

VII Jornadas Españolas de Documentación  
Bilbao 19-20-21 octubre 2000 - Palacio Euskalduna

**La Gestión del Conocimiento:  
retos y soluciones de los  
profesionales de la información**

# **La Gestión del Conocimiento: retos y soluciones de los profesionales de la información**

eman ta zabal zazu



Universidad  
del País Vasco  
servicio editorial

Euskal Herriko  
Unibertsitatea  
argitalpen zerbitzua

# Índice

Presentación .....	11
Comités .....	13
<b>1. Nuevos perfiles profesionales de la información y el conocimiento .....</b>	<b>17</b>
«Knowledge Management - An Evolving Professional Competence» <i>Dominic Kelleher</i> .....	19
«La gestión del proceso de encargo y creación de materiales didácticos en la intranet corporativa de la Universitat Oberta de Catalunya» <i>Albert Boada Abad; Carles Ramírez Sarrado</i> .....	21
«La Estadística como herramienta fundamental para el profesional de la información en el entorno empresarial y Centros de Investigación» <i>José A. Salvador Oliván</i> .....	29
«Informar en el archivo: características y recursos del servicio de referencia archivístico» <i>Diego Navarro Bonilla</i> .....	39
«Outsourcing documental: organización de futuro» <i>Pablo Lara Navarra; José Ángel Martínez Usero</i> .....	49
«El papel de los profesionales en el proceso de cambio: ser los agentes del cambio» <i>Carina Rey Martín</i> .....	57
«Ecología de la Información: ¿Nuevas Competencias, Nueva Ideología Profesional?» <i>Montserrat Sebastià i Salat</i> .....	69
«Cómo un Servicio de Documentación responde al nuevo usuario (Investigador, Auto-didacta, Autosuficiente y Experto): La experiencia de la Universitat de les Illes Balears a partir del acceso a las bases de datos de DialogWeb» <i>Daisy Chiang Martín; Tomàs Oliver Colom</i> .....	79
«Las nuevas competencias del profesional de la información: del control de documentos a la gestión del conocimiento» <i>Antonio Hernández Pérez; David Rodríguez Mateos</i> .....	89
«Dos vertientes de un mismo camino: una experiencia corporativa en gestión del conocimiento» <i>Yolanda Gonzalo Balmisa</i> .....	99

<b>2. Gestión y representación de la información</b> .....	103
«Las tecnologías de información y los desafíos de gestión y marketing en las bibliotecas latinoamericanas» <i>Lourdes Feria</i> .....	105
«Estructuración de la información mediante XML: un nuevo reto para la gestión documental» <i>Bonifacio Martín Galán; David Rodríguez Mateos</i> .....	113
«Unidad documental en los servicios de documentación de las televisiones» <i>Felipe Zapico Alonso; Jorge Caldera Serrano</i> .....	125
«Parámetros e indicadores de calidad para la evaluación de recursos digitales» <i>Lluís Codina</i> .....	135
«Normalización de referencias bibliográficas de documentos electrónicos» <i>M.<sup>a</sup> Dolores Borgoños Martínez</i> .....	145
«El uso de metainformación en los webs de las bibliotecas españolas» <i>José Antonio Merlo Vega; Ángela Sorti Rojo</i> .....	155
«Concepción, construcción y mantenimiento de una herramienta de indización para un centro de documentación sobre Gestión de la Innovación y el Conocimiento: una experiencia» <i>Mayte López Ferrer</i> .....	165
«El acceso a la información en soporte digital en la Fundación Sancho El Sabio (Vitoria-Gasteiz)» <i>Jesús Manuel Zubiaga; M.<sup>a</sup> Rosario Martínez</i> .....	173
«La organización hipertextual de textos legislativos con HTML y XML: una necesidad y las soluciones de presente y futuro» <i>J. Tomás Nogales; M. Carmen Arellano</i> .....	179
«Gestión documental en las empresas: una aproximación práctica» <i>Carlota Bustelo Ruesta</i> .....	189
«La implementación de metadatos y Dublin Core en sedes y páginas web de bibliotecas y centros de documentación de universidades y centros de investigación de la Red IRIS» <i>Francisco Javier Vidal Bordés; José Antonio Salvador Oliván</i> .....	197
«Metadatos y Tesoros: aplicación de XML/RDF a los sistemas de organización del conocimiento en Intranets» <i>Eva M.<sup>a</sup> Méndez Rodríguez</i> .....	211
«Localización, identificación y descripción de documentos web: tentativas hacia la normalización» <i>Eva M.<sup>a</sup> Méndez Rodríguez; José Antonio Merlo Vega</i> .....	221
«Indicadores hacia una evaluación no objetiva (cuantitativa) de sedes web» <i>Isidro F. Aguillo</i> .....	233
«Internet invisible o Infranet: Definición, clasificación y evaluación» <i>Isidro F. Aguillo</i> .....	249
«El documento electrónico: protección y seguridad de datos en la World Wide Web» <i>Elena Pérez Moreno; Mónica Lorience García</i> .....	271
«La supervivencia de los servicios de indización y resúmenes con la explosión de Internet» <i>Miguel Ángel López Alonso</i> .....	281
«La web de una biblioteca especializada en ciencias de la salud: criterios de calidad para recursos y servicios» <i>Ana Barbeta Márquez; M.<sup>a</sup> Jesús García Mateu; Carolina Roig Castelló</i> .....	291

