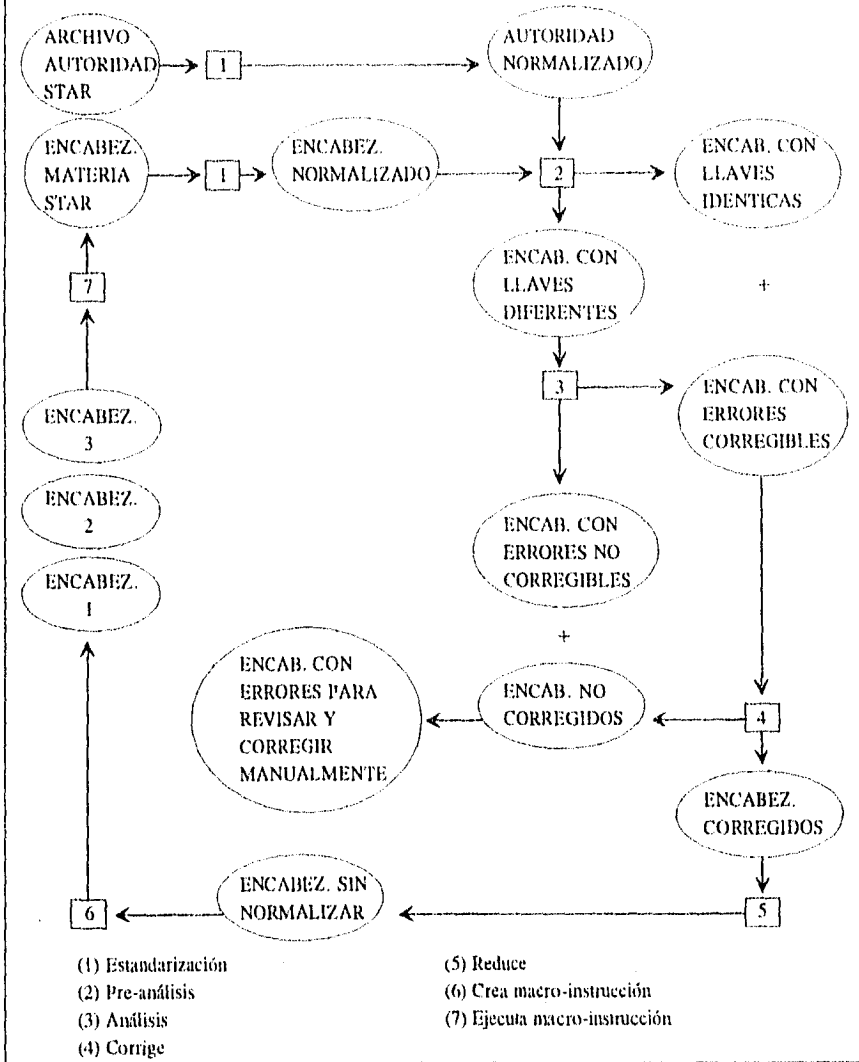
entrada en un registro de autoridad y los encabezamientos, también admitidos, a los que se remite). Este archivo fue creado con base en las tarjetas de envíos del catálogo (véase epígrafe 4.4.3), pero se completó después con los encabezamientos válidos que presentaban más de cien ocurrencias, con el fin de verificar un número mayor de encabezamientos mediante el algoritmo.

Como ya hemos mencionado la base de datos AUTORIMGEN sirve como regla de validación al primer subcampo (subcampo a) de los campos 690P y 690, por lo que extrajimos del catálogo todos los encabezamientos correspondientes al subcampo "a" de los campos 690P y 690.

La figura 21 muestra la secuencia de los diferentes procesos a los que se sometió tanto el archivo de autoridad como los encabezamientos del catálogo público seleccionados para la prueba de errores.

Figura 21

Funcionamiento del algoritmo de corrección de errores tipográficos en encabezamientos de materia



Se parte de STAR extrayendo por un lado el archivo de autoridad y por otro el conjunto de encabezamientos de materia que se van a someter al proceso de validación y corrección automática. Estos dos archivos pasan por un proceso de estandarización o "normalización" (1) que los convierte en cadenas de caracteres en mayúsculas y suprime los espacios, puntuación, paréntesis, acentos y diacríticos, creándose dos llaves, directa e inversa, que servirán como base de comparación entre encabezamientos (Véase figura 22). Así preparados, las longitudes de los encabezamientos de ambos archivos se comparan entre sí (2) dando como resultado dos archivos más.

El primero de estos archivos se compone de los encabezamientos de igual longitud con llaves directa e inversa idénticas. Esto quiere decir que se trata de los mismos encabezamientos pero que, en su forma original, pueden diferir en cuanto a espacios, puntuación, paréntesis, acentos y/o diacríticos (Véase figura 23).

El segundo de los archivos generados está formado por los encabezamientos restantes, es decir, por aquellos que difieren en la longitud y composición de sus llaves. Este último archivo se somete a un nuevo proceso de análisis (3) para detectar entre los encabezamientos otros posibles candidatos para su corrección automática. En esta ocasión, la comparación carácter por carácter de las llaves llevará a seleccionar encabezamientos en los que pudo darse la trasposición, omisión, sustitución o repetición de un carácter (Véase figura 24). Los encabezamientos seleccionados en esta forma se unirán al archivo de encabezamientos con llaves idénticas para ser procesados nuevamente con un programa de corrección (4). De allí saldrán un conjunto de encabezamientos corregidos y un archivo más de encabezamientos fallidos en el proceso de corrección (es el caso de los encabezamientos cuya repetición, trasposición, omisión o sustitución de un carácter coincide con otro encabezamiento válido y que, por tanto, no es posible determinar si es error o no). Este último archivo unido al archivo de encabezamientos en que no se encontraron candidatos para la corrección automática constituirán un archivo final de encabezamientos que requieren revisión y, en su caso, corrección manual o automática por otros procedimientos. Los encabezamientos corregidos deben ser ahora "des-estandarizados" mediante un programa (5) que los convierte en su forma original y genera un archivo que contiene una lista con los datos de: número de registro, encabezamiento incorrecto y encabezamiento correcto, para cada encabezamiento (Véase figura 25). Esta lista de encabezamientos pasa nuevamente por un proceso (6) que divide ese archivo en tres partes aproximadamente iguales y crea una macro-instrucción que será ejecutada (7) en STAR. La secuencia de esta macro-instrucción es la siguiente: busca el registro, busca el campo, busca la palabra incorrecta, borra la palabra incorrecta, escribe en su lugar la palabra correcta, inicia de nuevo el proceso con el siguiente registro de la lista.

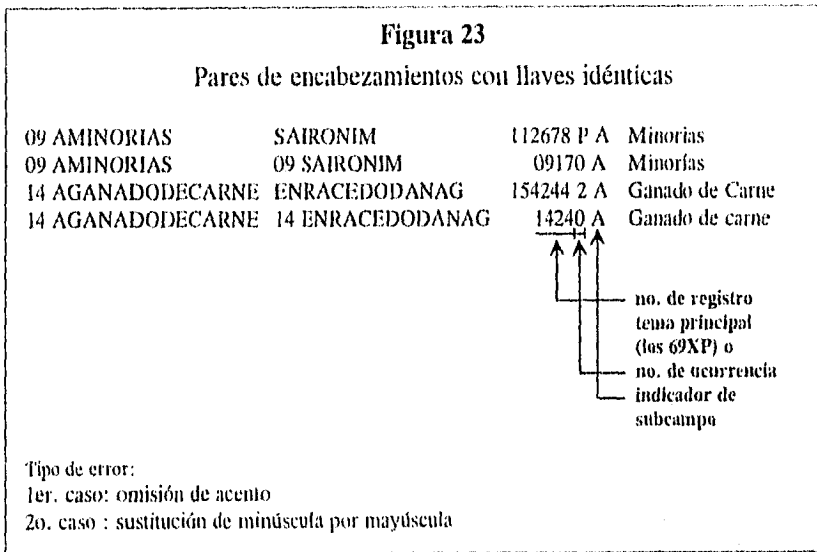
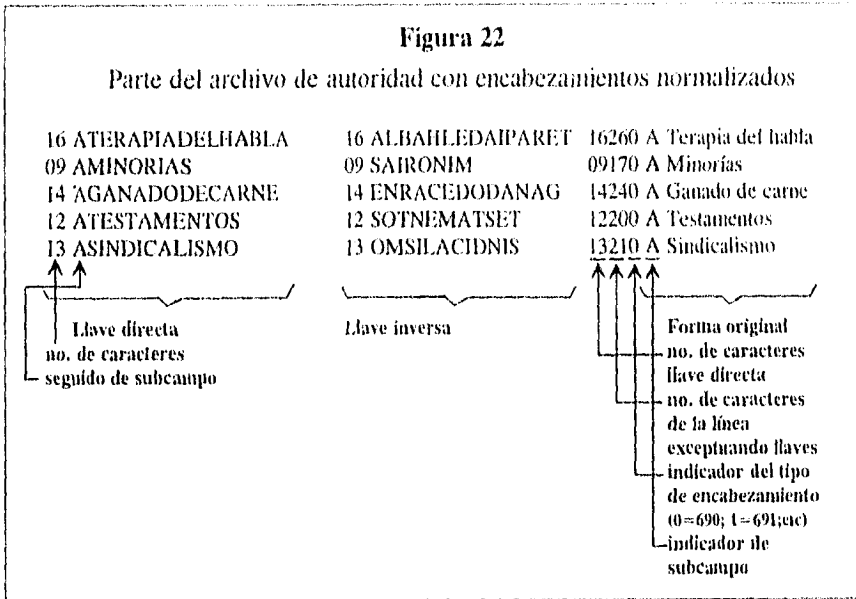


Figura 24

Pares de encabezamientos sujetos a posibles correcciones

12 ATESTAMENTOS	SOTNEMATSET	12201 1 A	Testametos
11 ATESTAMETOS	11 SOTEMATSET	12200 A	Testamentos
09 AHISTORIA	AIROTSIH	04921 P A	Histeria
09 AHISTERIA	09 AIRETSIH	09170 A	Historia

Tipo de error:
 1er. caso: omisión de un caracter
 2o. caso : sustitución de un caracter

Figura 25

Parte del archivo con pares de encabezamientos incorrecto/correcto

no. de registro	→ 102018 P	Tipo de error: repetición de un caracter
tema principal o	A Derecho penal	
no. de ocurrencia	A Derecho penal	
subcampo	↑	
	115077 3	Tipo de error: trasposición de caracteres
	A Donacoines	
	A Donaciones	

El catálogo se dividió en tres partes: 1) del registro 100,001 al 200,000; 2) del 200,001 al 285,691; y 3) del 1 al 100,000. La primera prueba real del algoritmo se realizó con la parte 1), es decir, con los registros del número 100,001 al 200,000. En esta prueba se comprobaron errores tanto de programación como de origen, es decir, provenientes del archivo de autoridad. El programa fue ajustado y subsanadas las fallas en dicho archivo para entonces volver a aplicar el algoritmo al resto de los encabezamientos del catálogo.

