

Aproximación a la catalogación de documentos electrónicos de acceso remoto

Manuel José Pedraza Gracia

Departamento de Ciencias de la Documentación
e Historia de la Ciencia
Universidad de Zaragoza

0.1. Resumen

Se analizan los problemas generales que existen para la organización y recuperación de la información en Internet y los diversos intentos para solventarlos. Se estudian posibles vías de actuación para el tratamiento de recursos informativos de Internet por parte de las bibliotecas de investigación. Se analiza la normativa española de catalogación de archivos de ordenador aplicadas a los archivos de ordenador de acceso remoto. (Autor)

Palabras clave: Internet. Organización y recuperación de la información. Catalogación. Archivos de acceso remoto.

0.2. Abstract

Analysis of general problems of Internet resources organization and retrieval and of the different attempts to solve it. The possible approaches available for research libraries are considered too. The Spanish standards and guidelines for remote access computer archives are analyzed.

Keywords: Internet. Information organization and retrieval. Cataloguing. Remote access computer files.

1. Introducción

El crecimiento exponencial de Internet debido al volumen de la información que se incorpora continuamente y su carácter altamente heterogéneo dificulta el acceso a la misma. Las diferencias de contenido, de intención y de calidad enmarañan la red que cada momento se torna más tupida e impracticable. Ha sido necesario que se intercalen entre el usuario y la red organizadores y recuperadores de información, mecanismos de búsqueda y recuperación de información, que permitan acceder a los contenidos informativos que interesan al usuario de la red.

Diversos esfuerzos se han dirigido a precisar los fines y modo de actuación de estos organizadores-recuperadores de información, centrándose especialmente en aspectos tan sustanciales como puedan ser los elementos informativos imprescindibles para la identificación del recurso, su estructuración, definición de sus contenidos, etc. Entre estos esfuerzos se encuentran, por una parte, iniciativas completamente novedosas en lo referente al tratamiento documental que han revelado aspectos muy positivos que habrán de estudiarse para ser aplicados en un futuro muy próximo; y, por otra, iniciativas que persiguen el tratamiento de los recursos de Internet con herramientas tradicionales. Estas iniciativas fueron frecuentemente cuestionadas. Sin embargo, con el transcurso del tiempo, ha resultado evidente que existe una importante demanda por parte de algunas instituciones de registros bibliográficos de recursos de Internet distribuidos a través de catálogos en línea.

OCLC (Olson, 1995) ha iniciado importantísimos trabajos en esta dirección. Las premisas básicas sobre las que se sustentan esos proyectos de OCLC para proveer el acceso a los recursos de Internet son:

- Existe una gran cantidad de información valiosa disponible a través del Internet. Existen colecciones valiosas de textos, imágenes y sonidos para determinados grupos de interés o de investigación -colecciones que pueden hasta ser el tema de discusión del estado de la cuestión en estos grupos- esa información en algunos casos existe solamente en forma electrónica.
- Existe una necesidad de organizar determinados recursos de Internet para facilitar su acceso.
- La utilización de las técnicas bibliotecarias y los procedimientos de creación de registros para la recuperación a través catálogos en línea existentes es el método más eficiente de acceder estos recursos.

Estas mismas premisas pueden servir como puntos de reflexión para nuestras bibliotecas científicas que, cada vez más, tienen que apreciar los recursos informativos de Internet como nuevos documentos que pueden enriquecer y actualizar su colección.

2. Dificultades para la organización y recuperación de la información en Internet

El principal problema que se presenta en lo que hace referencia a la organización y recuperación del material informativo existente en Internet es que éste no ha sido tamizado por una norma, no ha sido expuesto a los rigores de un editor que presenta sus publicaciones según esquemas más o menos rigurosos. Un segundo problema es la propia calidad efímera de gran parte de este material que ha eliminado la necesidad de su estructuración. El autor, por su parte, suele

obviar cualquier otro aspecto que no sea el de la inmediatez en proporcionar una determinada información. Por último, el catalogador se ha resistido a establecer normas descriptivas para un material que, en ese aspecto, resulta demasiado resbaladizo para establecer generalizaciones, donde incardinar una normativa -no es imprescindible la presencia de título, no es necesario precisar en el título los contenidos, se prefiere una dirección de correo electrónico que el nombre del autor, la mención de responsabilidad puede ser vaga, no es preciso indicar la versión o actualización, no es necesario indicar la fecha de creación, etc.-.

Sin embargo, los propios documentos resultan familiares desde la perspectiva de la normativas de catalogación, son, después de todo, publicaciones serias, monografías, publicaciones de hojas sueltas para ser actualizadas, materiales gráficos, ... De hecho, cuando se trata de simples versiones electrónicas de un documento que ya existe en papel, los retos a los que se enfrenta el catalogador no resultan tan insalvables. Este fenómeno se produce debido a que el reto no se encuentra en los propios objetos que se tratan, sino en la mutabilidad, el camaleonismo, los cambios de ubicación sin esfuerzo, el cambio de paternidad, el puntero desde el interior de un documento que hace del objeto original una parte de otra mayor, etc. (Hsieh-Yee, 1996).

Este dinamismo queda patente en aspectos como los que se enumeran a continuación:

- Un recurso puede moverse de un lugar a otro, o simplemente desaparecer totalmente sin previo aviso.
- Los contenidos de un documento pueden actualizarse bien ligeramente o bien sustancialmente en un breve período de tiempo.
- La responsabilidad de un documento en Internet puede variar con el transcurso del tiempo y con las modificaciones que puedan realizarse en ella.
- El patrocinio puede cambiar.
- Pueden añadirse después de la aparición de un documento aspectos sustanciales que modifiquen sus contenidos.

No obstante, se sigue procurando crear un gran sistema de información que organice toda la información existente en Internet. Si se lograra (Hsieh-Yee 1996) tal sistema ni sería efectivo ni posible porque:

- Hay demasiados recursos en Internet.
- La calidad de algunos recursos es discutible.
- Muchos recursos son efímeros por naturaleza y pueden tener un valor muy limitado para los usuarios.

3. Propuestas para el tratamiento de información en Internet

La necesidad de organización de información en Internet está atestiguada por los esfuerzos de diversos investigadores para controlar la información. Como se ha dicho anteriormente, algunos de estos esfuerzos en el tratamiento de la información en Internet son novedosos, entre estos destacan la utilización de motores de búsqueda y las realizaciones que han dado como resultado en 'enfoque meta-data'.

