

V COLOQUIO SOBRE AUTOMATIZACION DE BIBLIOTECAS

UNIVERSIDAD DE COLIMA

1991

"LOS CATALOGOS DE AUTORIDAD DE MATERIA EN EL CONTEXTO AUTOMATIZADO"



Pilar María Moreno
Y
Alvaro Quijano Solís

Biblioteca Daniel Cosío Villegas
El Colegio de México

El Colegio de México, A.C.
Camino al Ajusco No. 20 Col. Pedregal de Santa Teresa
Delegación Tlalpan, C.P. 10740
México, D.F.

INFOBILA

Donación

17 Oct 2011

2933

LOS CATALOGOS DE AUTORIDAD DE MATERIA

EN EL CONTEXTO AUTOMATIZADO

INTRODUCCIÓN

Este documento tiene por objeto presentar las modalidades que reviste la construcción de un catálogo automatizado de autoridades de materia en la Biblioteca Daniel Cosío Villegas de El Colegio de México. El proyecto surge en el entorno del proceso de automatización de la biblioteca, descrito en un trabajo anterior,¹ y con apoyo en la conversión de los registros manuales del catálogo público a registros legibles por computadora, en formato "Machine Readable Cataloging" (MARC).

LA AUTOMATIZACIÓN Y EL CATÁLOGO DE AUTORIDAD

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.²

Así pues, la estructuración del catálogo de autoridad es una fase previa en el proceso del control de autoridad. Un catálogo de

¹ Alvaro Quijano Solís, y Clotilde Tejeda Rodríguez. "La organización del sistema integral de automatización de la Biblioteca Daniel Cosío Villegas", ponencia presentada en las XXII Jornadas Mexicanas de Biblioteconomía, 1991, 16 p.

² Arlene G. Taylor, "Authority files in online catalogs: an investigation of their value", Cataloging and Classification Quarterly, vol. 4, No. 3, Spring 1984, p. 2.

autoridad deberá contener todas las entradas posibles para los encabezamientos, ya sea de autores, 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:³

- Creación de encabezamientos (puntos de acceso) uniformes que aseguren la colocación de todas las obras relacionadas entre sí bajo encabezamientos uniformes.
- 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.

El control de autoridad automatizado consiste en el uso de una computadora para manejar grandes bloques en el proceso de control de autoridad.⁴

En general, los procesos automatizados de control de autoridad que se han puesto en práctica en el contexto norteamericano caen dentro de las siguientes categorías:⁵

- Procesamiento electrónico de catálogos de autoridad.

³ Jean Dickson and Patricia Zadner, "Authority control and the authority file: a functional evaluation of LCNAE on RLIN", Cataloging and Classification Quarterly, vol. 9, No. 3, 1989, p. 57-73.

⁴ Ibid.

⁵ Arlene G. Taylor, Margaret F. Maxwell, and Carolyn O. Frost, "Network and vendor authority systems", Library Resources and Technical Services, vol. 29, April-June 1985, p. 195-205.

- Catálogos de autoridad en línea pero sin conexión con los catálogos bibliográficos.
- Sistemas de control de autoridad integrados a un sistema de bases de datos bibliográficos.

Este último procedimiento permite establecer la conexión necesaria entre el catálogo de autoridad y los registros bibliográficos. Consiste básicamente en crear un catálogo de autoridad aparte compuesto por encabezamientos tomados del catálogo bibliográfico, y donde ambos catálogos se relacionan entre sí de tal forma que cualquier cambio en el catálogo de autoridad inicia automáticamente un cambio en el catálogo bibliográfico y viceversa. En general, este sistema presenta las siguientes ventajas:

- Validación automática de los encabezamientos.
- Detección de errores.
- Despliegue de los términos relacionados.
- Reducción de los requerimientos de almacenamiento de datos,
- Capacidad para efectuar cambios globales.⁶

Es una verdad generalmente aceptada que el usuario común rara vez piensa como bibliotecario. 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 lineal del catálogo manual, representan una oportunidad importante para conciliar los intereses de usuarios y bibliotecarios. Esta conciliación deberá darse a partir del acercamiento entre los

⁶ David R. McDonald, "Data dictionaries, authority control, and online catalogs: a new perspective", Journal of Academic Librarianship, vol. 11, September 1985, p. 221.

longitud fija (o de control) y los campos de longitud variable. Cabe señalar que los dos primeros marcan la pauta que facilita 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 que la posición del carácter 6 nos va a indicar el tipo de registro con un código "2" que significa que se trata de un registro de autoridad, permitiéndonos así diferenciarlo de un registro bibliográfico.