Se obtuvieron los siguientes resultados:

Parte 2): registros del 200,001 al 285,691

Encabezamientos analizados	:	90 245	100.00 %
Encabezamientos validados	:	55 768	61.79 %
Encabezamientos corregidos	:	715	0.80 %
Encabezamientos dudosos	:	33 762	37.41 %
Parte 3): registros del 1 al 100,000			
Encabezamientos analizados	:	107 802	100.00 %
Encabezamientos validados	:	73 470	68.15 %
Encabezamientos corregidos	:	915	0.86 %
Encabezamientos dudosos	:	33 417	30.99 %
Total: Parte 2) más Parte 3)			
Encabezamientos analizados	:	198 047	100.00 %
Encabezamientos validados	:	129 238	65.26 %
Encabezamientos corregidos	:	1 630	0.82 %
Encabezamientos dudosos	:	67 179	33.92 %

Los archivos de encabezamientos dudosos pueden contener encabezamientos de tres tipos:

- 1) encabezamientos válidos que no se encuentran en el archivo de autoridad. La solución consiste en identificarlos y añadirlos al archivo. Con base en el nuevo archivo puede lanzarse de nuevo el proceso del algoritmo para los encabezamientos del archivo de dudosos.
- 2) encabezamientos correspondientes a otro campo temático, como puede ser 691, 692, etc., o sea, mal codificados. La solución consiste en identificarlos y corregirlos, ya sea manualmente o con procesos globales, en el catálogo.
- 3) encabezamientos que difieren en más de un carácter con la forma correcta del encabezamiento. La solución consiste en identificarlos y corregirlos en el catálogo.

Cabe aclarar que el bajo índice de encabezamientos corregidos se debe en buena medida a las revisiones y correcciones masivas efectuadas en los encabezamientos correspondientes a estos índices durante el programa emergente de correcciones previo a la edición del disco compacto del catálogo de la BDCV.

Es de destacar también la elevada cifra de encabezamientos validados por este procedimiento.

4.6.—Mantenimiento del catálogo de autoridad de materia

El mantenimiento del catálogo de autoridad de materia consiste en su actualización permanente, ya sea añadiendo encabezamientos nuevos, borrando encabezamientos obsoletos, o modificando encabezamientos cuando sea necesario, como es el caso de la inclusión de referencias nuevas de "véase" y "véase también" en registros de autoridad ya existentes.

En nuestro caso particular el procedimiento para ingresar registros nuevos constaría de las siguientes fases: (véanse Figuras 26 y 27)

1) Propuesta de candidatos por parte de los catalogadores

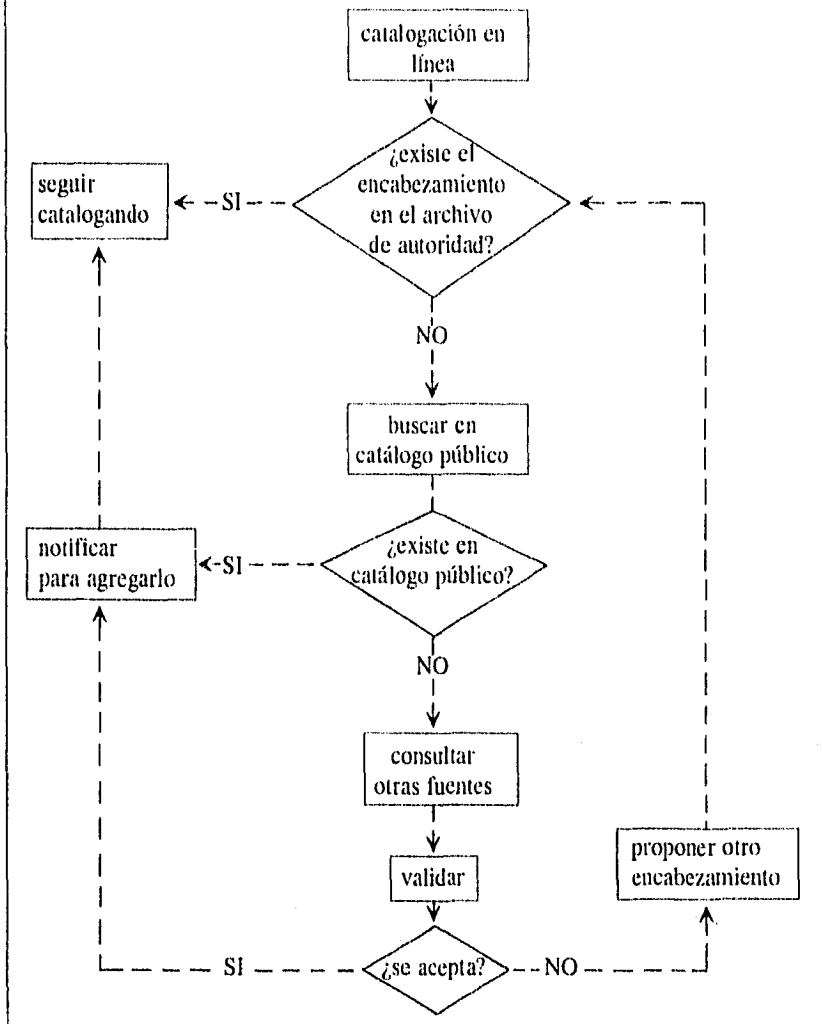
Durante la catalogación en línea, a la hora de asignar encabezamientos el catalogador recibirá una respuesta inmediata del sistema indicando si éstos se encuentran ya en las bases de datos de autoridad o no. En caso de que el encabezamiento no se encuentre en el catálogo de autoridad correspondiente, el catalogador verificará si éste ya fue usado anteriormente en el catálogo público. Como sólo se introdujeron a las bases de autoridad los encabezamientos con referencias de "véase" y "véase también" y/o notas y algunas traducciones al inglés, puede darse fácilmente la circunstancia de que el encabezamiento en cuestión haya sido usado anteriormente y no se encuentre en los archivos de autoridad. Cuando se verifica que efectivamente el encabezamiento ha sido usado en el catálogo público con anterioridad se turna una notificación a la persona que administra las bases de datos de autoridad para que se proceda a su captura. En caso de no encontrar el encabezamiento usado en el catálogo público se consultará otras fuentes para asegurarse de que no se trata de otra forma de un encabezamiento válido sino de uno verdaderamente nuevo. Una vez verificado que se trata de un encabezamiento nuevo se pasará para captura. Cuando se compruebe que el encabezamiento por usar debe ser otro se procederá con éste de la misma manera que con el anterior.

2) Captura de encabezamientos por el administrador de las bases de datos de autoridad

Cuando se quiera ingresar un encabezamiento nuevo a las bases de datos de autoridad lo primero que se hará es buscar en el catálogo público para verificar si existen una o varias formas equivalentes de ese encabezamiento y con qué frecuencia. Una vez detectadas las posibles inconsistencias se validará la forma elegida del encabezamiento y se procederá a su captura en un registro de autoridad nuevo. Posteriormente se corregirán las inconsistencias encontradas en el catálogo

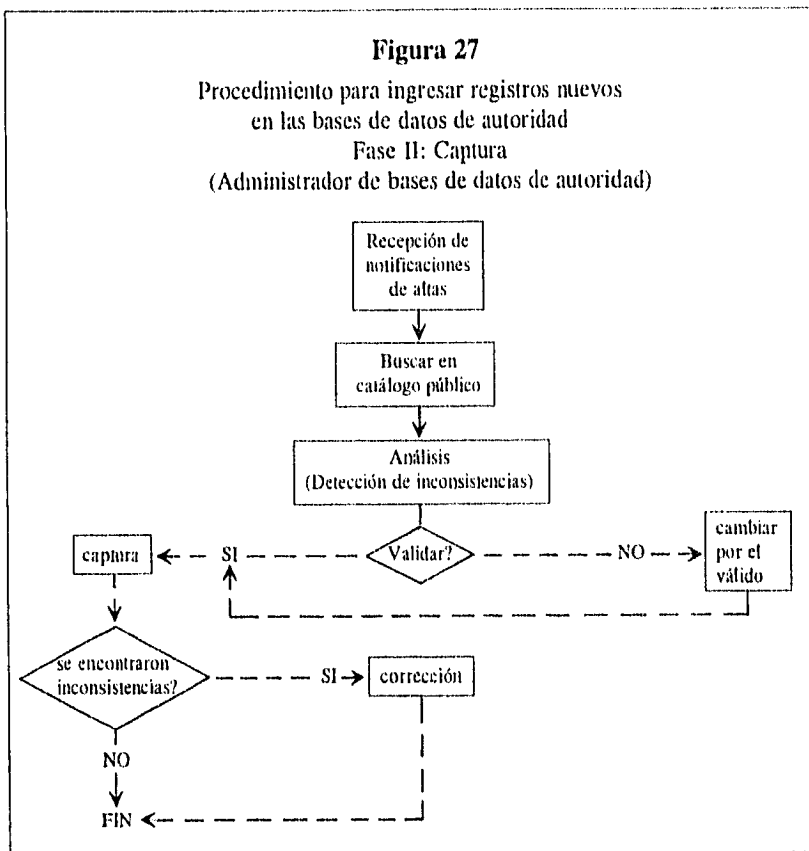
Figura 26

Procedimiento para ingresar registros nuevos
en las bases de autoridad
Fase I: Propuesta de candidatos
(catalogadores)

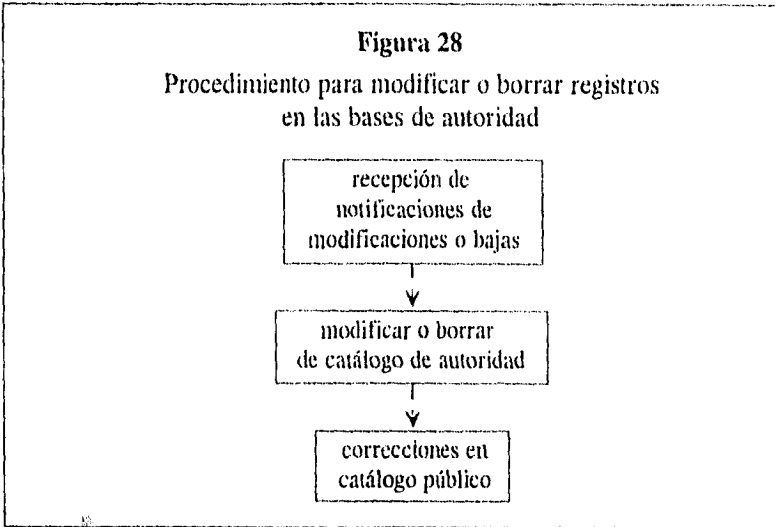


público en cuanto al uso de ese encabezamiento. Se recomienda capturar el encabezamiento en un registro de autoridad nuevo antes de empezar a corregir en el catálogo para así evitar la entrada de encabezamientos incorrectos entre la fase de corrección y captura.