<b>3. Accesibilidad y transparencia de la información</b> .....	301
«Bibliotecas en una sociedad global: todos en alianza con todos»	
<i>Sonsoles Celestino Angulo</i> .....	303
«Accesibilidad de los usuarios españoles a los fondos cinematográficos»	
<i>Susana Torrado Morales</i> .....	321
«Un programa para mejorar la calidad de las revistas científicas. Aplicación a las revistas españolas de biblioteconomía, documentación y archivística»	
<i>Elena Giménez Toledo; Luis Rodríguez Yunta; Adelaida Román Román</i> .....	331
«La accesibilidad y adaptabilidad de las bibliotecas para usuarios que se desplazan en silla de ruedas: hacia un mundo sin barreras arquitectónicas»	
<i>Pau Cazorla Cabeza; Josep Manuel Cazorla Cabeza</i> .....	343
«El acceso a la documentación de la Unión Europea en la Sociedad de la información: Indicadores de calidad y cartas de servicios. Una aproximación crítica»	
<i>Emilio Sánchez Blanco; Marina Vianello Osti</i> .....	355
«Estructura y accesibilidad de la información institucional en las páginas Web de las Comunidades Autónomas españolas: análisis y caso práctico»	
<i>Marina Vianello Osti; Emilio Sánchez Blanco</i> .....	369
«Comunidades virtuales especializadas: Un análisis comparativo de la información y servicios que ofrecen al usuario»	
<i>Elena Fernández; Isabel Fernández; Ángeles Maldonado</i> .....	383
«Información económica y empresarial: alianzas estratégicas en Internet»	
<i>Raquel Ruiz Crespo</i> .....	391
«Bibliotecas digitales españolas: análisis de las primeras realidades»	
<i>José Antonio Ontalba y Ruipérez</i> .....	399
«Productos y servicios para la formación de usuarios de bibliotecas universitarias: el uso de bases de datos en entornos Web»	
<i>Natalia Sastre Miralles</i> .....	411
«Gestión del conocimiento en ciencias sociales a través de la red: Las revistas electrónicas españolas»	
<i>Ángel Villagra; Ceferina Anta; Ana M.ª Valverde; María Ruiz Gálvez; María Espinosa</i>	419
«Colecciones digitales: análisis de usabilidad»	
<i>Ricardo Eito Brun</i> .....	435
«Implicaciones del uso de la firma electrónica en la validación de documentos»	
<i>José Luis Bonal Zazo</i> .....	445
«Métodos de evaluación de calidad de catálogos automatizados: análisis comparativo»	
<i>M.ª del Pilar Ortego de Lorenzo-Cáceres; José Luis Bonal Zazo</i> .....	453
«Los foros electrónicos como instrumentos para detectar conflictos éticos en el ámbito de la información»	
<i>Margarita Pérez Pulido; Aurora Gómez-Pantoja Fernández-Salguero</i> .....	461
«La clasificación: una estrategia de comunicación»	
<i>Maite Comalat; Constança Espelt</i> .....	473
<b>4. Gestión del conocimiento</b> .....	485
«Información y desarrollo industrial»	
<i>M.ª José Ordóñez Vergara</i> .....	487
«La gestión del capital intelectual: Nuevos parámetros de análisis para la economía de la información»	
<i>Teresa Allepuz Ros</i> .....	499