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

Los campos se identifican cada uno 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 (véase anexo 1). Los bloques 1xx, 4xx y 5xx guardan además ciertos paralelos en la designación de contenidos, como puede apreciarse por los dos últimos caracteres numéricos de sus etiquetas (véase anexo 2).

Aquí es importante hacer notar que a pesar de las semejanzas en la estructura de los campos de ambos formatos, bibliográfico y de autoridades, una etiqueta de uno no siempre designa el mismo contenido para ambos. Así, los campos lxx contienen la misma información en uno y otro, sin embargo, por ejemplo el campo 500 será una nota general en un registro bibliográfico, mientras que en un registro de autoridad contendrá una referencia de "véase también de".

Por último, señalaremos que entre las ventajas de la adopción de un formato MARC para control de autoridades se encuentran las siguientes:

- Permite establecer un vínculo o conexión entre los registros bibliográficos y los registros de autoridad correspondientes en forma automatizada.
- 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 da la posibilidad de establecer interdependencias entre los datos codificados para diferentes tipos de registros, constituyendo un instrumento más 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.

- La normalización de las prácticas para el establecimiento de encabezamientos y estructura de autoridades.
- La promoción de responsabilidades en la creación y difusión de registros de autoridad, y
- La creación de una infraestructura que posibilite el intercambio eficiente de datos de autoridad a nivel internacional.⁷

La IFLA, por su parte, ha promovido la creación y adopción de un formato de intercambio internacional de datos de autoridad, llamado "UNIMARC-Autoridades", semejante al formato MARC para intercambio de datos bibliográficos, y próximo a publicarse.

Asimismo, la Biblioteca del Congreso de los Estados Unidos ha venido trabajando desde 1976 un formato MARC de autoridades para su uso a nivel nacional, publicando una primera edición en esa fecha y tres actualizaciones sucesivas posteriormente, siendo la última de marzo de 1990.

El formato MARC para intercambio de datos de autoridad está diseñado para contener la información relativa a las formas establecidas de nombres y materias usadas como puntos de acceso en registros bibliográficos, las formas de esos nombres y materias que deben usarse como referencias a las formas establecidas, y las interrelaciones entre esas formas.

Este formato presenta una estructura semejante a la del formato MARC para datos bibliográficos. Así, consta básicamente de cuatro elementos: la guía (o líder), el directorio, los campos de

⁷ Tom Delsey. "Authority control in an international context", en Tillet, Barbara B. (ed.), Authority Control in the Online Environment: Considerations and Practices, p. 13-28.

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 que el lenguaje de sus preguntas esté o no ortodoxamente formalizado en términos bibliotecarios.

El reto para la biblioteca será lograr este objetivo sujeto a dos grandes restricciones: el tamaño de los archivos y el costo de mantener el control de autoridad.

LOS CATÁLOGOS DE AUTORIDAD Y EL FORMATO MARC

En el nivel internacional, la IFLA, a través de sus programas de Control Bibliográfico Universal (CBU) ha definido tres principios que deben seguirse en la descripción bibliográfica y el control de autoridades:

- Debe existir un canon mutuamente aceptado de principios, normas y prácticas que rijan la formulación y estructura de los datos catalográficos.
- Cada agencia bibliográfica nacional debe ejercer sus responsabilidades de forma que estén incluidos consistentemente los patrones comúnmente acordados.
- Debe existir una infraestructura que posibilite el intercambio eficiente de datos entre las diferentes agencias bibliográficas nacionales.

Estos principios implican la necesidad de implementar políticas nacionales para:

Algunos ejemplos de registros dentro del formato MARC se presentan en el anexo 3.

PROCEDIMIENTO PROPUESTO PARA GENERAR UN CATALOGO DE AUTORIDAD DE MATERIA

El proyecto de automatización de la Biblioteca comenzó en 1982 con el diseño de un sistema para controlar las adquisiciones por compra. En 1984 se inició la definición de un formato, compatible con el MARC, lo que permitió empezar a capturar las obras que ingresaron a la Biblioteca a partir de 1986.

En 1990, con un apoyo sustancial del gobierno federal, se continuó el proyecto de automatización de la Biblioteca, con la adquisición de un manejador de bases de datos denominado STAR, que se instaló en una computadora SUN, modelo SPARC Server 470 que, hasta la fecha, cuenta con 32 puertos, 32 MB en RAM, y 2.8 GB en disco duro del tipo IPI.