3.1. Los motores de búsqueda

Un motor de búsqueda, grosso modo, es una herramienta que permite acceder mediante palabras clave a los recursos de Internet indizados en su base de datos. La mayoría de los motores de búsqueda existentes fueron desarrollados por informáticos, y la excitación inicial generada por tales herramientas han llevado algunos usuarios a la conclusión de que el problema de organización de información en el Internet está resuelto. Sin embargo, estos motores de búsqueda en Internet tienen tanto puntos fuertes como débiles (Taylor y Clemson, 1996).

Entre los puntos fuertes figuran:

- Son gratuitos.
- Se obtiene la información sin procesamiento previo por parte del usuario.
- Permiten acceder a sus índices 24 horas al día, siete días a la semana.
- Los documentos recuperados del índice se encuentran inmediatamente disponibles.
- A pesar del dinamismo y el crecimiento continuo de Internet se accede a páginas recién creadas.
- Las páginas personales y otras de carácter efímero son también accesibles.

Entre los puntos débiles figuran:

- Se recuperan muchas entradas duplicadas en un mismo conjunto de búsqueda recuperado.
- Los recursos de búsqueda no muestran los contenidos de sus bases de datos ni proveen de una descripción de los criterios de indización utilizados.
- El vocabulario normalmente no está controlado, y no están normalizadas las reglas de puntuación y utilización de mayúsculas.
- Si el documento origen ha desaparecido o ha cambiado de servidor no es recuperable mediante el índice.
- Ofrece generalmente interfaces pobres y opciones de búsqueda limitadas.

- Presentan notables dificultades para la recuperación de información por nombres de persona y entidad, según han demostrado Taylor y Clemson (1996).

A todo esto hay que añadir que, a pesar de lo interesante que resultaría la utilización de motores de búsqueda de usuario que pudiesen recuperar mediante robots la información sobre un determinado tema en Internet, su utilización resultaría muy problemática ya que conducirían a la larga a la saturación de la red y a la inoperancia de la misma.

Pero después de que los efectos de la novedad se han consumido los usuarios han comprendido pronto la imprecisión de estas herramientas. Es patente que la obtención de resultados positivos desciende conforme aumenta el tamaño de Internet y que las bases de datos creadas por estas máquinas de búsqueda pueden volverse mayores que los propios recursos de la red. Pero, sobre todo, es evidente que ninguna de estas máquinas de búsqueda se puede adecuar a documentos no textuales: ejecutables, imagen, documentos comprimidos, grabaciones sonoras, o imagen en movimiento.

3.2. El enfoque Metadata

El término 'Metadata' se utiliza para designar el conjunto de elementos informativos que pueden utilizarse para describir y representar objetos de información. O lo que es lo mismo, metadata es información acerca de información. Un registro de catálogo o cualquier otra forma de descripción es metadata. Caplan (1995) al tratar el que se ha llamado Núcleo de Dublín indica que podría haberse llamado perfectamente catalogación pero que para algunos este término comportaba un excesivo equipaje -esencialmente Reglas de Catalogación y formato MARC-, hay que tener en cuenta, además, que son el resultado de las deliberaciones de informáticos, editores, proveedores, archiveros, otros gestores de información y, también, bibliotecarios, a pesar de que los catalogadores han estado durante mucho tiempo dedicados a la creación de 'metadata' para centros de información.

3.2.1. Características de Recurso normalizado (URC)

El URC está destinado a facilitar la recuperación de información por parte de máquinas. En estos momentos un URC contiene diez elementos opcionales: URN, autor, título, materia, resumen, ubicación, versión, acceso, firma electrónica, y revista. El URC puede contener además información: revisiones, restricciones de uso, etc. Su pretensión es asegurar la autenticidad de los recursos, y cierta capacidad de restricción de acceso. En este momento, los elementos de información en el URC están identificados por etiquetas descriptivas.

Entre los elementos que componen en URC destaca el URN o Nombre de Recurso Normalizado (Uniform Resource Name) -asignación de un número de identificación fijo, único, e independiente de la ubicación del documento en la red, similar al ISBN en el mundo de las monografías-. Si el URL o localizador universal de recursos (Universal/Uniform resource locator) o, si se quiere, dirección electrónica, cambia se puede asociar el nuevo URL al URN, es decir el URN permanece aún cuando el URL pueda cambiar.

3.2.2. *La cabecera de la Iniciativa para la codificación de textos (TEI) -Text encoding initiative*

La Iniciativa para la codificación de textos, utiliza SGML 'Standard Generalized Markup Language' como la base para codificar e intercambiar textos legibles por máquina entre comunidades de investigación. Dentro de la iniciativa y tiene importancia específica la provisión de información documental o bibliográfica acerca de textos electrónicos.

La cabecera TEI vendría a ser una portada electrónica para el texto electrónico. El esquema de codificación de la cabecera TEI es flexible, sus únicos elementos de información obligatorios son la declaración de título, la declaración de publicación, y de declaración de descripción fuente, que son los componentes de la descripción de archivo.

La fiabilidad de encabezados TEI para la recuperación de la información es variable debido a la flexibilidad de codificación. Sin embargo, cuando la cabecera está unida a un texto codificado, el usuario puede buscar además por texto completo y también por los *tags* de SGML.

3.2.3. *El Núcleo de Dublín*

La pretensión del Núcleo de Dublín es que el propio productor de información codifique sus propios documentos desde el mismo momento de su creación. Es un conjunto simple de trece elementos de información que persiguen el acceso al documento y no la descripción del mismo.

La estructura de datos del Núcleo de Dublín es simple:

- a) Tema: Materia.
- b) Título: El nombre del objeto.
- c) Autor: La persona o personas responsables del contenido intelectual del objeto.
- d) Editor: El responsable de hacer disponible el objeto .
- e) Otros responsables: La persona o personas, como editores literarios y transcritores, que haya hecho otras aportaciones intelectuales significativas para el trabajo.

- f) Fecha: La fecha de publicación.
- g) Tipo de objeto: El género del objeto, como novela, poema, o diccionario.
- h) Forma: La representación del objeto, como archivo ejecutable.
- i) Identificador: Número utilizado para identificar unívocamente el objeto.
- j) Relación: Relación con otros objetos.
- k) Fuente: Los objetos, impresos o electrónicos, de los que deriva este objeto.
- l) Lenguaje: El lenguaje del contenido intelectual.
- m) Cobertura: La ubicación espacial y la característica de la duración temporal del objeto.

Cada elemento de información es opcional y repetible y puede también tener subelementos que califican o explican el contenido del elemento.

La fiabilidad de la información descansa en un creador de gran calidad; puesto que no hay mejor conocedor de la información de un documento que su creador. Sin embargo, se introduce la disparidad de criterios, este aspecto hace que la fiabilidad desaparezca. Un "metadata" suministrado por el autor es mejor que ninguno, pero peor que un registro de catálogo profesionalmente preparado.