Una interesante posibilidad que brinda el manejo automatizado de los catálogos de autoridad es la de introducir encabezamientos nuevos, con la indicación de ser registros provisionales, y someterlos a una fase de prueba, antes de darlos por definitivos. La facilidad de poder efectuar cambios globales nos permite actualizar los registros bibliográficos correspondientes cuando se presente algún cambio en un registro de autoridad.



En cuanto a la modificación o eliminación de registros de la base se deben seguir los pasos indicados en la figura 28



Conviene destacar que STAR brinda la posibilidad de implementar un mecanismo automático que imposibilita que se produzcan duplicados e inconsistencias en las bases de datos de autoridad. En el momento de actualizar cualquier registro el sistema verifica que no exista previamente otro registro que contenga idéntica información en los campos 1XX's y 4XX's. Esto quiere decir que no es posible crear un registro idéntico a otro ya existente, ni crear un registro que contenga en un campo 1XX (encabezamiento establecido) la misma información que otro registro en un campo 4XX (referencia de véase).

Es importante mantener al corriente de todos los cambios de la base de datos de autoridad a los usuarios de la misma, principalmente catalogadores y referencistas. Para ello se generaría un informe bimestral, ya que los informes del Departamento de Control Bibliográfico se realizan con esa periodicidad y no es previsible que la frecuencia mensual de movimientos en la base sea muy alta. El informe se generaría automáticamente a partir de los campos de fecha de creación del registro y de última actualización del registro, así como de la información sobre registros borrados que conserva la computadora, y contendría la información referente a altas, bajas y modificaciones de registros de autoridad para ser

circulado en los Departamentos de Control Bibliográfico y de Servicios al Público.

El informe puede contener además los datos sobre el número total de registros en las bases de autoridad y el número total de encabezamientos admitidos contenidos en la base (campos 1XX's + 5XX's). Esta última cifra nos indicará cuántos encabezamientos del catálogo pueden ser sometidos al proceso de corrección de errores tipográficos en forma automática pues, como vimos, el "archivo de autoridad" que usa el algoritmo de corrección de errores tipográficos descrito en el epígrafe anterior se vale de estos encabezamientos.

4.7.—Intercambio de datos de autoridad de materia

Dado que fue necesario adaptar el formato MARC de autoridades para su óptimo funcionamiento con el manejador STAR, se requirió también definir los equivalentes correspondientes para fines de intercambio. En efecto, aunque muy similar al formato MARC, podemos decir que el resultado con STAR fue un formato MARCoide, convertible a MARC mediante unas cuantas operaciones.

Así pues para grabar una cinta MARC de datos de autoridad hay que:

- ♦ prescindir de los campos "Acción", "Mensaje" y "Subc.",
- ♦ omitir la primera ocurrencia de los campos 4XX's
- ♦ "pegar" los campos siguientes:
 - 1XX's ind + 1XX'1XX (menos 150)
 - 1XX's ind + 4XX's = 4XX (menos 450)
 - 1XX's ind + 5XX's = 5XX (menos 550)
- ♦ añadir antes del resto de los campos (excepto 045, 050, 082 y 083) los indicadores "bb" (blanco, blanco)

Además, habría que separar los registros que contengan traducciones al inglés de los encabezamientos (en ese caso aparecen siempre en la segunda ocurrencia de los 4XX's) para que pasen codificadas en los campos 7XX's correspondientes.

4.8.—Evaluación de resultados

Después de haber puesto en marcha todos los mecanismos de control de autoridad antes expuestos se procedió a repetir la prueba de errores en el catálogo para verificar en qué medida éstos habían sido subsanados. Para ello, se extrajo una muestra al azar de registros completos del catálogo y se realizó un análisis similar al del primer diagnóstico de errores. Se dividieron los errores en los mismos tipos

pero se añadió una categorización más para indicar si el error afectaba o no afectaba la recuperación. Este análisis arrojó unos resultados inesperados, ya que, exceptuando los errores tipográficos, todos los tipos de errores presentaron un aumento con respecto a la primera prueba. Analizando las causas de estos resultados, aparentemente contradictorios, llegamos a las siguientes conclusiones:

- ▣ hubo errores que no se consideraron la primera vez y la segunda sí (debido a cambios de políticas y usos de catalogación),
- ▣ hubo diferencias en la forma de computar errores la primera y la segunda vez,
- ▣ para poder comparar ambas muestras se debería proceder a analizarlas de nuevo con criterios iguales.

En la práctica esta última alternativa nos pareció complicada, por lo que creímos que tampoco los resultados serían muy confiables. Así es que decidimos extraer una muestra únicamente de los registros más recientes, para así medir el nivel de calidad alcanzado en los registros producidos recientemente.

En esta ocasión se seleccionó para un análisis de los registros completos una muestra aleatoria con el procedimiento de muestreo sistemático. Así, se partió de un número elegido al azar de una tabla de números aleatorios y se extrajo a partir de éste un total de 200 registros, avanzando a un intervalo de 34 registros. Expresado de otra forma:

Universo muestral $N = 6,768$ (todos los registros catalogados del 15 de julio al 30 de noviembre de 1994)

Tamaño de la muestra $n = 200$

Intervalo entre elementos de la muestra $N/n = k = 34$

El primer elemento de la muestra fue el registro 284,563 y al rebasar el número total de registros se continuó sumando al primer elemento elegible el resto del anterior.

Definimos como "error" toda desviación de las normas tipográficas, ortográficas, gramaticales o catalográficas, por mínima que ésta fuera, así como cualquier contravención de las normas y políticas de catalogación vigentes en la BDCV.

También en esta oportunidad se clasificaron los errores en los cuatro tipos ya definidos, es decir:

- Errores del tipo 1): errores tipográficos y ortográficos
- Errores del tipo 2): errores de codificación y etiquetado MARC
- Errores del tipo 3): errores en la forma y uso de las entradas
- Errores del tipo 4): otros

Además, clasificamos los errores en las categorías de:

- Errores del tipo A): errores en campos indizados y que afectan la recuperación del documento
- Errores del tipo AA): errores en campos que se rigen por catálogos de autoridad (autores y temas)
- Errores del tipo AN): errores en campos que no están sujetos a validación en catálogos de autoridad
- Errores del tipo B): errores en campos indizados que no afectan la recuperación del documento
- Errores del tipo C): errores en campos no indizados, o descriptivos, y que en ningún caso afectan la recuperación del documento.

Nuevamente, la revisión de los registros se hizo con base en la impresión de los mismos incluyendo todos los elementos codificados. Los registros fueron revisados por mí misma, con la colaboración de Brunilda Carretero, catalogadora veterana del Departamento de Control Bibliográfico.

En el análisis se investigó y verificó cada entrada sujeta a reglas de uniformidad, como son autores y temas, para no pasar por alto errores no aparentes. Asimismo, en algunos casos, se recurrió al documento original para verificar datos dudosos.

Los resultados fueron los siguientes:

Errores del tipo 1:	55
Errores del tipo 2:	46
Errores del tipo 3:	8
Errores del tipo 4:	25

TOTAL: $134 / 200 = 0.67 =$ índice de error promedio por registro

Errores del tipo B:	36
Errores del tipo C:	34
Errores que no afectan recuperación:	70
Errores del tipo AA:	23
Errores del tipo AN:	41

Errores que afectan recuperación (A): $64 / 200 = 0.32 =$ índice de error promedio por registro, afectando recuperación.

También se analizó el comportamiento por catalogador. Así:

Resultados por catalogador

	C1	C2	C3	C4	C5	C6	C7
Errores de todo tipo	18/40 = 0.45	12/26 = 0.46	28/44 = 0.64	41/35 = 1.17	9/8 = 1.12	20/25 = 0.8	6/22 = 0.27
Errores que afectan recuperación	13/40 = 0.32	5/26 = 0.19	17/44 = 0.39	11/35 = 0.31	4/8 = 0.5	10/25 = 0.4	4/22 = 0.18

En primer lugar, podemos apreciar que el índice de error promedio por registro (0.67) es menor al encontrado en el primer análisis (0.915). Esto quiere decir que los registros producidos recientemente presentan un nivel de calidad superior, lo cual es atribuible a todos los mecanismos preventivos que hemos expuesto anteriormente.

Como se ve, los errores más frecuentes siguen siendo los tipográficos, seguidos de los de codificación. En cuanto a los índices de error promedio, se aprecia que los errores que afectan la recuperación constituyen algo menos de la mitad del total de errores, variando esta proporción en el análisis por catalogador, donde se observan variaciones personales.

Cabe señalar que los catalogadores con índices más altos, por encima del promedio de 0.32, son precisamente los de más reciente ingreso al Departamento, es decir, los menos experimentados y, por tanto, sujetos a mayor riesgo de cometer errores.

Para la evaluación personal de los catalogadores sería necesario medir los índices de error en los "ANALIZA" sobre una base diaria, semanal, mensual, etc., o sobre muestras suficientemente amplias de los registros catalogados por un individuo en un período de tiempo determinado.

Comparando los índices de error computados en los "ANALIZA" con los índices de error encontrados posteriormente en el catálogo para un mismo período de tiempo, podremos tener una estimación de en qué medida éstos disminuyen y fijar nuestros niveles tolerables de error en la etapa previa al paso de los registros al catálogo.

Las diferencias entre los índices "antes" y "después" indicarían además, por un lado, el esfuerzo necesario para obtener determinados niveles de calidad en los registros, y por otro, la efectividad de la revisión previa de registros. Un índice previo menor al índice de error encontrado en el catálogo posteriormente estaría indicando una revisión deficiente.

5.—Conclusiones y recomendaciones

❖ No se puede hablar de patrones absolutos de medición de la calidad de las bases de datos, ya que éstos deben definirse en cada caso particular atendiendo a diferentes elementos como son las necesidades de los usuarios, las normas nacionales e internacionales y las políticas particulares de las instituciones.

❖ La calidad de las bases de datos se mide, entre otras cosas, por la ausencia de errores de diversos tipos. La medición de errores es sólo una medida más de la calidad de las bases, pero no debe ser la única.