A partir de septiembre de 1990, la Biblioteca inició la captura de los registros anteriores a 1986.³ En una primera etapa se ingresaron los registros de obras adquiridas entre 1984 y 1986. En abril de este año, con un equipo de 18 bibliotecólogos, se inició la codificación de la colección, la que se concluyó a principios de septiembre. A partir de ese momento, los mismos

³ Clotilde Tejeda Rodríguez, y Alvaro Quijano Solís. "Conversión retrospectiva: piedra angular de la automatización de la biblioteca", ponencia presentada en el Seminario Anual ABIESI, 1991, 10 p.

bibliotecólogos se han integrado al proceso de captura, que deberá terminarse en abril de 1992.

Actualmente, la Biblioteca ha ingresado a la computadora alrededor de 150 mil registros, de los que 85 mil constituyen el archivo de prueba para el diseño del catálogo público en el manejador STAR, que deberá estar listo a principios de noviembre.

El primer objetivo que se propuso la biblioteca al abordar la problemática de los catálogos de autoridad fue el de definir un formato MARC de autoridades que proporcionara el marco general adecuado para los diferentes tipos de catálogos. Teniendo en cuenta las ventajas que presentan los sistemas de control de autoridades integrados a un sistema de bases de datos bibliográficos, y que éstos pueden operar con MARC, optamos por un procedimiento de este tipo. El formato que se definió se basó fundamentalmente en el correspondiente "USMARC Format for Authority Data", incluyendo todas las actualizaciones hasta marzo de 1990. Las modificaciones que se hicieron estuvieron encaminadas a adaptar esta herramienta al uso y políticas catalográficas de nuestra biblioteca, tomando en cuenta especialmente la definición del formato llamado MARCOLMEX, que se usa en la actualidad para la codificación y procesamiento de los registros bibliográficos y, obviamente, excluyendo los campos destinados a usos propios del sistema norteamericano.

En una siguiente etapa nos planteamos la depuración del catálogo de materias y la creación del catálogo de autoridad correspondiente. Para ello, hemos iniciado la prueba de un algoritmo que permita a la máquina detectar errores de dos tipos en

encabezamientos 690 (encabezamientos de materia o temas) y 691 (encabezamientos de materias geográficos):

1. Inconsistencias en el uso de espacios, signos de puntuación y mayúsculas.
2. Errores tipográficos y de ortografía.

El algoritmo, propuesto por O'Neill y basado en la experiencia de OCLC,⁹ trabaja con frecuencias de uso y peso específico, criterios cuantitativos que permiten discernir entre dos encabezamientos aparentemente iguales.

Los errores incluidos en el inciso 1 pueden ser minimizados al pasar los encabezamientos a mayúsculas compactas, desapareciendo los espacios en blanco, y construyendo una cadena con los primeros 17 caracteres, seguidos de los últimos 17 en orden inverso.

La cadena así construida sirve como base de comparación carácter por carácter, para producir listados basados en la frecuencia, que contienen los encabezamientos aparentemente erróneos, y permite reducir los errores mencionados en el inciso 2.

El archivo final, idealmente, deberá reflejar encabezamientos con frecuencias y pesos relativos altos. Basados en el trabajo citado, se puede afirmar además que aquél reducirá su tamaño aproximadamente a una quinta parte de su tamaño inicial.

Este archivo depurado constituirá la base de un archivo MARC de autoridad de materia, que todavía deberá someterse a una revisión visual apoyada en combinaciones de información, tales como

⁹ Edward T. O'Neill, and Diane Vizine-Goetz. "Computer generation of a subject authority file", Proceedings of the ASIS Annual Meeting, vol. 19, 1982, p. 220-223.

clasificación vs. primer encabezamiento, para detectar errores de forma y de etiquetado. Además, necesitará completarse con la adición de las referencias y notas que relacionen entre sí los diferentes encabezamientos.

Procedimientos de este tipo deberán ser trabajados en profundidad más adelante para construir los catálogos de autoridad de autores, títulos y series.

CONSIDERACIONES FINALES

La construcción del catálogo de autoridad de materias es una tarea cuya dificultad reviste modalidades particulares cuando se trabaja en un ambiente automatizado. Muchas de las dificultades que se presentan podrían allanarse si existiese un catálogo de autoridades emitido a nivel nacional por alguna entidad responsable, acorde con las recomendaciones de la IFLA.

La investigación que se aplica a la solución de problemas que afectan la operación de las bibliotecas, si bien no forma parte de un programa formal de investigación académica, es fundamental para continuar la búsqueda de mejores servicios para los usuarios.