4. Las bibliotecas y el tratamiento de los recursos informativos de Internet

El paralelo entre estas nuevas especificaciones -metadata- y los registros de catalogación es evidente. Las normas y prácticas de las bibliotecas, más concretamente en lo referente a la catalogación, tienen mucho que contribuir a la organización de recursos de Internet.

Pero, además, las bibliotecas (Sha; Patrick; Kochtanek, 1996) garantizan el acceso público y el acceso para la formación de todos los ciudadanos a Internet - sólo el 8,4% de americanos de raza negra y el 18.3% de raza blanca tienen ordenadores en casa (*ibidem*)-.

Los principios de organización de información, esté ésta donde esté, incluyen:

- La determinación de los recursos existentes y la selección de recursos relevantes a las necesidades de los usuarios.
- La descripción de recursos seleccionados.
- La provisión de puntos de acceso y control de autoridades.
- El análisis del contenido de recursos seleccionados.

- La puesta a disposición del usuario de los 'metadata' para localizar estos recursos.

Por lo que se refiere a la determinación de recursos y selección de los mismos, los bibliotecarios, sin lugar a dudas, tienen más experiencia que los motores de búsqueda en la selección de recursos de información porque han tenido una larga experiencia en la selección de materiales en diversos formatos para sus usuarios. La experiencia de los bibliotecarios en la selección de recursos informativos y el conocimiento de las necesidades informativas de sus usuarios relativamente bien definidas garantizarán su éxito al evaluar y seleccionar recursos de Internet.

Por lo que se refiere a los otros cuatro principios, los catalogadores, en particular, deberían estar involucrados en organizar recursos de Internet porque han aplicado estos principios a la catalogación de materiales en diversos formatos y deberían poder aplicar estos principios a la catalogación de recursos de Internet con igual eficacia. Los profesionales de la información en el campo de las bibliotecas y de la bibliografía han desarrollado principios de control bibliográfico que han demostrado bastante éxito para organizar la información de modo que pueda ser recuperada de nuevo. Pueden diseñar una cita que presenta una descripción integrada y comprensible de un documento. Esta descripción indica claramente la relación de un documento con sus propias (u otras) manifestaciones así como su relación con otros documentos. En otras palabras el propósito es proveer información suficiente para permitir al investigador escoger la que es relevante y útil para ese investigador.

Además, con demasiada frecuencia hemos observado que la catalogación tiene su culminación en la realización de un registro mientras se prescinde del proceso intelectual más imperceptible de crear y aplicar normas, vocabularios, y los sistemas de descripción y clasificación que son inherentes en esos registros.

Desde esta perspectiva el catalogador tiene la oportunidad de perfeccionar una publicación de Internet en tres aspectos:

- a) Dentro de nuestras propias organizaciones estableciendo normas para desarrollo de páginas principales.
- b) En sistemas cooperativos desarrollando normas útiles para la comunidad.
- c) En uno y otro caso proporcionando palabras clave y otra información descriptiva para las páginas principales de los miembros de nuestra organización (Taylor y Clemson, 1996).

Nuestro reto es contribuir con nuestra experiencia en administración de recursos de información. Las normas de descripción han incorporado predecibi

lidad y eliminado la mayoría de la duplicación analizando los enlaces entre entidades.

Todo esto porque existen recursos informativos de gran valor y perdurabilidad entre los contenidos de la base de datos amorfa que es Internet, puesto que un modelo eficaz sería mantener un sistema que almacenase recursos cuya calidad hubiese sido evaluada. Las bibliotecas pueden seleccionar recursos relevantes a sus usuarios locales. El esfuerzo cooperativo de diversas bibliotecas permitiría a los bibliotecarios cubrir un gran número de recursos de información de Internet de calidad contrastada.

Como ha señalado Amanda Xu (1996) los recursos de internet organizados según la normativa ofrecen a beneficios importantes:

- Han sido seleccionados para satisfacer las necesidades de una comunidad de usuarios determinada.
- Han sido controlados mediante la descripción bibliográfica, la selección de puntos de acceso, el control de autoridades, y el análisis de materias.
- Los sistemas automatizados de bibliotecas en que se incorporan han sido desarrollados para gestionar búsquedas sofisticadas en cantidades de información muy grandes.
- Los recursos de Internet más importantes pueden estar integrados con las entidades bibliográficas ya indizadas con formatos MARC.

Pero, como puede deducirse, el trabajo que debe realizarse es muy importante, por esta causa es preciso establecer unas prioridades para la selección y catalogación de recursos de Internet:

- Organizar y catalogar los recursos de Internet poseídos y mantenidos por el sistema local.
- Organizar y catalogar materiales de investigación significativos que tienen interés para el usuario local.
- Organizar y catalogar herramientas significativas que puedan mejorar los servicios de referencia.
- Organizar y catalogar herramientas significativas que pueden actualizar el conocimiento y mejorar las habilidades del equipo técnico de biblioteca.

Las bibliotecas no puede perder la ocasión de ofrecer a sus usuarios el acceso a una documentación de importancia considerando, además, que a muchos recursos de Internet se puede acceder libremente y de forma gratuita, si no quieren que se encarguen de ello servicios comerciales. Desde esta perspectiva, cabría esperar que una vez localizado un recurso de Internet en un OPAC, debería accederse desde el OPAC al recurso en cuestión. El formato USMARC ha sol-

ventado el problema mediante la inclusión del campo 856 (Shieh, 1996). En apreciación de Morgan (1995), la adición de recursos de Internet en nuestro OPAC supondrá, en un primer momento, debido al aumento de registros a realizar, la necesidad de más catalogadores profesionales con un conocimiento a fondo de Internet y de tiempo para evaluar recursos de Internet con respecto a sus elementos bibliográficos; y, en un segundo momento, la necesidad de formar a nuestros usuarios para que conozcan y utilicen estos nuevos recursos y destierren de su mente que un OPAC es solamente una lista de libros.

5. Aplicación de la normativa española de catalogación al tratamiento de los recursos informativos de Internet

El formato MARC es un conjunto complejo de normas para la descripción, almacenamiento, intercambio, manipulación, y la recuperación de información bibliográfica legible por máquina. La inclusión en el formato USMARC del campo 856 ha sido adoptada para proveer la descripción, acceso y ubicación de recursos de la red. Este campo permite establecer un enlace entre el registro MARC y el texto electrónico que describe. A diferencia del resto de las nuevas normas 'metadata' incluye un formato de autoridades que facilita la recuperación. Si un recurso de Internet es de alta calidad, de interés para nuestros usuarios, y relevante para nuestras colecciones, debe pertenecer a nuestro catálogo. De este modo, presentamos a nuestros usuarios una imagen más exacta y completa de la información disponible en un área de tema determinado.

