

Modificaciones a las Reglas de Catalogación y la Infraestructura de OCLC para una Organización Efectiva de los Recursos de Internet

Ingrid Hsieh-Yee

Profesor, Escuela de Bibliotecología y Ciencias de la Información
Universidad Católica de América, Washington, D.C.
hsiehyee@cua.edu

Traducción por

Filiberto Felipe Martínez -Arellano

Investigador, Centro Universitario de Investigaciones Bibliotecológicas
Universidad Nacional Autónoma de México
felipe@cuib.laborales.unam.mx

Este trabajo originalmente apareció en "Proceedings of the OCLC Internet Cataloging Colloquium, San Antonio, Texas, January 1996" y es traducido y reproducido con la correspondiente autorización. Copyright, OCLC Online Computer Library Center, Incorporated, 1996, 1997. Todos los derechos reservados.

Título original: Modifying cataloging practice and OCLC infrastructure for effective organization of Internet resources

Contenido

- [Introducción](#)
 - [Porque las bibliotecas y los catalogadores deberán efectuar esta tarea](#)
 - [Como catalogar los recursos de Internet](#)
 - [Descripción](#)
 - [Puntos de Acceso](#)
 - [Análisis temático](#)
 - [La infraestructura de OCLC](#)
 - [Conclusión](#)
 - [Referencias](#)
-

Introducción

La dinámica naturaleza de los recursos de Internet representa un gran reto para los organizadores de la información. Una página web puede ser cambiada de lugar dentro del mismo sitio, desplazada a otro diferente, o bien, desaparecer por completo; su contenido puede ser actualizado ligera o substancialmente en un periodo muy corto; a medida de que la gente cambia, diferentes grupos de trabajo pueden tener la responsabilidad de las páginas; el auspicio institucional puede cambiar y nuevos métodos de acceso pueden ser agregados a un documento después de su aparición inicial. Adicionalmente, la falta de normalización en la presentación de la información dentro de la World Wide Web complica la tarea de los organizadores de la información. Por ejemplo, la mención de responsabilidad puede ser vaga o estar implícita dentro del mismo documento, el título puede no ser descriptivo o específico y la información referente a su publicación, a veces necesita ser inferida. Estos retos nos llevan a reconsiderar cómo los recursos de Internet pueden ser representados y organizados de una mejor manera.

Porque las bibliotecas y los catalogadores deberán de efectuar

esta tarea

La necesidad de organizar la información que existe en Internet es atestiguada por la serie de encomiables esfuerzos que varios investigadores han hecho para lograr un control de dicha información. La mayoría de los motores de búsqueda que actualmente existen fueron desarrollados por profesionales de la computación. La emoción inicial generada por estas herramientas motivó que algunos usuarios llegaran a la conclusión de que el problema que representaba organizar la información dentro de Internet estaba resuelto. Sin embargo, una vez que el efecto de la novedad había pasado, los usuarios se dieron cuenta, rápidamente, de la imprecisión de tales herramientas. Algunos diseñadores como Martin Coster [1] han enumerado sus limitaciones y advertido contra un a total confianza en ellas. Otro gran intento para organizar la información de Internet ha sido la identificación del núcleo de elementos que pueden constituir los metadatos de un documento para describirlo y hacerlo accesible. El "encabezamiento se mántico" [2], el "núcleo de Dublin" [3] y las " características uniformes de recursos" [4], son algunos de los ejemplos más conocidos del producto de las deliberaciones entre computólogos, bibliotecólogos, editores, vendedores, a rchivistas e ingenieros. Sin embargo, es un tanto incomprensible que muy pocos catalogadores hayan sido involucrados en estos esfuerzos, a pesar de que durante mucho tiempo son quienes han tenido a su cargo la identificación de los metadatos para elaborar los registros de información (actividad que ha sido llamada catalogación dentro de la Bibliotecología). El paralelismo que existe entre las especificaciones para registrar estos metadatos y aquellas utilizadas en los registros de catalogación salta a pri mera vista. Por otro lado, los esfuerzos que se han hecho para adaptar el encabezamiento TEI (Text Coding Initiative) al formato MARC [5], son evidencias de que las normas y métodos utilizados en la catalogación tienen mucho que contribuir a la organizac ión de los recursos que existen en Internet.