«Campo de acción de la documentación en actividades de vigilancia tecnológica» <i>Elena Giménez Toledo</i> . . . . .	507
«La aplicación de la norma ISO 9002 en el Servei de Biblioteques de la Universitat Autònoma de Barcelona» <i>Núria Balagué Mola</i> . . . . .	517
«Intranets corporativas y gestión del conocimiento: el caso de una empresa del sector de la energía» <i>José Antonio Sánchez Montero</i> . . . . .	529
«¿El conocimiento tácito es documentación? La experiencia de una empresa en la Gestión del Conocimiento» <i>Josep M. Rodríguez Rovira</i> . . . . .	539
«Los archivos en “Telépolis”: recoger, conservar y servir el documento electrónico» <i>Agustín Vivas Moreno</i> . . . . .	547
«Variables de decisión y factores clave en el proceso información - conocimiento - innovación» <i>Francisco Javier Odriozola Sagastume</i> . . . . .	559
«La Intranet del conocimiento» <i>José Manuel Fdez. de las Heras; Pako Díaz de Cerio</i> . . . . .	567
«Recortes de prensa digitales: todo un mundo de información a tu alcance» <i>Begoña Aguilera Caballero</i> . . . . .	575
«Los nuevos sistemas de innovación tecnológica en la Sociedad del Conocimiento: la gestión integrada de las tecnologías de la información en la UE» <i>M.ª Dolores Ayuso García; M.ª José Ayuso Sánchez</i> . . . . .	585
«Cooperación frente al caos en Internet: CORC, una propuesta de trabajo» <i>María Isabel Zubeldia Lauzurica</i> . . . . .	597
«Intranets para la gestión de una biblioteca: los flujos de información en un sistema de calidad ISO 9002» <i>Miquel Térmens Graells</i> . . . . .	607
«El sistema avanzado de digitalización de documentación administrativa del archivo de la Universidad de Alcalá (SADDA)» <i>Carmen de la Peña Montes de Oca; Concepción Serrano Murillo; David F. de Torres Sánchez</i> . . . . .	613
«La gestión de las redes virtuales interempresariales de intercambio de información» <i>Mercè Muntada; Blanca Viros</i> . . . . .	625
«Estrategias documentales para generar conocimiento. Representación y construcción de la realidad mediante el tesoro» <i>José Ignacio Aranes; Carmelo Landa</i> . . . . .	633
«Aportaciones del “PAKM”: Aspectos prácticos de la gestión del conocimiento» <i>Carmen M.ª Brugerolas Ros; Isabel Andreu Felipe; José Luis Cárceles Andreu; María Martín Hernández</i> . . . . .	639
Ponencia de clausura: «De la sociedad...» <i>Eduardo Bueno</i> . . . . .	647