Desde una perspectiva teórica, quizá el aspecto que más se ha destacado es la ausencia de una consistencia física del documento. Es este, sin embargo, un aspecto menor ya que únicamente afecta al área de descripción física. La ausencia de materialidad ha quedado perfectamente resuelta en la última edición de nuestras reglas de catalogación mediante la omisión del área 5 [R.C., 11.5].

Obviado este aspecto, que resulta altamente llamativo. El primer factor que se debe tener en cuenta es que en Internet se pueden encontrar todo tipo de documentos: textuales, gráficos, sonoros, audiovisuales, cartográficos, seriados, ..., e incluso archivos de ordenador. desde la perspectiva descriptiva, se trata, sin embargo, de archivos de ordenador cuyos contenidos son muy diversos y que presentan características diversas, datos matemáticos, en el caso de contenidos cartográficos, periodicidad, en caso de materiales seriados. Todo ello sin dejar de ser archivos de ordenador.

El segundo factor que se ha de plantear quien quiera catalogar los recursos de Internet es el que hace referencia al nivel de descripción. Las tendencias actuales tienden a utilizar los niveles más bajos de catalogación [R.C., 1.0.9], por tres aspectos fundamentales:

- a) La calidad del registro catalográfico no está en relación con su nivel, sino con las necesidades informativas del usuario.
- b) El investigador necesita sólo de forma atingente determinada información que la normativa incorpora.
- c) Un registro realizado en un nivel muy profundo de descripción necesita mucho más tiempo -gasto- lo que interfiere en la rapidez con la que el usuario de Internet quiere acceder a una información, que por otra parte es además muy dinámica.

El inicio de los documentos de Internet y el final de los mismos suelen contener, si existe, información suficiente para la descripción. El documento en si mismo actúa como fuente principal [R.C., 11.0.1], según la normativa establecida para los archivos de ordenador. Sin embargo, es bastante frecuente encontrarse con la ausencia de información para la descripción. Otro problema que se puede presentar en los recursos de Internet es que la información no puede ser obtenida por ilegible hasta que el documento no ha sido descomprimido.

Los recursos textuales de Internet pueden dividirse principalmente en dos grupos:

- a) Sistemas y servicios, incluyendo todos los directorios y subdirectorios que contienen más que una entidad bibliográfica, que se han de catalogar como monografías, es decir, colecciones con o sin título colectivo.
- b) Archivos que contienen solamente un recurso. Estos habrán de catalogarse como documentos seriados si fuesen publicaciones seriadas electrónicas y si los documentos individuales muestran intención de actualizaciones futuras o de estabilidad como documentos monográficos. Estas descripciones habrán de matizarse con las características propias de algunos materiales en función de su contenido: material gráfico, sonoro, etc.

Los documentos de Internet presentan las siguientes características propias analizados desde la perspectiva de cada una de las áreas de las ISBD y del formato Ibermarc aplicado a archivos de ordenador.

5.1. El área 1: título y mención de responsabilidad, campo 245 [R.C., 11.1]

En caso de no encontrar información el catalogador puede considerar la utilización del título de una página principal o de cualquier descripción publicada o de una cita sobre el archivo. En última instancia puede utilizarse el nombre del archivo, si no existe título. En última instancia, la ausencia de título requiere la redacción por parte del catalogador de un título facticio. En todos estos casos el título se indicará entre corchetes '[]'.

La Designación general de la clase de documento, subcampo \$h, [R.C., 11.1.4] es '[Archivo de ordenador]'. Algunos autores del mundo anglosajón (Hsieh-Yee, 1996) han propuesto para sus reglas de catalogación la posibilidad de indicar otra Designación General de la Clase de Documento propia para los archivos de ordenador remoto, de esta manera los usuarios podrían utilizar esta nueva Designación para dirigir sus búsquedas exclusivamente hacia los recursos de Internet. Se ubica como en el resto de los materiales inmediatamente después del título propio, subcampos \$a [R.C., 11.1.3] (título propio), \$n (número de la parte o sección de la obra) y \$p (nombre de la parte o sección de la obra).

Los títulos paralelos [R.C., 11.1.5] e informaciones complementarias sobre el título [R.C., 11.1.6], subcampo \$b, que se encuentren en la fuente de información se transcriben tras un signo igual '=' inmediatamente después de la Designación General de la Clase de Documento.

Las menciones de responsabilidad [R.C., 11.1.7], subcampo \$c, se indican, como en el resto de los materiales si figuran en fuente principal de información, si no ocurre así se omite ya que no es imprescindible y las informaciones que tengan que ver la responsabilidad no incluidas en el campo 245 pueden incluirse en una nota si se requieren entradas secundarias de las personas o entidades responsables. Las menciones de responsabilidad se transcriben tal y como figuran en la fuente. Un caso especial es el referido al administrador del servidor WWW Webmaster-, éste tiene una actuación similar a la de un editor literario, desde esta perspectiva debería figura entre las menciones de responsabilidad del área 1.

5.2. El área 2: Edición, campo 250 [R.C., 11.2]

Este área puede omitirse si no existiese información disponible sobre la edición ya que numerosos recursos de Internet no incluyen declaraciones de edición formales. Si figuran tienden a aparecer en la primera pantalla o la última.

Un aspecto peculiar de la documentación de Internet es su fácil actualización, estas actualizaciones no pueden, en manera alguna ser confundidas con una mención de edición. Si cada mención de actualización fuese tratada como una mención de edición, cada nueva versión exigiría un nuevo registro bibliográfico. La confusión resultante de tal práctica derrotaría el propósito de la catalogación. Tal información puede estar indicada en una nota general, campo 500.

Sin embargo, las reglas de catalogación españolas especifican con claridad que si existe duda sobre la existencia o no de una edición, bastará con que figure la mención de la palabra edición para que de hecho ésta exista.

Puede incluir otros elementos propios del área 2: menciones de edición paralelas [R.C., 1.2.4], menciones de responsabilidad sobre la edición [R.C., 1.2.5], subcampo \$b.

5.3. El área 3

La única dificultad estriba en que debido a las características de los documentos existentes en Internet podemos tener que utilizar dos áreas tres como ocurre con las revistas electrónicas.

5.3.1. Área de la clase y extensión del archivo [R.C., 11.3], campo 256, características del archivo

No difiere notablemente de lo establecido para la descripción de otro tipo de archivos de ordenador, especificando la clase de archivo. Para los recursos de Internet que pueden ser de interés a los efectos de una biblioteca, es decir documentos textuales, la mención que normalmente se habrá de incluir en este área será 'datos'.