❖ Aunque no hay un estándar específicamente definido sobre el nivel de calidad aceptable para los datos incluidos en las bases de datos bibliográficos, parece claro que éste debería marcarse por debajo de 1. Es decir, una frecuencia promedio mayor de un error por registro indicaría un nivel deficiente de la base, y cuanto más se acercara a cero dicha frecuencia la calidad de la base aumentaría.

❖ Definimos como error toda desviación de las normas tipográficas, ortográficas, gramaticales o catalográficas, así como cualquier contravención de las normas y políticas de catalogación vigentes en una institución determinada.

❖ Los errores más graves son los que pueden afectar la recuperación del documento. Otros errores también deben evitarse, aunque son menos importantes.

❖ Los datos capturados con escasa calidad afectan negativamente la productividad del sistema, mermando la precisión, integridad y/o consistencia de las bases de datos. Además, los errores pueden tener repercusiones económicas derivadas del aumento en el tiempo de búsqueda o de la generación de solicitudes de compra innecesarias.

❖ El efecto multiplicador de los errores, que pasan de unas bases a otras, puede provocar su perpetuación en las mismas.

❖ Aunque en el medio automatizado la detección y corrección de errores sea mucho más sencilla, rápida y económica que en los catálogos manuales, en muchas ocasiones resulta complicada, consumidora de tiempo y de recursos humanos.

❖ La prevención de errores es tanto o más importante que la corrección de éstos. Los sistemas de catalogación asistida, además de facilitar el trabajo de captura de los datos, constituyen mecanismos de prevención de errores que evitan el trabajo posterior de corregirlos.

❖ Los algoritmos de corrección de errores tipográficos representan un mecanismo eficaz no sólo para corregir este tipo de errores sino para obtener información valiosa acerca del número de encabezamientos correctos y de encabezamientos que requieren otro tipo de correcciones.

❖ Existen diversas técnicas de control de calidad en la producción de las bases de datos que van desde las más altas tecnologías, como el uso de complicados sistemas expertos, hasta rutinas de baja tecnología, siendo lo más recomendable la combinación de soluciones en diferentes niveles.

❖ Una forma de evaluar la calidad de las bases de datos bibliográficas consiste en extraer una muestra de registros al azar por el procedimiento de muestreo sistemático, contabilizar los errores encontrados y calcular el índice de error promedio por registro resultado de dividir el número de errores entre el número de registros revisados.

❖ Una medida más fina puede obtenerse contabilizando únicamente aquellos errores que afectan la recuperación, obteniéndose así el índice de errores promedio por registro que afectan la recuperación del documento. Con base en estas medidas se pueden hacer comparaciones entre muestras tomadas en diferentes tiempos de la misma base de datos (siempre que los criterios de medición de errores no varíen en ese período de tiempo), o entre diferentes bases de datos (sin olvidar respetar las normativas específicas de cada institución).

❖ La evaluación de los catalogadores puede hacerse teniendo en cuenta, además de la productividad, la calidad de los registros catalogados, medida por los índices de error promedio por registro.

❖ Los catalogadores noveles presentan un mayor riesgo de cometer errores. Por ello, es necesario una capacitación intensiva en áreas tales como rutinas específicas de los sistemas automatizados, políticas particulares de la biblioteca, etc. Debe evitarse la revisión cruzada de registros recién catalogados entre catalogadores de nuevo ingreso.

❖ Es recomendable incluir entre las políticas de catalogación de la biblioteca la definición de normas y criterios que rijan la redacción de encabezamientos nuevos.

❖ El diseño de un sistema de control de autoridad automatizado debe buscar soluciones creativas basadas en los supuestos teóricos, en las enseñanzas de otras experiencias similares, las características del "hardware" y "software" disponibles y las necesidades propias de la biblioteca o institución que lo vaya a usar.

El contar con un sistema de control de autoridad automatizado adecuadamente diseñado contribuirá a obtener una base de datos de mayor calidad.

❖ Los elementos que han de tenerse en cuenta para la implementación de un sistema de control de autoridad automatizado son los siguientes:

1) Fuente de los registros de autoridad. Pueden ser los catálogos manuales previos a la automatización, o catálogos transferidos electrónicamente, o una combinación de ambas opciones. En cualquier caso habrá que tener en cuenta la consistencia de las fuentes y el trabajo de edición posterior requerido.

2) Almacenamiento de los registros de autoridad. Independientemente del formato elegido, un registro de autoridad debe dar cabida a las formas establecidas de los encabezamientos, referencias cruzadas y cualquier otra información relevante, como las notas de alcance y la fuente del encabezamiento. También es conveniente poseer registros aparte para encabezamientos simples, encabezamientos con subdivisiones y subdivisiones flotantes. Para propósitos de intercambio de datos de autoridad es fundamental poseer un formato tipo MARC.

3) Mantenimiento de la base de datos de autoridad. La adición, modificación o supresión de registros en la base de datos de autoridad puede realizarse durante las actividades normales de catalogación o por una actualización diaria, semanal o mensual.

También se recomienda la emisión de informes periódicos para detectar posibles inconsistencias o errores.

4) Relación entre los registros de autoridad y los registros bibliográficos. En general, puede decirse que es más conveniente que los archivos bibliográficos y de autoridad se encuentren ligados.

Las principales ventajas que ello puede reportar son:

- producción de cambios automáticos en el catálogo bibliográfico como consecuencia de haberse efectuado cambios en el de autoridad,

- validación automática de encabezamientos durante la catalogación, y
- recuperación directa de registros bibliográficos a través del catálogo de autoridad.

5) Estructura de las referencias. Incluso en los catálogos en línea con capacidad para realizar búsquedas por palabras clave, búsquedas booleanas y truncamientos, es necesaria una estructura de referencias cruzadas. Las referencias cruzadas sirven para relacionar encabezamientos y para recuperar aquéllos que pueden presentarse bajo múltiples formas, ya que las computadoras no pueden reconocer asociaciones entre términos que no estén explícitamente ligados.

6) Productos y evaluación. Las listas e informes impresos constituyen herramientas para el diagnóstico de la calidad de la base de datos de autoridad, facilitando la corrección y actualización de la misma.

❖ El hecho de que no exista, como sería deseable, una agencia con autoridad reconocida a nivel nacional que se responsabilice de la creación y difusión de registros de autoridad, provoca la duplicación de esfuerzos y la falta de uniformidad en los criterios para la creación y mantenimiento de catálogos de autoridad que podrían ser compartidos por las bibliotecas del país.

NOTAS

- 1 Arlene G. Taylor, "Authority files in online catalogs: an investigation of their value", *Cataloging and Classification Quarterly*, v.4, no.3, Spring 1984, p.2.
- 2 Jean Dickson y Patricia Zadner, "Authority control and the authority file: a functional evaluation of LCNAE on RLIN", *Cataloging and Classification Quarterly*, v.9, no.3, 1989, pp.57-58.
- 3 Véase: Barbara B. Tillet, "Considerations for authority control in the online environment", *Cataloging and Classification Quarterly*, v.9, no.3, 1989, pp.3-4.
- 4 No encuentro un equivalente en español que me convenza. Traduciendo el término de la forma más literal podría ser "trabajo de autoridad", pero no me parece que exprese claramente en español lo mismo que en inglés. Por eso prefiero usar la forma entrecomillada del inglés.
- 5 Marta Stiefel Ayala, Reynaldo Ayala y Jesús Lau, *Technical dictionary of Library and Information Science = Diccionario técnico de Bibliotecología y Ciencias de la Información: English/Spanish, Spanish/English*. New York, Garland, 1993. p. 21.
- 6 José Martínez de Sousa, *Diccionario de Bibliología y Ciencias afines*. Madrid, Fundación Germán Sánchez Ruipérez, 1989. p. 478.
- 7 Gernot Wersig y Ulrich Neveling (comps.), *Terminology of documentation = Terminología de la Documentación: a selection of 1,200 basic terms published in English, French, German, Russian and Spanish*. Paris, UNESCO, 1976. p. 103. Negritas en el original.
- 8 *Encyclopedia of Library and Information Science*, v.2, p. 132.
- 9 *Op. cit.*, Negritas en el original.
- 10 *Op. cit.*, v.2, p.136.
- 11 Arlene G. Taylor, *art. cit.*, p.2.

- 12 Arlene G. Taylor, Margaret F. Maxwell y Carolyn O. Frost, "Network and vendor authority systems", *Library Resources and Technical Services*, v.29, April-June 1985, pp.195-205.
- 13 David R. McDonald, "Data dictionaries, authority control, and online catalogs: a new perspective", *Journal of Academic Librarianship*, v.11, September 1985, p. 221.
- 14 Michele I. Dalehite, "Vendor- supplied automated authority control: what it is and how to get it", *Law Library Journal*, v.81, no.1, Winter 1989, pp.120-121.
- 15 Véase por ejemplo: Robert H. Burger, *Authority work: the creation, use, maintenance, and evaluation of authority records and files*. Littleton, Libraries Unlimited, 1985, p.10.
- 16 Edward T. O'Neill y Diane Vizine-Goetz, "Quality control in online databases", *Annual Review of Information Science and Technology*, v.23, 1988, pp.125-156.
- 17 Elizabeth S. Aversa, "Quality control in database production: a presentation on some present concerns and future prospects", trabajo presentado en el *Segundo Foro de Información Latinoamericana LATINBASE'93*, Guadalajara, 28 de nov. a 1 de dic. de 1993, p.10.
- 18 Peter S. Graham, "Quality in cataloging: making distinctions", *The Journal of Academic Librarianship*, v.16, n.4, 1990, pp.213-218.
- 19 Doris H. Clack, "Authority control and linked bibliographic databases", *Cataloging and Classification Quarterly*, v.8, 1988, p.41.
- 20 Simone Klugman, "Failures in subject retrieval", en: Holley, Robert P., ed. *Subject control in online catalogs*. New York, Haworth Press, 1989, p.10.
- 21 Tomado de: F. F. Martínez Arellano. *Impacto del uso de un catálogo en línea de una biblioteca universitaria*. México, Universidad Nacional Autónoma de México, 1993. Tesis para optar al grado de Maestría en Bibliotecología, pp. 30-32.
- 22 Alexis J. Jamieson, Elizabeth Dolan y Luc Declerck, "Keyword searching vs. authority control in an online catalog", *Journal of Academic Librarianship*, v.12, no.5, November 1986, pp.277-283.
- 23 Adeline Wilkes y Antoinette Nelson, "Subject searching in two online catalogs: authority control vs. non-authority control", *Cataloging and Classification Quarterly*, v.20, no.4, 1995, pp.57-79.