# Localización, identificación y descripción de documentos web: tentativas hacia la normalización

Eva M.<sup>a</sup> Méndez Rodríguez

Universidad Carlos III de Madrid. Dpto. Biblioteconomía y Documentación

José Antonio Merlo Vega

Universidad de Salamanca. Dpto. Biblioteconomía y Documentación

**Resumen:** La explosión del número de usuarios en Internet y el crecimiento continuado del volumen de los contenidos accesibles a través de la World Wide Web hacen que la valiosa información electrónica alojada en la web sea inconsistente y difícil de localizar. Por este motivo se están desarrollando distintos proyectos de normalización de la información distribuida en Internet, tanto para la identificación de los recursos, como para su estructuración, acceso y referencia. En la presente comunicación se analizan estas tentativas de estandarización relativas a los sistemas de identificación y localización, a la descripción de los contenidos y a la descripción formal de los recursos web. Este trabajo pretende ofrecer una panorámica del estado de la cuestión en materia de localización, identificación y descripción de documentos web, así como de reflexionar en torno a la normalización en este sentido.

**Palabras clave:** Normalización; Internet; Información web; Localización de recursos; Descripción de recursos; Identificación de recursos.

## 1. Introducción

La documentación ha sido desde siempre una disciplina con un alto componente normativo, tanto para la identificación como para la descripción del conocimiento contenido en las diversas publicaciones. Los bibliotecarios y/o documentalistas hemos utilizado y utilizamos normas por varios motivos evidentes: para incrementar la calidad y la consistencia de la información, para mejorar la compatibilidad de estructuras de datos o para facilitar tanto la recuperación como el intercambio de información. Las normas bibliográficas tradicionales que hemos adoptado en la labor de la gestión de la información son aquellas que, al aplicarse a una colección de información almacenada, organizan y facilitan el acceso a los registros de conocimiento y a su contenido intelectual. Estos, amén de la tradición normativa generalizada, son aspectos que nos llevan a pensar en normas similares para la Red. Sin embargo, el proceso de estandarización formal, tal y como está establecido en el mundo impreso, es insuficiente para ajustarse a las necesidades y al

desarrollo acelerado de la información contenida en Internet; por ello proliferan distintas tentativas de normalización que pretenden convertirse en estándares *de facto* para la web.

La WWW no es más que un universo de información complejo, distribuido, heterogéneo, cambiante, dinámico, y poco uniforme. Todas estas características se justifican desde el punto de vista que en la web, se puede acceder a distintos tipos de información (texto, imágenes, vídeo, sonido, software, etc.) en distintas lenguas, sobre las temáticas más heterogéneas, sin diferenciar la información científica, de la comercial o de la meramente divulgativa, la imagen del sonido, lo importante de lo accesorio y siempre a través de un ordenador. Además toda esa información heteróclita se encuentra distribuida en distintos nodos por todo el mundo, sin ningún criterio documental apriorístico de organización de la información; a ello se une el aumento exponencial de la web que dificulta aún más su estructuración; día a día varía el número de páginas en Internet, los contenidos de las existentes, otras desaparecen, y toda esta mutación informacional es constante, diaria y vertiginosamente cambiante. Quizás esta falta de uniformidad de esa maraña de información web se deba a que no existe ninguna estandarización universalizada que estructure la información identificadora, ni de la forma ni del contenido de las páginas web.

Por todo ello, en esta comunicación pretendemos dar a conocer sistemáticamente las principales tentativas de estandarización que están ocupando muchos Mb de información tanto en la web como en los distintos foros de discusión relacionados con ellas. En primer lugar revisaremos las iniciativas que tratan de solucionar el problema de la inconsistencia de ubicación de los recursos web, esto es, las «normas» que están surgiendo en torno a la localización e identificación de documentos electrónicos; en segundo lugar nos ocuparemos del prolijo mundo de los metadatos a través del análisis de los distintos modelos que se están proponiendo para la estructuración y definición del contenido web; y por último, pero no por ello menos importante, reflejaremos la problemática que comporta la referencia y descripción formal de estos recursos electrónicos a través de los proyectos normativos que se están desarrollando al respecto.