5.3.2. Área de numeración [R.C., 12.3], campo 362, mención alfabética, numérica y/o cronológica del primer y último fascículo

No existen diferencias con la utilización de este campo en la catalogación de publicaciones seriadas, campo 362. Podrán omitirse los datos correspondientes al primer y último número publicados que resulten desconocidos.

5.4. El área 4: publicación, producción, distribución, etc. [R.C., 11.4], campo 260

El lugar de edición [R.C., 11.4.3], subcampo \$a, y el nombre del editor [R.C., 11.4.4], subcampo \$b, pueden ser identificado en Internet con la sede de la entidad que sustenta el servidor en el que se sitúa el documento y el nombre de esa entidad. La exposición razonada para este tratamiento es que se provee de un foro para el autor presentar sus ideas, es decir, vendría a ser el paralelo más claro con el editor. No cabe identificar al administrador de un servidor WWW con el editor, ya que la función de éste se aproxima más a la de un editor literario. En caso de no encontrarse ningún dato pueden sustituirse por las abreviaturas [S.l.] y [s.n.].

Por lo que respecta a la fecha de publicación [R.C., 11.4.6], subcampo \$c, ya que muchos recursos de Internet son actualizados frecuentemente, la fecha de actualización es inestable y no debería utilizarse ésta como la fecha de publicación. Se debería aplicar un tratamiento similar al que se da a las publicaciones de hojas sueltas para ser actualizadas (Hsieh-Yee, 1995). Se podrían crear asientos abiertos con la fecha inicial del documento, si se conoce, seguida de una raya para indicar que está en curso de publicación. Si la fecha inicial no figura en la fuente, debería indicarse entre corchetes.

5.5. El área 5: Descripción física

Debido a la ausencia de un soporte físico, este área se ha omitido en la última revisión de las reglas españolas (p. 320, nota 1) cuando se catalogan archivos de ordenador de acceso remoto. Sin embargo es un campo obligatorio en el formato Ibermarc.

5.6. El área 6 [R.C., 11.6]: Serie, campo 4xx

Se registran los elementos referidos a la serie y subserie [R.C., 1.6.3 a 1.6.6 y 1.6.9 a 1.6.12], subcampo \$a, el referido a la numeración de la serie o de la subserie [R.C., 1.6.8 y 1.6.14], subcampo \$v, y el referido al número normalizado de la serie o de la subserie [R.C., 1.6.7 y 1.6.13], subcampo \$x.

Esta información es infrecuente en los recursos de Internet. Si bien es cierto que en la red se pueden encontrar numerosas publicaciones seriadas, las descripciones de las partes componentes de éstas darían lugar a asientos analíticos.

5.7. El área 7: Notas [R.C., 11.7], campo 5xx

El área 7 provee información valiosa que complementa información registrada en las primeras 6 áreas del registro y que no ha podido incluirse en ninguna de ellas. La mayoría de las notas que se indican a continuación son optativas.

5.7.1. Nota general, campo 500

Este campo permite introducir información para la que no se ha previsto un campo de nota especializado.

1. *Dirección electrónica* : Un dato transcendente de un archivo de ordenador en Internet para la que no ha sido previsto un campo especial es la dirección electrónica, por esta causa, este dato debería incluirse en este campo. En otros formatos -USMARC- se ha previsto un campo especial -el campo 856, ubicación y acceso electrónico- que no solo contiene esta información sino que permite además el acceso directo al documento, lo que permite el acceso al documento original desde el propio OPAC.
2. *Fuentes del título propiamente dicho* [R.C., 11.7.4. A c)]: Permite indicar la fuente de donde se ha tomado el título cuando no procede del archivo de ordenador.
3. *Nota de variante de título* [R.C., 11.7.4. A d)] Se puede indicar aquí el título de HTML u otros títulos o variantes de título obtenidos de la propia fuente.
4. *Nota sobre las menciones de responsabilidad no especificadas en el área uno* [R.C., 11.7.4. A e)]: Permiten indicar las personas o entidades que no fueron transcritas como menciones de responsabilidad pero con-

siderados importantes para identificación del documento.

5. *Nota de descripción* basada en [R.C., 12.7.3 H]: Es una nota propia de publicaciones seriadas y se ha de utilizar de manera similar. Esta nota ayudará a los usuarios a entender cuando ha sido creado el registro bibliográfico y explica cualquier discrepancia entre el registro bibliográfico y un recurso de Internet.
6. *Indicación de actualización*: Sirve para indicar las actualizaciones posteriores del recurso lo que ayuda al usuarios a entender los cambios.
7. *Periodicidad* [R.C. 12.7.3. A], campo 315 y 008 / 18 : Se trata de una nota obligatoria en las publicaciones seriadas. Permite indicar la periodicidad con la que aparece la publicación. No se descarta como periodicidad posible la mención 'irregular'. No se considera publicación periódica la que supera la periodicidad anual.
8. *Notas sobre la lengua y escritura* [R.C., 11.7.4 A b)], campo 008 / 35-37: Informa sobre la lengua y escritura del contenido del archivo de ordenador.

5.7.2. *Nota de "con", campo 501 [R.C., 11.7.4 B f)]*

Esta nota permite incluir los títulos del resto de las partes cuando se describe una de ellas con título independiente de un archivo con varias partes sin título colectivo. No se utilizará esta nota si no se describen partes individualmente.

5.7.3. *Nota de tesis, campo 502 [R.C., 1.7.4 B a)]*

Permite indicar si el archivo contiene una tesis y en que centro universitario fue defendida.

5.7.4. *Nota de edición e historia bibliográfica, campo 503 [R.C., 1.7.4 A b)]*

Especifica la relación del documento que se describe con otros documentos.

5.7.5. *Nota de bibliografía, campo 504 [R.C., 1.7.4 B d)]*

Informa sobre las referencias bibliográficas que incluye el archivo de ordenador.

5.7.6. *Nota de contenido, campo 505 [R.C., 1.7.4 B d)]*

Permite incluir un listado de las partes que contiene un archivo de ordenador. Se aplica en el caso de colecciones.

5.7.7. *Nota de acceso restringido, campo 506 [R.C., 11.7.4. B b)]*

Es una nota que permite indicar si el uso del material está restringido a determinados grupos de personas o se prohíbe su uso por un período de tiempo deter-

minado. Si se tiene en cuenta el tipo de recursos que se seleccionan en las bibliotecas de investigación, será de aplicación en escasas ocasiones.