- 24 Nirmala S. Bangalore, "Authority files in online catalogs revisited", *Cataloging and Classification Quarterly*, v.20, n.3, 1995, p.89.
- 25 F.W. Lancaster, "Vocabulary control in information retrieval systems", en *Advances in Librarianship*, ed. por Melvin J. Voight y Michael H. Harris. New York, Academic Press, 1977, pp.1-40.
- 26 James Tilio Maccaferri, "Managing authority control in a retrospective conversion project", *Cataloging and Classification Quarterly*, v.14, nos.3-4, 1992, p.145.
- 27 Sarah Hager Johnston, "Current offerings in automated authority control: a survey of vendors", *Information Technology and Libraries*, September 1989, pp.236-237.
- 28 Michele I. Dalehite, *art. cit.*, p.120.
- 29 Véase, por ejemplo: R. Conrad Winke, "Discarding the main entry in an online cataloging environment", *Cataloging and Classification Quarterly*, v.16, no.1, 1993, pp.53-70.
- 30 David R. McDonald, *art. cit.*, p.221.
- 31 Sarah Hager Johnston, *art. cit.*
- 32 Es el caso de Dynix e Inlex.
- 33 Véase: Dale Irwin, "Local systems and authority control", *Cataloging and Classification Quarterly*, v.16, no.2, 1993, p.62.
- 34 Por ejemplo, para una base de datos que contenga 300.000 registros el proceso concluido para autores y materias se entrega en una seis semanas con un costo aproximado de cinco mil dólares.
- 35 Véase: Roberto Garduño Vera. *Los formatos MARC y CCF y su aplicación en unidades de información mexicanas*. México, UNAM, 1990, p.25.
- 36 Tom Delsey, "Authority control in an international context", en: Tillet, Barbara B., ed., *Authority control in the online environment: considerations and practices*. New York, Haworth Press, 1989. pp.14-15.
- 37 International Federation of Library Associations and Institutions. Universal Bibliographic Control and International MARC Programme. *UNIMARC/Authorities: universal format for authorities*. New York, K. G. Saur, 1991.
- 38 Ellen Gredley y Alan Hopkinson. *Exchanging bibliographic data: MARC and other international formats*. Ottawa, Canadian Library Association, 1990, p.105.

- 39 Véanse: *USMARC format for authority data*. Cum. vol. (base text and updates 1 and 2). Washington DC, Library of Congress Network Development and MARC Standards Office, 1989; y Update 3, 1990.
- 40 Existen además los códigos "f", "i" y "n" para esa posición, pero su aplicación no está relacionada con el ejemplo que queremos ilustrar.
- 41 International Federation of Library Associations and Institutions. *Guidelines for authority control and reference entries*. London, IFLA International Programme for UBC, 1984.
- 42 Tomado de: International Federation of Library Associations and Institutions. Op. cit., p.17.
- 43 Los bloques 6-- y 7-- fueron intercambiados por error en la edición del formato. Véanse en la obra citada las p.32 y 67-69.
- 44 Es el caso, por ejemplo, del sistema control de autoridad de la Biblioteca Nacional de Escocia. Véase: Robin Frederick Guy, "Evolution of automation in a national library: the experience of the National Library of Scotland from 1978-1989", *Program*, v.24, n.1, January 1990, p. 11.
- 45 Edward T. O'Neill y Rao Ahuri. *A method for correcting typographical errors in subject headings in OCLC records*. Dublin, OCLC Office of Research, 1980.
- 46 Edward T. O'Neill y Diane Vizine-Goetz, "Computer generation of a subject authority file", *Proceedings of the ASIS Annual Meeting*, v.19, 1982, pp.220-223.
- 47 Sally A. Rogers, "Automated control of subject headings at the OSU libraries", *Information Technology and Libraries*, March 1989, p.80.
- 48 Judith A. Fox y Kay Kanafani, "Global change capabilities to improve authority control in an online catalog", *Information Technology and Libraries*, September 1989, p.276.
- 49 Lorene E. Ludy, "OSU libraries' use of Library of Congress Subject Authorities File", *Information Technology and Libraries*, v.4, n.2, p.159.
- 50 Véanse: Karen M. Drabentstott, "The need for machine-readable authority records for topical subdivisions", *Information Technology and Libraries*, v.11, n.2, June 1992, pp.91-104.
"Facilitating geographic subdivision assignment in subject headings", *Library Resources and Technical Services*, v.36, n.4, 1992, pp.411-425.
"Period subdivisions in the Library of Congress Subject Headings System:

- some thoughts and recommendations for the future", *Cataloging and Classification Quarterly*, v.15, n.4, pp. 19-45.
- 51 Véase: Catherine M. Thomas, "Authority control in manual versus online catalogs: an examination of 'see' references", *Information Technology and Libraries*, v.3, December 1984, p. 393-398; M. R. Watson y Arlene G. Taylor, "Implications of current reference structure for authority work in online environments", *Information Technology and Libraries*, v.6, 1987, p.10-19; y Jennifer A. Younger, "After Cutter: authority control in the twenty-first century", *Library Resources and Technical Services*, v.39, n.2, April 1995, pp.133-141.
- 52 Douglas Anderson, "Automatically generated references in minimal-level authority records", *Information Technology and Libraries*, v.10, n.4, December 1991, p.252.
- 53 Véase: Karen Markey Drabenstott, ed. *Subject authorities in the online environment: papers from a Conference Program held in San Francisco, June 29, 1987*. Chicago, American Library Association, 1991. pp.39-40 y 54-55.
- 54 Véase: Walt Crawford. *The online catalog book: essays and examples*. New York, G.K. Hall, 1992. pp.16-17.
- 55 Chu, Clara M. et al., "The case of Tweddle-dee and Tweddle-dum: authority control in a retrospective database", *The Canadian Journal of Information Science*, v.12, n.1, 1987, p.47.
- 56 Lorene A. Ludy, art. cit., p.155.
- 57 Véase: Clotilde Tejeda Rodríguez y Alvaro Quijano Solís, "Conversión retrospectiva: piedra angular de la automatización de la biblioteca", ponencia presentada al Seminario Anual ABIESI 1991.
- 58 Pilar María Moreno Jiménez y Shirley Ainsworth, *Informe de diagnóstico de errores en el catálogo automatizado de la BDCV*. Documento interno de trabajo. México, BDCV, 1993.
- 59 Véase: Edward T. O'Neill, y Diane Vizine-Goetz. "Computer generation of a subject authority file", *Proceedings of the ASIS Annual Meeting*, vol. 19, 1982, pp. 220-223.
- 60 Llamamos "tema principal" al primer encabezamiento de materia del registro. La notación 69Xap significa el subcampo "a" de un tema principal (indicado por la letra "p").

- 61 Richard Reeb, "A quantitative method for evaluating the quality of cataloging", *Cataloging and Classification Quarterly*, v.5, no. 2, Winter 1984, pp. 21-26.
- 62 Véase: Gloria Escamilla González, *México: lista y códigos de lenguas indígenas, códigos de áreas geográficas*, México: CONACYT, 1981.

BIBLIOGRAFÍA

- Aliprand, Joan M., "Linking of alternate graphic representation in USMARC authority records", *Cataloging and Classification Quarterly*, v.18, no.1, 1993, pp.27-62.
- Anderson, Douglas, "Automatically generated references in minimal-level authority records", *Information Technology and Libraries*, v.10, no.4, 1991, pp.251-262.
- Autoridades: un formato MARC; especificaciones para las cintas magnéticas que contienen registros de autoridad*. Edición preliminar que incluye hasta el addendum 8. México, Consejo Nacional de Ciencia y Tecnología, 1982.
- Aversa, Elizabeth S., "Quality control in database production: a presentation on some concerns and future prospects", trabajo presentado en el *Segundo Foro de Información Latinoamericana LATINBASE'93*, Guadalajara, 28 de noviembre al 1 de diciembre de 1993. México, Universidad Nacional Autónoma de México, Centro de Información Científica y Humanística, 1993.
- Ayres, F.H. et al., "QUALCAT: automation of quality control in cataloguing", *British Library Research and Development Department Report*, 6068, 1994.
- Baer, Nadine L. y Karl E. Johnson, "The state of authority", *Information Technology and Libraries*, June 1988, pp.139-153.
- Bangalore, Nirmala S., "Authority files on online catalogs revisited", *Cataloging and Classification Quarterly*, v.20, no.3, 1995, pp.75-94.
- Bechtel, Joan M., "An authority control alternative for small colleges", *College and Research Libraries*, November 1992, pp.485-498.
- Belonogov, G. G. et al., "Automating the discovery and correction of errors in text", *Scientific and Technical Information Processing*, v.18, no.4, 1991, pp.50-52.
- Burger, Robert H. Authority control. New York, Haworth, 1996.
- _____. *Authority work: the creation, use, maintenance, and evaluation of authority records and files*. Littleton, Colo. Libraries Unlimited, 1985.

- Byrne, Deborah J. *MARC manual: understanding and using MARC records*. Englewood, Libraries Unlimited, 1991.
- CannCasciato, D., "Growing while catching up: some benefits of an authority backlog", *OCLC Systems and Services*, v.9, no.4, 1993, pp.24-27.
- Carter, J.A., "PRISM/PASSPORT, questioning authority files", *OCLC Systems and Services*, v.10, no.2-3, 1994, pp.25-30.
- Chamis, Alice Yanosko. *Vocabulary control and search strategies in online searching*. New York, Greenwood, 1991.
- Chan, Lois Mai. *Library of Congress subject headings: principles and application*. 2 ed. Littleton, Colo., Libraries Unlimited, 1986.
- _____, Phyllis A. Richmond y Elaine Svenonius. *Theory of subject analysis: a sourcebook*. Littleton, Colo., Libraries Unlimited, 1985.
- Chu, Clara M. et al., "The case of Tweddle-dee and Tweddle-dum: authority control in a retrospective database", *The Canadian Journal of Information Science*, v.12, no.1, 1987, pp.35-57.
- Cinneide, C.N., "A foolish consistency...: the obsolescence of authority control in the online catalogue", *An Leabharlann*, v.9, no.4, 1993, pp.99-116.
- Clack, Doris H., "Authority control and linked bibliographic databases", *Cataloging and Classification Quarterly*, v.8, 1988, pp.35-46.
- _____. *Authority control: principles, applications, and instructions*. Chicago, American Library Association, 1990.
- _____, "Authority control: issues and answers", *Technical Services Quarterly*, v.33, no.1-2, Winter 1985-86, pp.127-140.
- Coyne, Fumiko H., "Automated authorities maintenance at the Western Library Network", *Technical Services Quarterly*, v.5, no.1, 1987, pp.33-47.
- Crawford, Walt. *The online catalog book: essays and examples*. New York, G.K. Hall, 1992.
- Cuadra Associates, Inc. *Using STAR. Version 3.2*. Los Angeles, Ca., 1993.
- Cuadra Associates, Inc. *Managing STAR databases. Version 3.1B*. Los Angeles, Ca., 1992.
- Dalehite, Michele I., "Vendor-supplied automated authority control: what it is and how to get it", *Law Library Journal*, v.81, no.1, Winter 1989, pp.117-129.
- Di Lauro, Anne y Maureen Sly. *Guidelines for the building of authority files in development-information systems*. Ottawa, International Development Research Centre, 1985.