## 2. La localización e identificación de documentos web

Desde antaño han existido iniciativas para identificar los documentos y poder localizarlos entre un conjunto. Bibliografías, catálogos, etc. han pretendido la identificación de los documentos, al tiempo que indicaban dónde podían encontrarse. Para ello se desarrollaron normas de descripción y catalogación, así como sistemas de acceso al documento. Para la identificación y localización de los archivos telemáticos se parte de objetivos semejantes; incluso, la necesidad es aún mayor, dada la profusión de documentos que están presentes en Internet y la facilidad con la que se incorporan páginas nuevas. Parece lógico, por tanto, que se desarrollen métodos eficaces para localizar un documento concreto en la gran «biblioteca» de



Internet. Son muchos los proyectos llevados a cabo con esta finalidad, incluso algunos se están aplicando con gran éxito y difusión. En este apartado se comentarán las principales actuaciones realizadas por organismos de relevancia para identificar a un documento web de forma única, inequívoca y persistente y para poder localizarlo de forma sencilla. A continuación se exponen los proyectos más importantes:

- *Uniform Resource Identifiers* (URI): sistema ideado por el Internet Engineering Task Force (IETF) que pretende ser un sistema global para identificar recursos en la web, ya sean documentos, imágenes, archivos de programas, correos electrónicos, etc. Es un método genérico, que combina URLs y URNs para la identificación de un documento. El grupo de trabajo de la IETF que estudia este tema publicó su último borrador sobre la sintaxis de los URIs en agosto de 1998 (RFC 2396) y en el mismo los definía como una cadena compacta de caracteres para identificar un recurso físico o abstracto. El término más significativo de esta técnica es el de «identificadores», que no es más que una secuencia de caracteres con sintaxis controlada mediante la cual se reconoce un documento de forma única.
- *Uniform Resource Name* (URN): se trata de un tentativa normativa aportada por el IETF (Documento RFC 1737) con la intención de suplir las carencias del sistema URL. Parte de la idea de conseguir un método global de identificación de recursos, que a la vez sea único y persistente, que pueda crecer a medida que sea necesario y que no incumpla la legislación en materia de tratamiento de datos, además de otras características. Su sintaxis se estructura en tres bloques separados por dos puntos: el identificador URN, el NID o nombre de la categoría en la que se incluye el documento que se pretende identificar y el NSS o cadena específica. Por ejemplo, un documento web se identificaría así: urn:inet:dstc.edu.au:tr0088, donde «urn» es el código que indica el tipo de localizador empleado, «inet» explica que se trata de un documento de Internet y «dstc.edu.au:tr0088» es la ruta y el documento en cuestión. Una característica importante de este sistema es que trabaja a la par del *Uniform Resource Characteristics* (URC), esquema para la descripción de metadatos.
- *Uniform Resource Locator* (URL): es el sistema de localización de documentos común en la World Wide Web, ya que su aparición coincidió con la explosión del web. En realidad no consiste en establecer nombre para los objetos, sino en indicar la forma de acceso a los mismos. Este método indica, en primer lugar, el protocolo mediante el cual se localizará el documento; a continuación, se informa del servidor en el que está alojado, para lo que se sirve del Sistema de Nombre de Dominios. Además, un URL explica el directorio o directorios en los que se encuentra el documento, indicando en último lugar el nombre y extensión del archivo que se pretende identificar y/o localizar. Dada la inconsistencia de este forma de identificación se prevé que sea sustituida por el sistema URN, aunque para ello es necesario que los navegadores soporten esta nueva forma de localizar la información.