Los principios sobre los que se basa la organización de la información [6] son:

1. La determinación de qué recursos existen y la selección de aquellos que son relevantes para las necesidades de los usuarios.
2. La descripción de los recursos seleccionados.
3. El proporcionar los puntos de acceso necesarios, incluyendo el control de autoridad de los mismos.
4. El análisis del contenido de los recursos seleccionados.
5. El proporcionar la información necesaria para la localización de esos recursos.

En comparación con los motores o los servicios de búsqueda en Internet, las bibliotecas están mejor preparadas para llevar a cabo la selección de recursos electrónicos debido a la experiencia que han obtenido en la adquisición de materiales en diversos fo rmatos para sus usuarios. La experiencia de los bibliotecólogos en la selección de los recursos y el conocimiento de los componentes de su comunidad, les aseguraran el éxito en la evaluación y selección de los recursos que existen en Internet. En particular, los catalogadores deberán de ser involucrados en la organización de estos recursos puesto que han aplicado los principios de la organización bibliográfica a la catalogación de materiales en diversos formatos, y por lo tanto, son capaces de aplicar est os principios a la catalogación de los recursos de Internet con igual eficiencia.

Como catalogar los recursos de Internet

¿Como deberá de ser realizada esta actividad? A partir de la catalogación de 160 documentos de Internet durante un periodo de dos meses, nos dimos cuenta que utilizar un nivel de catalogación completo era una tarea que requería de mucho tiempo. Adicional mente, algunos de los elementos señalados para el segundo nivel de descripción en las RCA2, podrían ser de poca utilidad para los usuarios. Resultaba evidente que para adaptarse a la naturaleza dinámica de los recursos de Internet, un nivel de descripción diferente debería de ser derivado de las actuales reglas. Por lo anterior y con

el fin de lograr un balance entre la rapidez en la creación de un registro y su calidad, así como entre esa rapidez y un pronto acceso a los recursos de Internet, un nivel de catalogación mínima es propuesto. Este nivel (mencionado de aquí en adelante como nivel de catalogación M) tiene como objetivo apoyar el papel del catálogo como una herramienta de búsqueda, una herramienta de evaluación y un mecanismo de localización de los recursos de Internet a través de la inclusión de aquellos datos esenciales para su identificación y ubicación temática. El nivel de catalogación M sigue las reglas de puntuación especificadas en las RCA2 e incluye las 8 áreas señaladas para el segundo nivel de descripción. Sin embargo, algunos elementos han sido simplificados. Los problemas y soluciones (P&S) que se discuten a continuación, muestran la forma en la cual pueden ser registrados los datos.

Descripción

Area 1: Título y mención de responsabilidad

Fuentes prescritas de información: La pantalla más antigua.

La transcripción del título propiamente dicho deberá de ser realizada en la forma prescrita por la regla 1.1B1 de las RCA2. La fuente del título deberá de ser incluida en una nota para ayudar al usuario a entender el registro.

P&S: Si la mención de responsabilidad no está prominentemente mencionada en la fuente principal de información, ésta puede ser omitida del Area 1. Evite la utilización de información entre paréntesis para este elemento. En lugar de lo anterior, utilice una nota general para incluir aquella información sobre los autores que se encuentre implícita en cualquier otra parte del documento.

Los llamados "diseñadores web," quienes codifican un documento, deberán de ser tratados como editores. Inclúyalos en el Area 1 si son mencionados de una manera prominente en la fuente principal de información. Las RCA2 referentes a la mención de hasta tres individuos, son aplicables para registrar a los autores y editores de este tipo de documentos.