5.7.8. Nota de citas o referencias bibliográficas, campo 510

Esta nota permite indicar los documentos en los que el archivo de ordenador ha sido citado.

5.7.9. Nota de peculiaridades de la numeración, campo 515 [R.C., 12.7.3 D]

Se trata de una nota que frecuentemente aparece en publicaciones seriadas. Como en otros casos son de aplicación directa en las revistas electrónicas.

5.7.10. Nota de tipo de archivo o de datos, campo 516

En este campo pueden incluirse los datos sobre la naturaleza y alcance del archivo [R.C., 11.7.4. A a)]. Esta nota debe redactarse cuando no se desprende del resto de la descripción.

5.7.11. Nota de sumario, resumen, etc., campo 520 [R.C., 11.7.4. B e)]

Se incluye aquí un sumario breve del contenido del archivo si no se deduce del resto de la descripción.

5.7.12. Nota de tipo de usuario al que se dirige la publicación, campo 521, [R.C., 11.7.4. B a)]

Se indica en este campo el público al que está destinado el archivo. Se aplicará en raras ocasiones, debido las características de los materiales seleccionados en bibliotecas de investigación.

5.7.13. Nota de ámbito geográfico, campo 522

Se utiliza para indicar el ámbito geográfico del documento.

5.7.14. Nota sobre el período de tiempo del contenido, campo 523 [R.C., 11.7.4 A f)]

Indica la información sobre el período de tiempo cubierto por el contenido de un archivo de ordenador.

5.7.15. Nota de formato físico adicional disponible. campo 530 [R.C., 11.7.4. B d)]

Permite indicar otros formatos físicos en los que se encuentra disponible el documento.

5.7.16. Nota sobre la obra original, campo 534

En principio esta nota no es aplicable a archivos de ordenador. No obstante no son extraños los documentos de Internet que reproducen obras originales.

5.7.17. Nota de patrocinador, campo 536 [R.C., 1.1.7. J b)]

Permite indicar la entidad patrocinadora de un documento o de su publicación electrónica.

5.7.18. Nota de fuente de los datos, campo 537 [R.C., 11.7.4 A f)]

Contiene información relativa a la fuente de los datos.

5.7.19. Nota de datos técnicos, campo 538 [R.C., 11.7.4 A i)]

Incluye los datos técnicos sobre el archivo.

5.7.20. Nota de índices acumulativos u otros instrumentos de localización, campo 555 [R.C., 12.7.3 C n)]

Este campo sirve para indicar la existencia de índices acumulativos en las publicaciones seriadas, en consecuencia se utiliza en las publicaciones electrónicas seriadas.

5.7.21. Nota de información sobre documentación, campo 556

Normalmente no es de aplicación en los documentos de acceso remoto que no aportan documentación, y por la misma causa tampoco es de aplicación el campo 581. Sin embargo, si existiesen archivos separados en los que se incluyese documentación -del tipo léame- su existencia se expresará en esta nota.

5.7.22. Nota metodológica, campo 567

Contiene información relativa a las características metodológicas significativas del material.

5.7.23. Nota de responsable personal, campo 570 [R.C., 12.1.7 C]

Este campo se aplica a publicaciones seriadas. Permite indicar información sobre determinados responsables personales que se consideran de importancia para la descripción.

5.7.24. Nota de relación compleja con otros documentos, campo 580

Permite indicar una relación compleja entre el documento descrito y otros documentos

5.7.25. Nota de archivos relacionados, campo 582

En esta nota se expresa la existencia de archivos de ordenador relacionados.

5.8. El área 8: Número normalizado, campo 022

Se registran en este área los numeros normalizados, ISBNs e ISSNn que figuran el documento [R.C. 11.8.3 A]. Siguiendo la normativa [11.8.3. B] podrían registrarse los URNs -Nombres Uniforme de Recurso- cuando estén disponibles.

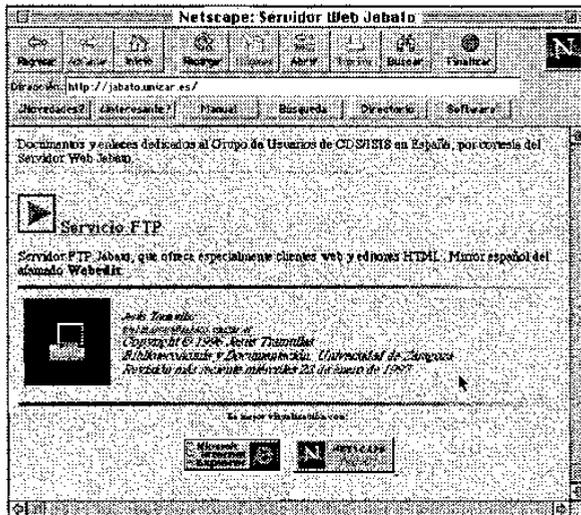
5.9. Puntos de acceso y formulación de encabezamientos, campos 080, 1xx, 6xx, 7xx, 80x-83x [R.C.,14-16]

La normativa sobre puntos de acceso no difiere notablemente de los establecido para el resto de los materiales ni en lo referente a la elección de puntos de acceso ni a la formulación de encabezamientos. De hecho, no se han establecido normas especiales para los archivos de ordenador sobre estos temas. Las herramientas de tratamiento de los puntos de acceso de materias pueden ser los mismos que los utilizados para el resto de los materiales. Es esencial que todos los tipos de puntos de acceso estén sometidos a control de autoridades.

6. Ejemplos.

Se tratan, a modo de ejemplo, tres recursos pertenecientes a cada uno de los tipos más frecuentes en Internet.

6.1 Tratamiento de un sistema o servicio: Servidor web Jabato.





6.1.1. Formato Ibermarc.

```

008 970218m199599999espn000f000m000000000spa0d
019 00Z 3397-96
080 00$a002
080 00$a902
245 00$aServidor web Jabato $h[Archivo de ordenador]
/$c[administrador] Jesús Tramullas.
256 00$aDatos y programas
260 00$aZaragoza :$bUniversidad de Zaragoza,
Biblioteconomía y Documentación $c1995.
500 00$aUrl: http://jabato.unizar.es
505 00$aContiene: Servicio Retiarius. Software para
Internet. Ciencias de la Información y la documenta-
ción. Proyecto San Telmo. Proyecto Aureus.
516 00$aProgramas, Datos gráficos y texto.
600 14$aDocumentación.
600 24$aInternet.
    
```

600 24\$aArqueología.

700 10\$aTramullas, Jesús.

710 21\$aUniversidad de Zaragoza\$bÁrea de
Biblioteconomía y Documentación.