- *Persistent Uniform Resource Locator* (PURL): método desarrollado por OCLC partiendo de los mismos objetivos que el sistema URN. De hecho OCLC está trabajando también en el establecimiento del sistema URN. Partiendo del protocolo HTTP realiza una base de datos de URLs, en la que se identifican los nombres de los documentos y los servidores que los albergan. En el momento en que un documento cambie de dirección bastará con indicarlo en la base de datos, de tal manera que se produzca el redireccionamiento de forma automática. Formalmente un PURL es un URL, ya que su estructura es idéntica. En realidad se trata de un intermediario entre la dirección antigua de un documento web y su nueva dirección. A diferencia del sistema URN, éste ya está operativo en el servicio OCLC PURL.
- *Digital Object Identifier* (DOI): una docena de editores están probando este método de identificación de objetos digitales. No se trata de un sistema como URN o PURL, aunque coincide con estas iniciativas en que también persigue la identificación única y persistente; en realidad se trata de una manera de intercambiar información, sobre todo con fines comerciales. Originariamente fue desarrollado por la *American Association of Publishers* y la *Corporation for National Research Initiatives*, pero en la actualidad se responsabiliza de su mantenimiento la *International DOI Federation*. Se estructura en un prefijo, en el que se identifica el directorio concreto en el que se encuentra un documento, y un sufijo, en el que se expresa el código del objeto digital en cuestión.
- Otras iniciativas: aparte de las experiencias presentadas, existen otra serie de proyectos en marcha también relacionados con la identificación de recursos web; entre ellos destacan WEBDAV, prototipo en el que están trabajando empresas multinacionales de gran prestigio en el mundo de la informática y la telecomunicaciones y que está siendo desarrollado por el IETF. Del mismo modo, son de interés sistemas como *Human Friendly Names*, *Serial Item and Contribution Identifier* (SICI) que es desde 1996 una norma ANSI Z39.56, *Book Item and Contribution Identifier* (BICI), *Publishers Item Identifier* (PII), *Real Names o Handle System*, entre otros. Muchos de estos sistemas no se centran en la identificación de documentos web, aunque su lógica de funcionamiento permitiría que también fueran aplicados con este fin.

### 3. La descripción de contenidos web

Hablar de la descripción de contenidos web, implica hablar del «tan de moda» y complejo mundo de los metadatos o metainformación. Los recursos electrónicos pueden y deben ser una parte integral de las colecciones de las bibliotecas modernas, por ello, establecer sistemas robustos de acceso a esa información electrónica tiene una importancia crítica en el desarrollo de las incipientes «bibliotecas digitales». Desde hace aproximadamente cinco años, diversos colectivos están desarro-

lizando modelos o pseudo-estándares de metadatos, con distintos niveles de difusión, para describir sus colecciones digitales, otros están esperando a la madurez de estos formatos para implemetarlos de una forma más fácil y sólida en sus sistemas de información, otros servicios tratan de solucionar la incorporación de documentos web a través del campo 856 del formato USMARC.

A una definición tan sencilla y manida como «datos sobre los datos», podemos añadir que los metadatos en el contexto de Internet, son informaciones estructuradas sobre la información distribuida, todos aquellos datos comprensibles por el ordenador, estructurados de tal forma que sirven para localizar, identificar y describir el contenido de un documento web, y por ende, para recuperar información en la Red. No obstante, los diversos modelos de metadatos no son normas en el amplio sentido de la palabra<sup>1</sup> responden más bien a proyectos de desarrollo de normas o a estándares *de facto* en determinadas comunidades de usuarios. Analizaremos a continuación algunos de los más importantes modelos en este sentido.