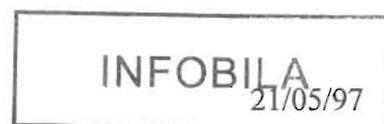
Designación general del material: Un nuevo término, [*Archivo de Internet*], es utilizado para distinguir los recursos de Internet de los archivos de computadora almacenados en otro tipo de dispositivos para acceso directo. Los usuarios podrán utilizar este nuevo DGM para restringir sus búsquedas a los recursos de Internet.

Area 2: Edición

Fuentes prescritas de información: Las pantallas que se encuentran antes y después del cuerpo del documento (este concepto es similar a los "preliminares" y "colofón" dentro de las monografías.)

P&S: Frecuentemente, los recursos que existen en Internet no incluyen una mención formal de edición. Esta área puede ser omitida si no existe ninguna información disponible. Cuando los datos de la edición son incluidos en un documento de Internet, éstos tienden a aparecer en las primeras o en las últimas pantallas. El registro de las menciones formales de edición sigue la regla 1.2B1.

Como diversos participantes en *Intercat* señalan, si cada mención de actualización fuese tratada como una mención de edición, sería necesario crear un nuevo registro bibliográfico cada vez que un documento fuese puesto al día [7]. La confusión resultante de tal práctica desvirtuaría el objetivo de la catalogación. Por lo anterior, a diferencia de lo mencionado en las RCA2 y en el documento *Cataloging Internet Resources : A Manual and Practical Guide* [8], las actualizaciones no deberán de ser tratadas como otras ediciones; tal información puede ser incluida en una nota general (véase las indicaciones en la parte correspondiente).



Area 3: Características del archivo

Los términos especificados dentro de las RCA2 para esta área no proporciona información única para identificar o localizar los recursos de Internet. Por lo anterior, esta área es omitida del nivel de catalogación M.

Area 4: Publicación, Distribución, etc.

Fuentes prescritas de información: Las pantallas que se encuentran antes y después del cuerpo del documento (este concepto es similar a los "preliminares" y "colofón" dentro de las monografías.)

P&S: En muchas ocasiones, la información correspondiente al área de publicación necesita ser inferida y a veces la información incluida en la pantalla que contiene el título es diferente a la correspondiente al sitio donde se encuentra el documento. Por ejemplo, un autor cuya institución es la *Catholic University of America*, podría tener una página en el web de la *University of North Carolina*. La solución a esta situación es tomar el sitio donde se encuentra la página como el editor e incluir, opcionalmente, la afiliación del autor en una nota general. La fundamentación para este tratamiento es que el sitio anfitrión proporciona un foro para que el autor presente sus ideas. Si la información correspondiente al área de publicación no aparece en las fuentes prescritas de información, ésta deberá de ser registrada entre corchetes.

Fecha de Publicación : Dado que muchos de los recursos de Internet son actualizados frecuentemente, la fecha de actualización es cambiante y por lo tanto no deberá de ser utilizada como la fecha de publicación [9]. Siguiendo el tratamiento que se aplica a las publicaciones que se actualizan constantemente a través de hojas sustituibles, es recomendable que la fecha inicial de aparición de un documento web, si es conocida, se registre con un guión para indicar que éste se encuentra en constante actualización. Si la fecha de publicación no se incluye en las fuentes prescritas de información, ésta deberá de ser registrada entre paréntesis.

Area 5: Descripción Física

Esta área es omitida del registro para recursos de acceso remoto.

Area 6: Serie

Fuentes prescritas de información: Cualquier parte del documento. Unicamente se registran el título formal de la serie y la información correspondiente a la numeración.

Area 7: Notas

Fuentes prescritas de información: Cualquier parte del documento

El área 7 proporciona información de gran utilidad, la cual complementa la información incluida en las primeras 6 áreas del registro.

P&S: Las notas deberán de ser incluidas, de acuerdo a su importancia, en el siguiente orden:

7.11 Campo 856. Localización y acceso electrónico

Dado que el principal interés de muchos usuarios es como acceder un documento, esta nota deberá de ir en primer lugar.