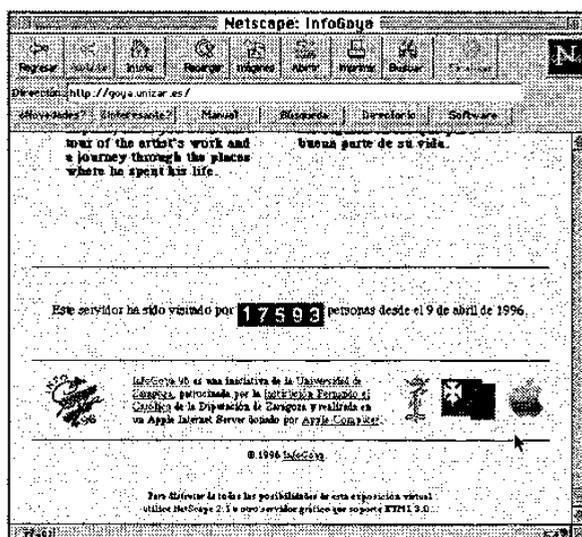
6.1.2. *Formato en pantalla según ISBD.*

SERVIDOR web Jabato [Archivo de ordenador] / [administrador] Jesús Tramullas. - Datos y programas. - Zaragoza : Universidad de Zaragoza, Biblioteconomía y Documentación, 1995.

Url: <http://jabato.unizar.es>. - Contiene: Servicio Retiarius. Software para Internet. Ciencias de la Información y la documentación. Proyecto San Telmo. Proyecto Aureus. - Actualización: 10 de febrero de 1997.

1. Documentación. 2. Internet. 3. Arqueología.
I. Tramullas, Jesús. II. Universidad de Zaragoza.
Área de Biblioteconomía y Documentación.

6.2 Tratamiento de un recurso monográfico: INFOGOYA'96.



6.2.1. *Formato Ibermarc.*

008 970218c199600000espn0000g0000m0000000000spa0d

041 10\$aspaeng

080 00\$a75Goya, Francisco de.

245 00\$aExposición virtual InfoGoya'96 \$h[Archivo de ordenador] :\$bVirtual exhibition InfoGoya'96 /\$ces una iniciativa de la Universidad de Zaragoza ; patrocinada por la Institución Fernando el Católico.

256 00\$aDatos

260 00\$aZaragoza :\$bUniversidad de Zaragoza,\$c1996.

500 00\$aUrl: <http://goya.unizar.es>

516 00\$aDatos gráficos y texto.

600 11\$aGoya, Francisco de\$jexposiciones.

710 21\$aUniversidad de Zaragoza.

710 21\$aInstitución Fernando el Católico (Zaragoza)

6.2.2. *Formato en pantalla según ISBD.*

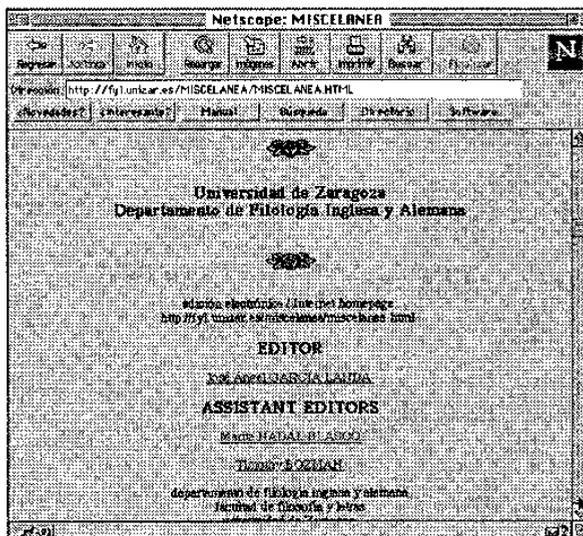
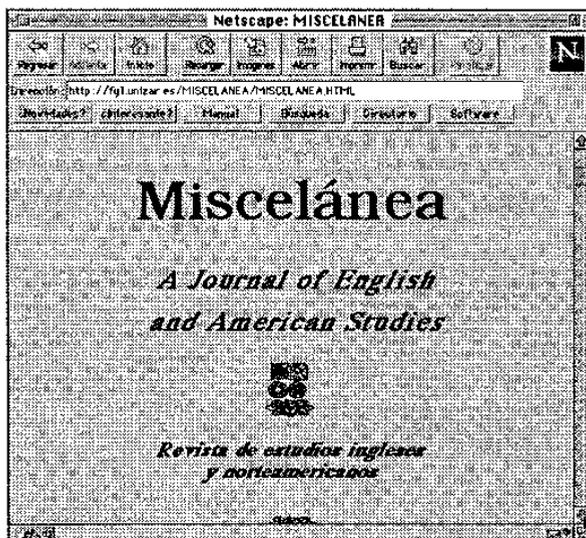
EXPOSICION virtual InfoGoya'96 [Archivo de ordenador] = Virtual exhibition InfoGoya'96 / es una iniciativa de la Universidad de Zaragoza ; patrocinada por la Institución Fernando el Católico. - Datos. - Zaragoza : Universidad de Zaragoza, 1996.

Url: <http://goya.unizar.es>

1. Goya, Francisco de - Exposiciones. I. Universidad de Zaragoza. II. Institución Fernando el Católico (Zaragoza).

75Goya, Francisco de.

6.3. Tratamiento de recurso seriado: Miscelánea.



6.3.1. Formato Ibermarc.

008 970218c19uu9999espar00f000t00000000eng0d

080 00\$a811.111

080 00\$a821.111

245 00\$aMiscelánea \$h[Archivo de ordenador] :\$ba
journal of english and american studies = revista de
estudios ingleses y norteamericanos /\$cUniversidad de
Zaragoza, Departamento de Filología Inglesa y Alemana.

256 00\$aDatos

260 00\$aZaragoza :\$bUniversidad de Zaragoza,
Departamento de Filología Inglesa y Alemana,\$c[1979?]-

315 00\$aAnnual.

500 00\$aUrl: <http://fyl.unizar.es/MISCELANEA/MISCELANEA.HTML>

500 00\$aDescripción basada en: n.17 (1996).

503 00\$aEdición electrónica de la publicación del
mismo título.

516 00\$aTexto.

600 14\$aLengua inglesa\$jrevistas.

600 14\$aLiteratura inglesa\$jrevistas.

710 21\$aUniversidad de Zaragoza\$bDepartamento de
Filología Inglesa y Alemana.

6.3.2. Formato en pantalla según ISBD.

MISCELANEA [Archivo de ordenador] : a journal of english and american studies = revista de estudios ingleses y norteamericanos / Universidad de Zaragoza, Departamento de Filología Inglesa y Alemana. - Datos. - . - Zaragoza : Universidad de Zaragoza, Departamento de Filología Inglesa y Alemana, [1979?]- .