El formato MARC para autoridades es una base para diseñar los catálogos de materias, ya que proporciona una estructura formal lógicamente desarrollada para incluir en ella toda la información que exista sobre cada encabezamiento. Sin embargo, cada biblioteca deberá decidir la profundidad de la información que incluirá en su catálogo de autoridad. En todo caso, el MARC es solamente una guía

para definir la liga entre el catálogo bibliográfico y el de autoridad.

Una gran ventaja al contar con una base de datos bibliográficos en la computadora la representa el poder transferir rápidamente los temas de los registros bibliográficos hacia un archivo que permita efectuar todo tipo de depuraciones de manera automática, con base en algoritmos como el que se describe en este trabajo.

El control de autoridades, a partir de la construcción de su catálogo, se convierte en una herramienta poderosísima para el catalogador y, cuando se realiza en un ambiente automatizado, facilita considerablemente el trabajo de catalogación, clasificación y asignación de temas. Todos estos procesos, al agilizarse, permiten una mayor precisión y eficiencia en la entrega de productos a los usuarios, objetivo último de cualquier biblioteca.

BIBLIOGRAFIA

Autoridades: un formato MARC; especificaciones para las cintas magnéticas que contienen registros de autoridad. Edición preliminar que incluye hasta el addendum 8. México, CONACYT, 1982.

Holley, Robert P. (ed.). Subject control in online catalogs. New York, Haworth, 1989.

Tillett, Barbara B. (ed.). Authority control in the online environment: considerations and practices. New York, Haworth, 1989.

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.

ANEXO 1

IDENTIFICADORES DE FUNCION DE LOS DATOS

- 0XX Información de control, números de identificación, clasificación, etc.
- 1XX Encabezamientos.
- 2XX Referencias complejas de véase.
- 3XX Referencias complejas de véase también.
- 4XX Trazados de véase de.
- 5XX Trazados de véase también de.
- 6XX Tratamientos específicos, notas, etc.

ANEXO 2

IDENTIFICADORES DEL TIPO DE INFORMACION DE LOS DATOS

- X00 Nombres personales.
- X10 Nombres corporativos.
- X11 Nombres de reunión.
- X30 Títulos uniformes.
- X50 Materias.
- X51 Nombres geográficos.

Registro de encabezamiento establecido (temas)

GUIA ***** nzbbb22 ***** nbb4500

001 exb86004405b

005 19860808072450.5

008 860719 inban znnba bnbbb bbbbb bbbab anabb bbbu

040 bb\$a NUC \$C CM \$d CM

150 b0\$a Computadoras analógicas electrónicas

450 b0\$a Máquinas calculadoras analógicas

550 b0\$a Computadoras analógicas

Despliegue:

Máquinas calculadoras analógicas

véase

Computadoras analógicas electrónicas

Computadoras analógicas

véase también

Computadoras analógicas electrónicas

008 860327 bnban znnba bnbbb bbbbb bbbab anabw bbbu

040 bb\$a NUC \$c CM \$d CM **UNION DE BIBLIOTECAS**

151 b0 \$a Roma \$x Antigüedades

360 bb\$1 la subdivisión \$a Antigüedades romanas \$1 bajo nombres
de países, ciudades, etc.

450 b0 \$a Antigüedades romanas

SECRETARIA DE SEGURIDAD DE EXTERNA

TEXTO DE TEXTO

Registro de referencia (temas con trazados)

Registro de referencia:

GUIA ***** nzbbb22 ***** nbb4500

001 exb99123457b

005 19860327111910.5

008 860211 nbnbn znbb babbb bbbbb bbbnb anbb bbbu

040 bbSa NUC SC CM SD CM

150 b0Sa Antigüedades romanas

260 bbSi la subdivisión Sa Antigüedades romanas Si bajo nombres
de países, ciudades, etc.

Registro de subdivisión

.
.
.

008 870520 inbdn znnbb babbb bbbbb bbbnb annbb bbbu

040 bb \$a NUC \$C CM \$d CM

150 b0 Control y prevención

680 bb \$i Usese bajo los nombres de enfermedades transmisibles
y bajo el encabezamiento \$a Enfermedades transmisibles.

Despliegue:

Roma-Antigüedades

Véase además la subdivisión

- Antigüedades romanas, bajo nombres de países, ciudades, etc.

Antigüedades romanas

véase

Roma - Antigüedades

Véase además la subdivisión

- Antigüedades romanas, bajo nombres de países, ciudades, etc.

Despliegue:

- Control y prevención

Usese bajo los nombres de enfermedades transmisibles y bajo el encabezamiento

Enfermedades transmisibles