7.12 Múltiples campos 856: Una nota deberá de ser empleada para para cada uno de los diferentes métodos de acceso.

7.13 Mantenimiento de los URL: En ocasiones, los URL cambian o son modificados. Su

mantenimiento puede ser hecho a través del PURL (Persistent Uniform Resources Locator) de OCLC [10].

7.21. Fuente del título. Una nota requerida.

7.22 Variantes del título: Incluye información como el título HTML o los títulos tomados de otras fuentes.

7.31 Información sobre el autor: Si la mención de responsabilidad no es registrada en el Area 1, una nota sobre el autor deberá de ser incluida si esta información se localiza en cualquier otra parte del documento. Los editores y diseñadores de webs pueden ser incluidos en esta nota, la cual deberá de contener las fuentes de donde se tomó tal información. La regla para la mención de hasta tres individuos, también es aplicable a autores, editores y diseñadores de webs.

7.32 Otras personas u organismos: Aquí se registran las personas u organismos no incluidos en la mención de responsabilidad pero considerados importantes para la identificación de los documentos .

7.41 Actualidad de la información [11]: De manera similar al tratamiento de las publicaciones seriadas y las publicaciones actualizadas por hojas sustituibles, se deberá de incluir una nota para indicar la fecha de creación del registro. Por ejemplo:

Descripción basada en : versión de Febrero 15, 1995.

Esta nota le indicará al usuario la fecha en la cual el registro fue creado y le ayudará a entender y aclarar cualquier discrepancia entre el registro y un documento de Internet recuperado posteriormente por el usuario.

7.42 Actualizaciones al registro: Sí el catalogador decide actualizar un registro, la información correspondiente a la fecha de la última actualización deberá de ser incluida para ayudar al usuario a entender los cambios en el registro. Por ejemplo:

Actualización del registro basada en: versión de Junio 30, 1995.

Los catalogadores que decidan no actualizar los registros no necesitan utilizar esta nota.

7.51. Requerimientos del sistema (campo 538): Esta nota puede ser simplificada incluyendo únicamente el equipo y programación especiales que son necesarios para acceder un documento. Por ejemplo, impresora *postscript* o interfase para gráficas

7.52 Modo de acceso (Campo 538): Esta información es duplicada con la incluida en el campo 856 y por lo tanto puede ser omitida.

7.6 Nota de Contenido/Resumen: Algunos catalogadores han optado por incluir las partes que integran un documento (contenido); sin embargo, un resumen cuidadosamente elaborado, generalmente es más informativo que una larga lista de contenido y por lo tanto deberá de preferirse.

Area 8. Números normalizados y términos de disponibilidad

Esta área no se encuentra actualmente en uso. Se podrá utilizar para registrar el URN (Nombre Uniforme de Recursos) cuando éste se encuentre disponible.

Con base en todo lo anteriormente expuesto, un registro para un documento de Internet puede ser tan completo como un registro bibliográfico de segundo nivel de catalogación, si todos los elementos pueden ser identificados y registrados; o bien, puede ser tan corto como el registro que se muestra a continuación:

Título propio [Archivo de Internet] / Mención de responsabilidad
 . -- [Lugar de la institución anfitriona del documento: Nombre
 de la institución anfitriona del documento, Fecha inicial del docu-
 mento -]
 Nota de localización y acceso. (Campo 856)
 Fuente del título. (Nota 500)
 Editores u organismos importantes para fines de identificación.
 (Nota 500)
 Actualidad de la información. (Nota 500)
 Actualizaciones del registro. (Nota 500)
 Requerimientos del sistema. (Nota 538)
 Resumen. (Nota 520)

Cambios en los campos 500 (notas) del formato MARC: Para simplificar el mantenimiento de la información incluida en las diversas etiquetas 500, una serie de valores numéricos deberán de ser creados para utilizarse como indicadores y especificar la naturaleza de cada una de las notas registradas en las diferentes etiquetas 500. Si esto es difícil de lograr, otro campo deberá de ser utilizado para registrar la actualidad y actualización de la información correspondiente a cada una de las notas.