- Drabensstott, Karen Markey (ed.) *Subject authorities in the online environment: papers from a Conference program held in San Francisco, June 29, 1987*. Chicago, American Library Association, 1991.
- _____, "Facilitating geographic subdivision assignment in subject headings", *Library Resources and Technical Services*, v.36, no.4, 1992, pp.411-425.
- _____, "The need for machine-readable authority records for topical subdivisions", *Information Technology and Libraries*, v.11, no.2, 1992, pp.91-104.
- _____, "Period subdivisions in the Library of Congress Headings System: some thoughts and recommendations for the future", *Cataloging and Classification Quarterly*, v.15, no.4, 1992, pp.19-45.
- Escamilla González, Gloria. *Lista de encabezamientos de materia*. 2 ed. México, Instituto de Investigaciones Bibliográficas, Universidad Nacional Autónoma de México, 1978.
- Eyler, W. y M. Mooney, "Designing an innovative authority control system in a PC environment", *OCLC Systems and Services*, v.10, no.2-3, 1994, pp.38-41.
- Fox, Judith A. y Kay Kanatani, "Global change capabilities to improve authority control in an online catalog", *Information Technology and Libraries*, September 1989, pp.273-283.
- García Hidalgo, María Isabel y Laura Freidberg Gojman. *Más allá de los manejadores de bases de datos: una aplicación bibliográfica*. México, El Colegio de México, 1991.
- García López, Carlos et al., "Control en los registros catalográficos a través de los catálogos de autoridad automatizados", ponencia presentada al VII Coloquio sobre Automatización de Bibliotecas. Colima, Universidad de Colima, 1995.
- Garduño Vera, Roberto. *Los formatos MARC y CCF: su aplicación en unidades de información mexicanas*. México, Universidad Nacional Autónoma de México, Centro Universitario de Investigaciones Bibliotecológicas, 1990.
- Gorman, M., "An administrator's view: factors influencing organized subject access", *Technicalities*, v.15, no.10, 1995, p.1.
- Grady, Agnes M. "Online maintenance features of authority files: survey of vendors and in-house systems", *Information Technology and Libraries*, March 1988, pp.51-55.
- Graham, Peter S., "Quality in cataloging: making distinctions", *Journal of Academic Librarianship*, v.16, no.4, 1990, pp.213-218.

- Gredley, Ellen y Alan Hopkinson, *Exchanging bibliographic data: MARC and other international formats*. Ottawa, Canadian Library Association, 1990.
- Guy, Robin Frederick, "Evolution of automation in a national library: the experience of the National Library of Scotland from 1978-1989", *Program*, v.24, no.1, 1990, pp.1-19.
- Halverson, Jacque A., Joni Gomez y Jonathan C. Marner, "Creation and implementation of an automated authority section at the Texas A&M University Library", *Cataloging and Classification Quarterly*, v.15, no.3, 1992, pp.57-68.
- Hays, Kathleen M., "Changing standards of quality control in an automated system", *Technical Services Quarterly*, v.6, no.3-4, 1989, pp.49-55.
- Holley, Robert P. (ed.) *Subject control in online catalogs*. New York, Haworth, 1989.
- International Federation of Library Associations and Institutions. Universal Bibliographic Control and International MARC Programme. *UNIMARC/Authorities: universal format for authorities*. New York, K.G. Sauer, 1991.
- Irwin, Dale, "Local systems and authority control", *Cataloging and Classification Quarterly*, v.16, no.2, 1993, pp.55-69.
- Jamieson, Alexis J., Elizabeth Dolan y Luc Declerck, "Keyword searching vs. authority control in an online catalog", *Journal of Academic Librarianship*, v.12, no.5, November 1986, pp.277-283.
- Jaramillo, George R., "Authority control: is it needed for the 80's?", *Colorado Libraries*, v.11, n.1, Spring 1985, pp.24-28.
- Johnston, Sarah Hager, "Current offerings in automated authority control: a survey of vendors", *Information Technology and Libraries*, September 1989, pp.236-264.
- Kascus, M. (ed.). *Outsourcing cataloging, authority work and physical processing: a checklist of considerations*. Chicago, American Library Association, 1995.
- Kent, Allen y Harold Lancour (eds.) *Encyclopedia of Library and Information Science*. New York, M. Dekker, 1969, v.2.
- Kidder, Ione Marion (comp.) *Encabezamientos de materia, castellano-inglés, inglés-castellano, basados en la lista de los encabezamientos de materia utilizados en el catálogo de la Biblioteca Nacional, Caracas, Venezuela, con el equivalente en inglés tomado de 'Subject headings used in the dictionary catalogs of the Library of Congress', edición de 1943*. México, Biblioteca Benjamin Franklin, 1947.

- Kranz, Jack, "Writing technical specifications for database authorization", *Library Resources and Technical Services*, v.33, no.4, October 1989, pp.358-365.
- Krieger, Michael T., "Subject authority control on the Dynix system", *Cataloging and Classification Quarterly*, v.12, no.1, 1990, pp.103-115.
- Laucaster, F. W. "The evaluation of library services: a concise review of the existing literature", *Investigación bibliotecológica*, v.9, no.18, 1995, pp.25-37.
- LeBlanc, James D., "Cataloging in the 1990s: managing the crisis (mentality)", *Library Resources and Technical Services*, v.37, no.4, 1993, pp.423-433.
- Library Technologies, Inc. *Authority control services*. Abington, Library Technologies, 1992.
- Ludy, Lorene E. "OSU libraries use of Library of Congress subject authorities file" *Information Technology and Libraries*, v.4, no.2, 1985, pp.155-160.
- _____ y Sally A. Rogers, "Authority control in the online environment", *Information Technology and Libraries*, v.3, no.3, September 1984, pp.262-266.
- Maccaferri, James Tilio, "Managing authority control in a retrospective conversion project", *Cataloging and Classification Quarterly*, v.14, no.3-4, 1992, pp.145-167.
- Marnier, J.C., "Measuring the success of keyword search strategy in an online catalog", *Technical Services Quarterly*, v.11, no.2, 1993, pp.1-11.
- Martínez Arellano, Felipe F. et al., "Catálogos de autoridad de la Dirección General de Bibliotecas", ponencia presentada en el *Seminario Anual ABIESI 92*. México, Asociación de Bibliotecarios de Instituciones de Educación Superior e Investigación, 1992.
- Mbaakanyi, D.M.; F.N. Ubogu y E. Lumande, "Subject authority control in a computerised system: use of CD-MARC subjects in an academic library", *Electronic Library*, v.11, no.4/5, 1993, pp.311-317.
- McDonald, David R., "Data dictionaries, authority control, and online catalogs: a new perspective", *Journal of Academic Librarianship*, v.11, no.4, September 1985, pp.219-222.
- McGrath, Ellen, "Developing priorities for authority work at the University at Buffalo libraries", *Cataloging and Classification Quarterly*, v.16, no.4, 1993, pp.5-23.
- Moreno Jiménez, Pilar María y Alvaro Quijano Solís, "Los catálogos de autoridad de materia en el contexto automatizado", ponencia presentada al V

- Coloquio sobre Automatización de Bibliotecas*, Colima, Universidad de Colima, 1991.
- _____, "El control de autoridad y los sistemas automatizados para bibliotecas: criterios de evaluación", ponencia presentada a las *XXIII Jornadas Mexicanas de Biblioteconomía*, Mérida, Universidad Autónoma de Yucatán, 1992.
- _____, "Estudio de diagnóstico de errores en el catálogo automatizado de la biblioteca de El Colegio de México", ponencia presentada al *VI Coloquio sobre Automatización de Bibliotecas*, Colima, Universidad de Colima, 1993.
- _____, "Lenguajes de indización libres o controlados: la experiencia de la Biblioteca Daniel Cosío Villegas de El Colegio de México", ponencia presentada en el *XII Coloquio de Investigación Bibliotecológica*, México, Universidad Nacional Autónoma de México, Centro Universitario de Investigaciones Bibliotecológicas, 1994.
- _____ y Shirley Ainsworth, *Informe de diagnóstico de errores en el catálogo automatizado de la BDCV*. Documento interno de trabajo. México, El Colegio de México, Biblioteca Daniel Cosío Villegas, 1993.
- Nye, Julie Blume, "User interaction with the authority structure of the online catalog: results of a survey", *Information Technology and Libraries*, September 1988, pp.313-316.
- Oddy, Pat, "Authority control in the local, national and international environment", En: *Standars for the international exchange of bibliographic information: papers presented at a course held at the School of Library, Archive and Information Studies, University College, London, 3-18 August 1990*, ed. by I. C. McIlwaine. London, Library Association, 1991. pp.66-72.
- O'Neill, Edward T. y Rao Aluri. *A method for correcting typographical errors in subject headings in OCLC records*. Dublin, OCLC, 1980.
- _____, W. Michael Oskins y Kerre A. Kammerer, "OCLC authority control", *Annual Review of OCLC research July 1992-June 1993*, pp.30-37.
- _____ y Diane Vizine-Goetz, "Computer generation of a subject authority file", *Proceedings of the ASIS Annual Meeting*, v.19, 1982, pp.220-223.
- _____, "Quality control in online databases" *Annual Review of Information Science and Technology*, v.23, 1988, pp.125-156.
- Palumbo, K.M., "Experiencies in performing authority work with a NOTIS catalog", *Colorado Libraries*, v.20, no.1, 1994, pp.51-54.