Url: <http://fyl.unizar.es/MISCELANEA/MISCELANEA.HTML>. - Anual. - Descripción basada en: n.17 (1996). - Edición electrónica de la publicación seriada del mismo título.

1. Lengua inglesa.- Revistas. 2. Literatura inglesa - Revistas. I. Universidad de Zaragoza. Departamento de Filología Inglesa y Alemana.

811.111

7. Referencias

- Anderson, Bill ; Hawkins, Les (1996). Development of CONSER Cataloging Policies for Remote Access Computer File Serials // The Public-Access Computer Systems Review 7 (1996), n. 1. URL : <http://info.lib.uh.edu/pacsrev.html>.
- Beck, Melissa, et al. (1995). Module 31, Remote Access Computer File Serials. // CONSER Cataloging Manual. Washington: Library of Congress, 1995. URL: <http://www.library.vanderbilt.edu/ercelawn/eserials.html>
- Butterfield, Kevin (1996). Catalogers and the Creation of Metadata Systems : A Collaborative Vision at the University of Michigan. // InterCat : A Catalog of Internet Resources : proceedings of the OCLC Internet Cataloging Colloquium : San Antonio, Texas, January 19, 1996. URL : <http://www.oclc.org/oclc/man/colloq/butter.htm>
- Caplan, Priscilla (1995). "You Call It Corn, We Call It Syntax-Independent Metadata for Document-Like Objects". // The Public-Access Computer Systems Review 6 (1995), n. 4. URL : <http://www.nlc-bnc.ca/documents/libraries/cataloging/caplan3.txt>
- Daniel, Ron (1994). Proposed URC External Representation. URL : <http://www.acl.lanl.gov/URI/ExtRep/urc0.html>
- Formato Ibermarc para registros bibliográficos (1995). 5ª ed. Madrid: Biblioteca

- Nacional, 1995. 2 vols.
- Frías, José Antonio (1996). La catalogación de los documentos electrónicos : problemática, propuestas e iniciativas. // Jornadas españolas de documentación automatizada (5a. 1996. Cáceres). Sistemas de información : balance de 12 años de jornadas y perspectivas de futuro: actas. Cáceres : FESABID, 1996, vol. 1, p. 431-438.
- Hsieh-Yee, Ingrid (1996). Modifying Cataloging Practice and OCLC Infrastructure for Effective Organization of Internet Resources. // InterCat : A Catalog of Internet Resources : proceedings of the OCLC Internet Cataloging Colloquium : San Antonio, Texas, January 19, 1996. URL : <http://www.oclc.org/oclc/man/colloq/hsieh.htm>
- InterCat (1996): A Catalog of Internet Resources : proceedings of the OCLC Internet Cataloging Colloquium : San Antonio, Texas, January 19, 1996. URL: <http://www.oclc.org/oclc/man/colloq/toc.htm>
- Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas (1994). ISBD(CF) : Descripción Bibliográfica Internacional Normalizada para Archivos de Ordenador. Madrid : ANABAD : Arco/Libros, 1994.
- Koster, Martijn (1995). Robots in the Web: Threat Or Treat. // ConneXions, 9 (1995), n. 4. URL : <http://info.webcrawler.com/mak/projects/robots/threat-or-treat.html>
- López Alonso, Miguel A. (1996). El futuro de la identificación de la información en Internet. // Jornadas españolas de documentación automatizada (5a. 1996. Cáceres). Sistemas de información : balance de 12 años de jornadas y perspectivas de futuro: actas. Cáceres : FESABID, 1996, vol. 1, p. 513-518.
- Morgan, Eric Lease (1995). Adding Internet resources to our OPACs. // Serials Review 21 (1995), n. 4. URL : <http://www.lib.ncsu.edu/staff/morgan/adding-internet-resources.html>
- Olson, Nancy B. (eds.) (1995). Cataloging Internet Resources : A Manual and Practical Guide, ed. Nancy B. Olson. Dublin: OCLC, 1995. URL: <http://www.oclc.org/oclc/man/9256cat/toc.htm>
- Reglas de catalogación españolas (1995). Ed. refundida y rev., reimp con corr. Madrid: Dirección General del Libro, Archivos y Bibliotecas, 1995.
- Sha, Vianne T. ; Patrick, Timothy B. ; Kochtanek, Thomas R. (1996). The Traditional Library and the National Information Infrastructure. // InterCat : A Catalog of Internet Resources : proceedings of the OCLC Internet Cataloging Colloquium : San Antonio, Texas, January 19, 1996. URL : <http://www.oclc.org/oclc/man/colloq/sha.htm>
- Shieh, Jackie (1996). Does It Really Matter?: The Choice of Format, Order of Note Fields, and Specifics of 856 // InterCat: A Catalog of Internet Resources : proceedings of the OCLC Internet Cataloging Colloquium : San Antonio, Texas, January 19, 1996.
- Taylor, Arlene G. ; Clemson, Prattice (1996). Access to Networked Documents: Catalogs? Search Engines? Both? // InterCat: A Catalog of Internet Resources : proceedings of the OCLC Internet Cataloging Colloquium : San Antonio, Texas, January 19, 1996. URL : <http://www.oclc.org/oclc/man/colloq/taylor.htm>
- Tramullas Saz, Jesús (1996). Servidor web Jabato : un servidor web para la formación en ciencias de la información. El servicio Retiarus. // Jornadas españolas de documentación automatizada (5a. 1996. Cáceres). Sistemas de información : balance de 12

- años de jornadas y perspectivas de futuro : actas. Cáceres: FESABID, 1996, vol. 1, p. 295-303.
- Vaquero Pulido, Raúl ; García Figuerola, Carlos (1996). Motores de búsqueda en Internet. // Jornadas españolas de documentación automatizada (5a. 1996. Cáceres). Sistemas de información : balance de 12 años de jornadas y perspectivas de futuro : actas. Cáceres : FESABID, 1996, vol. 2, p. 621-629.
- Weibel, Stuart L. (1995). Metadata: The Foundations of Resource Description. // D-Lib Magazine, 1995 7. URL : <http://www.dlib.org/dlib/July95/07weibel.html>
- Xu, Amanda (1996). Access Information on the Internet : A Feasibility Study of MARC Formats and AACR2. // InterCat: A Catalog of Internet Resources : proceedings of the OCLC Internet Cataloging Colloquium : San Antonio, Texas, January 19, 1996. URL: <http://www.oclc.org/oclc/man/colloq/xu.h>