Puntos de Acceso

Los puntos de acceso para cada uno de los autores, editores y organismos (no más de tres de cada tipo) directamente relacionados con la producción de un documento, deberán de ser incluidos. Adicionalmente, el trabajo de autoridad respectivo deberá de ser realizado.

Análisis temático

Dado que el contenido de la mayoría de los documentos que existen en Internet tiende a ser amplio, la opción tradicional de sintetizar dicho contenido parece ser la adecuada para este tipo de recursos. Los catalogadores necesitarán combinar el uso de un vocabulario controlado en los encabezamientos de materia con la utilización de un lenguaje natural dentro del resúmen y proporcionar a los usuarios diversas opciones en la recuperación de los recursos de Internet a través del tema que abordan. Adicionalmente, todos los términos utilizados como elementos de recuperación temática deberán de estar sujetos a un trabajo de control de autoridad. Las listas de encabezamientos tradicionales como LCSH, MeSh y Sears, así como los esquemas de clasificación de Dewey y LC, deberán de continuar siendo utilizados para lograr una integración de los documentos de Internet con otros recursos que existen en la colección y lograr una ubicación por materia uniforme.

La infraestructura de OCLC

Aunque es deseable contar con un sistema de información que organice todo lo que existe en Internet, actualmente tal sistema no es factible ni podría ser efectivo por las siguientes razones:

1. Existen demasiados recursos en Internet.
2. La calidad de muchos de esos recursos es cuestionable.
3. Muchos recursos son de naturaleza efímera y pueden ser de escaso valor para los usuarios.

Un modelo que podría tener un mayor costo-efectividad, sería un sistema que incluyera aquellos recursos cuya calidad hubiese sido evaluada previamente. La actual infraestructura de OCLC permite que las bibliotecas seleccionen los recursos que son relevantes para sus usuarios y contribuyan con los registros originales de esos recursos, en el caso de que no existan en OCLC. A través de este proceso, las bibliotecas tienen la ventaja de utilizar el Archivo de Autoridad de Autores y el Catálogo

de Autoridad de Materia de la Biblioteca del Congreso, así como el módulo de creación de registros de OCLC. Por otra parte, las reglas seguidas para la catalogación de los recursos de Internet son esencialmente las mismas que aquellas utilizadas para la creación de registros originales en OCLC, excepto que la parte descriptiva ha sido substancialmente simplificada.

Los esfuerzos cooperativos de las bibliotecas permitirán a los bibliotecólogos identificar una gran cantidad de los recursos de Internet que son de calidad. Los registros bibliográficos incluirán aquellos puntos de acceso e información temática que les son de utilidad a los usuarios en la búsqueda y selección de los recursos que existen en Internet. Asimismo, contarán con las relaciones o ligas, actualizados a través de PURLs, para recuperar fácilmente aquellos recursos seleccionados por los usuarios. Un sistema de información con tales características, podrá ser reconocido por la calidad de sus registros y efectividad para recuperar la información, y por lo tanto, preferido por los usuarios. Dado que tal sistema sería potencialmente rentable, las bibliotecas y OCLC podrían compartir su administración. Mínimamente, tres categorías de usuarios, los cuales tendrían que cubrir una serie de cargos con base en su relación con OCLC, pueden ser identificados. Primeramente, las bibliotecas que contribuyan con registros originales al nuevo sistema obtendrían los créditos correspondientes y las cuotas de acceso al sistema podrían ser substancialmente reducidas. Estas bibliotecas podrían tener la opción de no incluir ligas o relaciones desde sus OPACs. Segundo, las bibliotecas que no contribuyeran con registros a OCLC, pero desearan copiarlos a sus respectivos OPACs o únicamente efectuar búsquedas en el sistema, tendrían que cubrir una cuota. Tercero, aquellos usuarios finales que prefieren tener acceso directo a OCLC a través de Internet, cubrirían una cuota fija por un conjunto de búsquedas para estimular e incrementar el uso del sistema. El sistema podría ser colocado en Internet y competir con otros motores y servicios de búsqueda. Las utilidades que el sistema pueda generar, serían utilizadas para apoyar su desarrollo y a las bibliotecas que contribuyan con registros.