- Park, Amey L., "Automated authority control: making the transition", *Special Libraries*, Spring 1992, pp.75-85.
- Piternick, Anne B., "Vocabularies for online subject searching", *Encyclopedia of Library and Information Science*, v.45, supplement 10, 1990, pp.399-420.
- Pontificia Universidad Católica de Chile. Dirección del Sistema de Bibliotecas. *Proyecto "Implementación de una base de datos de autoridades para un sistema de catalogación automatizada, Chile": informe final*. Santiago de Chile, 1986.
- Pontificia Universidad Católica de Chile. Sistema de Bibliotecas. Departamento de Catalogación. Tratamiento de las autoridades. Documento interno. Santiago de Chile, 1996.
- Preece, B.G. y B. Henigman, "Shared authority control: governance and training", *Technical Services Quarterly*, v.11, no.3, 1994, pp.19-31.
- Quijano Solís, Alvaro y Clotilde Tejeda Rodríguez, "La organización del sistema integral de automatización de la Biblioteca Daniel Cosío Villegas", Ponencia presentada a las *XXII Jornadas Mexicanas de Biblioteconomía*, 1991. México, Asociación Mexicana de Bibliotecarios, 1991.
- _____. *Proyecto de automatización de la Biblioteca: informe de avance al 18 de octubre 1991*. Documento interno. México, El Colegio de México, Biblioteca Daniel Cosío Villegas, 1991.
- Quiroga Carmona, Luz Marina. *Una base de datos para apoyo al proceso de catalogación bibliográfica en línea*. Tesis para obtener el grado de Maestra en Ciencias de la Computación. México, Universidad Nacional Autónoma de México, 1984.
- Roe, K., "Enhanced authority control: is it time?", *Archivaria*, no.35, 1993, pp.119-129.
- Rogers, Sally A. "Automated control of subject headings at the OSU libraries", *Information Technology and Libraries*, March 1989, pp.79-82.
- Smith, P., "Book vendor-supplied cataloging: impacts of technical services", *Colorado Libraries*, v.20, no.3, 1994, pp.14-16.
- Stehelmar, S., "Vocabulary control and the humanities: a case study of the MLA International Bibliography", *Reference Librarian*, no.47, 1994, pp.61-78.
- Taylor, Arlene G., "Authority files in online catalogs: an investigation of their value", *Cataloging and Classification Quarterly*, v.4, n.3, Spring 1984, pp.1-17.
- _____, Margaret F. Maxwell y Carolyn O. Frost, "Network and vendor authority systems", *Library Resources and Technical Services*, v.29, n.2, April-June 1985, pp.195-205.

- Tejeda Rodríguez, Clotilde y Alvaro Quijano Solís, "Conversión retrospectiva: piedra angular de la automatización de la Biblioteca", Ponencia presentada al *Seminario Anual ABIESI 1991*. México, Asociación de Bibliotecarios de Instituciones de Educación Superior e Investigación, 1991.
- Thomas, Catherine M., "Authority control in manual versus online catalogs: an examination of 'see' references", *Information Technology and Libraries*, v.3, December 1984, pp.393-398.
- Thomas, Sarah E. y Jennifer A. Younger, "Cooperative cataloging: a vision for the future", *Cataloging and Classification Quarterly*, v.17, no.3/4, 1993, pp.237-257.
- Tillet, Barabara B., (ed.) *Authority control in the online environment: considerations and practices*. New York, Haworth, 1989.
- Truitt, Marc. "USMARC to UNIMARC/Authorities: a qualitative evaluation of USMARC data elements", *Library Resources and Technical Services*, v.36, no.1, 1992, pp.37-58
- USMARC format for authority data*. Cumulated volume (base text and updates 1 and 2). Washington, Library of Congress Network Development and MARC Standards Office, 1989.
- Update 3. Washington, Library of Congress Network Development and MARC Standards Office, 1990.
- Update 4. Washington, Library of Congress Network Development and MARC Standards Office, 1991.
- Ed. 1993. Washington, Library of Congress Network Development and MARC Standards Office, 1993.
- Ed. 1993. Update 1. Washington, Library of Congress Network Development and MARC Standards Office, 1995.
- Valentine, Phyllis A., "Increasing production in cataloging while decreasing cost", *Technicalities*, v.4, n.7, July 1984, pp.10-13.
- Van Pulis, Noelle y Lorene E. Ludy, "Subject searching in an online catalog with authority control", *College and Research Libraries*, v.49, no.6, November 1988, pp.523-533.
- Wajenberg, Arnold S., "Authority work, authority records, and authority files", en: Michael Gorman *et al. Technical services today and tomorrow*. Englewood, Libraries Unlimited, 1990.
- Watson, Mark R. y Arlene G. Taylor, "Implications of current references structures for authority work in online environments", *Information Technology and Libraries*, v.6, no.1, March 1987, pp.10-19.

- White, Marilyn y Eileen G. Abels, "Measuring service quality in special libraries: lessons from service marketing", *Special Libraries*, v.86, no.1, 1995, pp.36-45.
- Wilkes, Adeline y Antoinette Nelson, "Subject searching in two online catalogs: authority control vs. non-authority control", *Cataloging and Classification Quarterly*, v.20, no.4, 1995, pp.57-79.
- Winke, R. Conrad, "Discarding the main entry in an online cataloging environment", *Cataloging and Classification Quarterly*, v.16, no.1, 1993, pp.53-70.
- Working Group on an International Authority System. *Guidelines for authority control and reference entries*. London, International Federation of Library Associations, International Programme for Universal Bibliographic Control, 1984.
- Younger, Jennifer A., "After Cutter: authority control in the twenty-first century", *Library Resources and Technical Services*, v.39, no.2, 1995, pp.133-141.

ANEXO:

Formato MARCOLMEX para registros de autoridad

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
001 (NR)	No tiene	No tiene	Núm. de control (generado automáticamente por el sistema)
003 (NR)	No tiene	No tiene	Identificador del número de control. Valor =CM
005 (NR)	No tiene	No tiene	Fecha y hora de la última transacción (generado automáticamente por el sistema)
008 (NR) /00-05	No tiene	No tiene	Datos de longitud fija Fecha entrada al archivo (generado automáticamente por el sistema)

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
/06			División geográfica directa o indirecta Valores: b - no subd. geográficamente d - directa i - indirecta n - no aplicable
/07			Esquema de romanización Valores: a - norma internacional b - nacional e - local f - de origen desconocido g - romanización convencional o forma convencional del nombre en español n - no aplicable
/08			Posición no definida
/09			Tipo de registro Valores: a - encabezamiento establecido

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
/09			f - encabez. establecido y subdivisión b - referencia sin trazados c - referencia con trazados g - referencia y subdivisión d - subdivisión e - nódulo temático
/10			Reglas de catalogación descriptiva Valores: b - RCAA1 c - RCAA2 z - Políticas de la BDCV n - No aplicable
/11			Lista de encabezamientos de materia o tesaurus Valores: z-Kidder/COLMEX n - No aplicable
/12			Tipo de serie Valores: a - serie monogr.

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
/12			b - obra en varios vols. c - serie-frase z - otra n - No aplicable
/13			Serie numerada o no numerada Valores: a - numerada b - no numerada c - num. var. n - no aplicable
/14			Uso del encabezamiento - entrada principal o secundaria Valores: a - apropiado b - no apropiado
/15			Uso del encabezamiento - encabezamiento de materia Valores: a - apropiado b - no apropiado
/16			Uso del encabezamiento - entrada de serie

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
/16			Valores: a - apropiado b - no apropiado
/17			Tipo de subdivisión de materia Valores: a - materia b - forma c - cronológica d - geográfica e - lengua n - no aplicable
/18-27			Caracteres no definidos
/28			Tipo de organismo gubernamental Valores: a - componente autónomo o semi-autónomo de Malasia b - no gubernam. c - multilocal i - internacional intergubernam. f - nacional /federal l - local s - estatal,

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
/28			provincial, territorial, dependiente, etc. o - gubernamental u - desconocido z - otra
/29			Evaluación de referencias Valores: a - trazados consistentes con el encabezamiento b - trazados no necesariamente consistentes n - no aplicable
/30			Posición no definida
/31			Actualización del registro Valores: a - en uso b - registro actualizándose
/32			Nombre personal indiferenciable Valores: a - diferenciable b - indiferenc. n - no aplicable

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
/33			Estado del registro Valores: a - encabez. completamente establecido b - memorandum c - provisional d - preliminar n - no aplicable
/34-37			Caracteres no definidos
/38			Registro modificado Valores: b - no modificado s - reducido x - faltan caracteres
/39			Fuente de catalogación Valor fijo: d
014 (R)	b b	a (NR)	Registro bibliográfico ligado Número de control del reg. bibliográfico relacionado Liga
		6 (NR)	

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
020 (R)	b b	a (NR) 6 (NR)	ISBN ISBN Liga
022 (R)	b b	a (NR) 6 (NR)	ISSN ISSN Liga
035 (R)	b b	a (NR) z (R) 6 (NR)	Número de control Núm. control Núm. inválido o cancelado Liga
040 (NR)	b b	a (NR) b c (NR) d (NR)	Fuente de catalogación Agencia de catalogación original spa Agencia que transcribe Agencia que modifica
042 (NR)	b b	a (R)	Código de autenticación
043 (R)	b b	a (NR)	Código de área geográfica

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
045 (NR)	Primero: Tipo de período de tiempo en los subcampos b ó c Valores: b - no hay subc. b ó c 0 - una sola fecha/tiempo 1 - múltiples fecha/tiempo solos 2 - cadenas de fechas/tiempos Segundo: b	a (R) b (R) c (R)	Período de tiempo Código cronológico Fechas desde 9999 a.C. hasta toda la era cristiana Fechas anteriores a 9999 a.C.
050 (R)	b 0	a (NR) b (NR) d (NR) 6 (NR)	Signatura LC Clasificación Cutter Vols. o fechas Liga
052 (R)	b b		Código de clasificación geográfica

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
052		a (NR)	Código de clasificación geográfica
		b (R)	Subcódigo de clasificación geográfica
053 (R)	b b	a (NR)	Clasificación LC
066 (NR)	b b	a (NR) b (NR) c (R)	Conjuntos de caracteres presentes No-ASCII G0 No-ANSEL G1 Identificación de un grupo de caracteres gráficos especiales
073 (NR)	b b	a (R) 6 (NR)	Uso de subdivisión Liga
082 (R)	Primero: Tipo de edición Valores: 0 - completa 1 - abreviada Segundo: Fuente de la signatura		Signatura Dewey

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
082	Valores: 0 - LC 4 - BDCV	a (NR) b (NR) d (NR) 2 (NR) 6 (NR)	Clasificación Cutter Vols. o fechas No. edición Liga
083 (R)	Primero: Tipo de edición Valores: 0 - completa 1 - abreviada Segundo: b	a (NR) 2 (NR) 6 (NR)	Clasificación Dewey No. edición Liga
100 (NR)	0,1,2,3 b	a, b, c, d k, f, g, l n, p, s x, y, z	Encabezamiento nombre personal Tipo de elemento de entrada (Igual a formato bibliográfico) (Igual a formato bibliográfico) (Igual a formato bibliográfico)
110 (NR)	0,1,2 b	a, b, c, d e	Encabezamiento nombre corporativo Tipo de elemento de entrada (Igual a formato bibliográfico)

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
110		g, k, n t, f, l, p s x, y, z	(Igual a formato bibliográfico) (Igual a formato bibliográfico) (Igual a formato bibliográfico)
111 (NR)	0, 1, 2 b	a, q, d, c e, g, b t, f, k, l p, s x, y, z	Encabezamiento nombre de reunión Tipo de elemento de entrada (Igual a formato bibliográfico) (Igual a formato bibliográfico) (Igual a formato bibliográfico)
130 (NR)	b 0-9	a, n, p, l f, k, s, d h, g, t x, y, z	Encabezamiento título uniforme Caracteres que no alfabetizan (Igual a formato bibliográfico) (Igual a formato bibliográfico)