Conclusión

Los recursos que existen en Internet necesitan ser evaluados, seleccionados, descritos y analizados temáticamente para facilitar su acceso. Por mucho tiempo, los catalogadores han estado agregando un valor añadido a los medios para registrar información y los principios utilizados dentro de la organización de la información pueden ser efectivamente aplicados para organizar los recursos de Internet. Los registros que integren el catálogo de estos recursos, deberán de incluir aquellos elementos que son esenciales para su identificación y localización. Debido a la gran cantidad de este tipo de recursos, la mejor opción para crear este catálogo es a través de un esfuerzo cooperativo. Para lograr lo anterior, un nivel de catalogación mínimo ha sido propuesto y el papel que OCLC podría jugar en esta tarea ha sido señalado. Los bibliotecólogos con su experiencia en las actividades de catalogación y OCLC con su infraestructura de información, podrían ser aliados para mejorar la búsqueda, el acceso y la utilización de los recursos que existen en Internet.

Referencias

1. Martijn Koster, "Robots in the Web: Threat Or Treat," [<http://info.wehcrawler.com/mak/projects/robots/threat-or-treat.html>]
2. Binp C. Desai, "WWW Workshop: Navigation Issues," [<http://www.cs.concordia.ca/bcd/navigate.html>]
3. Para una descripción del "Núcleo de Dublin", véase Priscilla Caplan, "You Call It Corn, We Call It Syntax-Independent Metadata for Document-Like Objects," *Public-Access Computer Systems Review* 6, no. 4 (1995): 19-23; y un resumen del OCLC/NCSA Metadata Workshop at [http://www.cnri.reston.va.us/home/dlib/July_95/07weibel.html]
4. Ron Daniel, "Proposed URC External Representation," [<http://www.acl.lanl.gov/URI/ExtRep/urc0.html>]

5. Véase por ejemplo, "TEI Guidelines for Electronic Text Encoding and Interchange," [<http://etext.virginia.edu/TEI.html>]
6. Principios similares fueron discutidos por Arlene G. Taylor en "The Information Universe: Will We Have Chaos or Control?" *American Libraries* 25 (1994): 629-32.
7. El problema del sitio web y la mención de responsabilidad fueron discutidos en la lista de *InterCat* en Octubre de 1995 por Neil R. Hughes, Angelina G. Joseph, Glenn Patton, y Ellen McGrath. Los mensajes se encuentran almacenados en el archivo de la lista de discusión en <http://ftplaw.wuacc.edu/listproc/intercat/9510/maillist.html>
8. Cataloging Internet Resources: A Manual and Practical Guide, ed. Nancy B. Olson. Dublin: OCLC, 1995.
9. Adele Hallam, *Cataloging Rules for the Description of Looseleaf Publications: with Special Emphasis on Legal Materials*. 2a. ed. Washington, D.C.: Office for Descriptive Cataloging, Library of Congress, 1989.
10. El concepto de PURL fue introducido por Erik Jul y discutido en la lista de *InterCat* en Enero de 1996 por Erik Jul, Vianne Sha, Priscilla Caplan, y Stuart Weibel. Los mensajes están almacenados en el archivo de la lista en <http://ftplaw.wuacc.edu/listproc/intercat/9601/maillist.html>
11. Hallam, *Cataloging Rules for the Description of Looseleaf Publications*.
12. Agradezco a mis estudiantes del curso sobre Organización de los recursos de Internet la ayuda para desarrollar algunas de las ideas incluidas en este documento.

El contenido de este documento refleja las opiniones del autor y traductor y no necesariamente las del Centro Universitario de Investigaciones Bibliotecológicas



Regresar a Portada



Diseño y elaboración del sitio
Filiberto Felipe Martínez Arellano
felipe@cuib.laborales.unam.mx
con la colaboración de
Sergio Arreguín Meneses
Abril 23, 1997