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
150 (NR)	b b	a, b x, y, z	Encabezamiento de materia general (Igual a formato bibliográfico)
151 (NR)	b b	a x, y, z	Encabezamiento de materia nombre geográfico (Igual a formato bibliográfico)
180 (NR)	b b	x, y, z,	Subdivisión general (igual a 150)
181 (NR)	b b	x, y, z,	Subdivisión geográfica (igual a 150)
182 (NR)	b b	x, y, z,	Subdivisión cronológica (igual a 150)
260 (R)			Referencia compleja de véase - materia

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
260	b b	a (R) i (R)	Encabezamiento Texto explicativo
360 (R)			Referencia compleja de véase también - materia
	b b	a (R) i (R)	Encabezamiento Texto explicativo
400 (R)			Véase de - nombre personal
	Igual a 100	Igual a 100 i (NR)	Frase de instrucción de referencia Subcampo de control Liga
		w (NR) 6 (NR)	
410 (R)			Véase de - nombre corporativo
	Igual a 110	Igual a 110 i (NR)	Frase de instrucción de referencia Subcampo de control Liga
		w (NR) 6 (NR)	

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
411 (R)	Igual a 111	Igual a 111 ₃ i (NR) w (NR) 6 (NR)	Véase de - conferencia o reunión Frase de instrucción de referencia Subcampo de control Liga
430 (R)	Igual a 130	Igual a 130 i (NR) w (NR) 6 (NR)	Véase de - título uniforme Frase de instrucción de referencia Subcampo de control Liga
450 (R)	Igual a 150	Igual a 150 i (NR) w (NR) 6 (NR)	Véase de - materia Frase de instrucción de referencia Subcampo de control Liga

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
451 (R)	Igual a 151	Igual a 151 i (NR) w (NR) 6 (NR)	Véase de - nombre geográfico Frase de instrucción de referencia Subcampo de control Liga
480 (R)	Igual a 180	Igual a 180 i (NR) w (NR) 6 (NR)	Véase de - Subdivisión general Frase de instrucción de referencia Subcampo de control Liga
481 (R)	Igual a 181	Igual a 181	Véase de - subdivisión geográfica

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
481		i (NR)	Frase de instrucción de referencia
		w (NR)	Subcampo de control
		6 (NR)	Liga
482 (R)	Iguales a 182	Iguales a 182	Véase de - subdivisión cronológica
		i (NR)	Frase de instrucción de referencia
		w (NR)	Subcampo de control
500 (R)	Iguales a 100	Iguales a 100	Véase también de -nombre personal
		i (NR)	Frase de instrucción de referencia
		w (NR)	Subcampo de control
		6 (NR)	Liga

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
510 (R)	Iguales a 110	Iguales a 110 i (NR) w (NR) 6 (NR)	Véase también de - nombre corporativo Frase de instrucción de referencia Subcampo de control Liga
511 (R)	Iguales a 111	Iguales a 111 i (NR) w (NR) 6 (NR)	Véase también de - conferencia o remiún Frase de instrucción de referencia Subcampo de control Liga
530 (R)	Iguales a 130	Iguales a 130 i (NR)	Véase también de - título uniforme Frase de instrucción de referencia

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
530		w (NR)	Subcampo de control
		6 (NR)	Liga
550 (R)	Igual a 150	Igual a 150 i (NR)	Véase también de - materia
		w (NR)	Frase de instrucción de referencia
		6 (NR)	Subcampo de control Liga
551 (R)	Igual a 151	Igual a 151 i (NR)	Véase también de -nombre geográfico
		w (NR)	Frase de instrucción de referencia
		6 (NR)	Subcampo de control Liga
580 (R)	Igual a 180	Igual a 180	Véase también de - subdivisión general

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
580		i (NR)	Frase de instrucción de referencia
		w (NR)	Subcampo de control
		6 (NR)	Liga
581 (R)	Igual a 181	Igual a 181 i (NR)	Véase también de - subdivisión geográfica
		w (NR)	Frase de instrucción de referencia
		6 (NR)	Subcampo de control
			Liga
582 (R)	Igual a 182	Igual a 182 i (NR)	Véase también de - subdivisión cronológica
		w (NR)	Frase de instrucción de referencia
		6 (NR)	Subcampo de control
			Liga

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
640 (R)	Primero: Forma de la nota Valores: 0 - estándar 1 - libre Segundo: b	a (NR) z (NR) 6 (NR)	Fechas de publicación y/o vols. de serie Fechas o volúmenes Fuente de la información Liga
641 (R)	b b	a (NR) z (NR) 6 (NR)	Peculiaridades en la numeración de series Nota Fuente de información Liga
642 (R)	b b	a (NR) d (NR) 6 (NR)	Ejemplo serie numerada Ejemplo Vols./fechas Liga
643 (R)	b b	a (R) b (R) d (NR) 6 (NR)	Lugar y editor de la serie Lugar Editor Vols./fechas Liga

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
644 (R)	b b	a (NR) b (NR) d (NR) 6 (NR)	Política de análisis de series Política general Excepciones Vols./fechas Liga
646 (R)	b b	a (NR) d (NR) 6 (NR)	Clasificación de series Política de clasificación de series Valores: s - vols. separados c - vols. en colección m - vols. en subserie Vols./fechas Liga
663 (NR)	b b	a (R) b (R) t (R) 6 (NR)	Referencia compleja de véase también - nombres Texto explicativo Encabezamiento Título Liga

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
664 (NR)	b b		Referencia compleja de véase - nombres
		a (R)	Texto explicativo
		b (R)	Encabezamiento
		t (R)	Título
		6 (NR)	Liga
665 (NR)	b b	a (R)	Referencia de historia
		6 (NR)	Liga
666 (NR)	b b		Referencia explicativa general - nombres
		a (R)	Referencia
		6 (NR)	Liga
667 (R)	b b	a (NR)	Nota general interna
		6 (NR)	Liga
670 (R)	b b		Fuente de datos encontrados
		a (NR)	Fuente
		b (NR)	Información
		6 (NR)	Liga
675 (NR)	b b		Fuente de datos no encontrados
		a (R)	Fuente
		6 (NR)	Liga

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
678 (NR)	b b	a (R) 6 (NR)	Epítome Liga
680 (R)	b b	a (R) i (R) 6 (NR)	Nota general pública Encabezamiento Texto explicativo Liga
681 (R)	b b	a (R) i (R) 6 (NR)	Nota de ejemplo - temas Encabezamiento Texto explicativo Liga
682 (NR)	b b	i (R) a (R) 6 (NR)	Información sobre encabez. borrados Texto explicativo Nuevo encabez. Liga
688 (R)	b b	a (NR) 6 (NR)	Nota de aplicación histórica Nota Liga
700 (R)	Primero: 1,2,3 Segundo: 0,1,2,3,4,5,6,7	Igual a 100 v (R)	Encabezamiento establecido entrada relacionada - nombre personal Número de control del registro

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
700		w (NR)	Subcampo de control
		2 (NR)	Fuente del encabezamiento o término
		6 (NR)	Liga
		8 (NR)	Número relacionado
710 (R)	Primero: 0,1,2 Segundo: 0,1,2,3,4,5,6,7		Encabezamiento establecido entrada relacionada -nombre corporativo
		Iguales a 110	
		v (R)	Número de control del registro
		w (NR)	Subcampo de control
		2 (NR)	Fuente del encabezamiento o término
		6 (NR)	Liga
		8 (NR)	Número relacionado
711 (R)	Primero: 0,1,2 Segundo: 0,1,2,3,4,5,6,7		Encabezamiento establecido entrada relacionada -nombre conferencia
		Iguales a 111	

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
711		v (R)	Número de control del registro
		w (NR)	Subcampo de control
		2 (NR)	Fuente del encabezamiento o término
		6 (NR)	Liga
		8 (NR)	Número relacionado
730 (R)	Primero: b Segundo: 0,1,2,3,4,5,6,7	Igual	Encabezamiento establecido
		a	entrada relacionada
		130	- título uniforme
		v (R)	Número de control del registro
		w (NR)	Subcampo de control
		2 (NR)	Fuente del encabezamiento o término
		6 (NR)	Liga
8 (NR)	Número relacionado		
750 (R)	Primero: b Segundo: 0,1,2,3,4,5,6,7		Encabezamiento establecido
			entrada relacionada
			- materia general

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
750		Igual a 150 v (R) w (NR) 2 (NR) 6 (NR) 8 (NR)	Número de control del registro Subcampo de control Fuente del encabezamiento o término Liga Número relacionado
751 (R)	Primero: b Segundo: 0,1,2,3,4,5,6,7	Igual a 151 v (R) w (NR) 2 (NR) 6 (NR) 8 (NR)	Encabezamiento establecido entrada relacionada -materia geográfica Número de control del registro Subcampo de control Fuente del encabezamiento o término Liga Número relacionado

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
780 (R)	Primero: b Segundo: 0,1,2,3,4,5,6,7	Igual a 180 v (R) w (NR) 2 (NR) 6 (NR) 8 (NR)	Encabezamiento de subdivisión general entrada relacionada Número de control del registro Subcampo de control Fuente del encabezamiento o término Liga Número relacionado
781 (R)	Primero: b Segundo: 0,1,2,3,4,5,6,7	Igual a 181 v (R) w (NR) 2 (NR) 6 (NR) 8 (NR)	Encabezamiento de subdivisión geográfica entrada relacionada Número de control del registro Subcampo de control Fuente del encabezamiento o término Liga Número relacionado

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
782 (R)	Primero: b Segundo: 0,1,2,3,4,5,6,7	Igual a 182 v (R) w (NR) 2 (NR) 6 (NR) 8 (NR)	Encabezamiento de subdivisión cronológica entrada relacionada Número de control del registro Subcampo de control Fuente del encabezamiento o término Liga Número relacionado
788 (NR)	Primero: b Segundo: 0,1,2,3,4,5,6,7	a (R) i (R) 2 (NR) 6 (NR) 8 (NR)	Entrada relacionada compleja Encabezamiento al que se refiere Texto explicativo Fuente del encabezamiento o término Liga Número relacionado
880 (R)	Primero: Igual al campo asociado		Representación de grafía alternativa

EtiquetaMARC	Valor de indicadores	Código subcampo	Descripción del campo, indicador o subcampo
880	Segundo: Igual al campo asociado	a-z	Igual al campo asociado
		2, 8	Igual al campo asociado
		6 (NR)	Liga
902 (NR)	b b	a (NR)	Catalogador que dio de alta el encabezamiento