- *Dublin Core* (DC): es el formato de metainformación más divulgado y más citado a nivel general, al menos en el ámbito estrictamente bibliotecario. Desde sus comienzos (1995), el DC fue diseñado para promover un estándar de propósito general, sencillo y descriptivo de los recursos web de cualquier materia. En la propuesta original el DC facilitaría una indización consistente sin la necesidad de una gran experiencia catalogadora, poco a poco (a través de sus diferentes workshop donde se ha desarrollado su evolución) se entendía como una herramienta de estructuración de la información para facilitar la interoperabilidad semántica en la web<sup>2</sup>. Sintéticamente el DC es un conjunto de quince elementos, opcionales y repetibles, para describir el contenido web según tres tipos de información: siete elementos, que podríamos denominar estrictamente de contenido (title, subject, description, source, language, relation, coverage); elementos con información relativa a la propiedad intelectual del recurso (creator, publisher, contributor, rights); y elementos relativos a la temporalidad y formato del documento así como su identificación (data, type, format, identifier). La especial adaptabilidad de este modelo es un factor atractivo para la descripción/»catalogación» de

---

<sup>1</sup> Sin embargo, Altavista <<http://www.altavista.com>> tiene 6,920 páginas indizadas bajo la descripción de «metadata standards». Esto nos hace una idea de la vocación normativa que se le presume a estos modelos de metadatos, sin embargo no abundan normas internacionales al respecto. En este sentido destacan los trabajos de la ISO/IEC JTCT1/SC32 para la normalización de los elementos de datos para facilitar la transparencia y la funcionalidad en el intercambio de información entre distintas bases de datos. P. ej.: *ISO CD 8459-5: Bibliographic data element directory Part 5. Data elements for the exchange of cataloguing and metadata*, de mayo de 1999

<sup>2</sup> Esta función del DC se empieza discutir en el 2.º workshop celebrado en la ciudad Inglesa de Warwick (1996). Vid. Carl LAGONZE. «The Warwick Framework: A container architecture for diverse sets of metadata». *D-Lib Magazine* [en línea], July/August 1996. Disponible en Internet <<http://www.dlib.org/dlib/july96/lagonze/07lagonze.html>> [Consulta: 20 julio 2000]

recursos web, más aún cuando se está promoviendo el trabajo conjunto con el W3C para una descripción formal para la codificación de los metadatos-DC en el *Resource Description Framework*.

- *Text Encoding Initiative* (TEI): surge en 1987 como un proyecto del área de las humanidades, promovido en un congreso de la *Association for Computers in the Humanities* (ACH), pero su publicación efectiva para la codificación de textos, fundamentalmente literarios, es posterior (1994). Se trata de una DTD (definición del tipo de documento) madura y bien formada de SGML para asegurar un formato estándar con amplias capacidades de marcado para la indización y el intercambio de información textual. A pesar de sus orígenes de propósito lingüístico, las recomendaciones TEI servirán para la estructuración de la información electrónica de todo tipo. Uno de los componentes más significativos del esquema TEI es la cabecera TEI (TEIH) que permite definir una descripción bibliográfica detallada para cada texto codificado como parte de su codificación original. Desde 1998 se está trabajando en una versión de la TEIH en XML, que sin duda ampliará el espectro de aplicación futura de este formato de descripción de contenido web.
- *Resource Description Framework* (RDF): es sin duda alguna el estándar en desarrollo más importante para la descripción de contenidos web. Su importancia reside, por un lado en la entidad que tiene Consorcio Web (W3C) quien fomenta el proyecto, y por otro, el hecho de ser una aplicación de metadatos que utiliza XML a fin de proporcionar un marco estándar para la interoperabilidad entre distintos modelos de metadatos para la descripción de contenido, como p. ej. el DC. Está basado en los trabajos de varios colectivos como otras iniciativas del propio W3C (PICS para el control de contenidos o P3P destinado a salvaguardar la privacidad en la web) y de otros trabajos previos en el ámbito de los modelos de descripción de contenido, como el Warwick Framework del DC. Debemos señalar los tres aspectos de la semántica funcional de RDF: un modelo de datos, una sintaxis y un esquema. La especificación del modelo y la sintaxis —Recomendación del Consorcio Web desde febrero de 1999—, es un estándar estructural de metainformación diseñado para servir como fundamento para la interoperabilidad en el procesamiento de metadatos. La especificación del esquema —desde el 27 de marzo del 2000, Candidato para la Recomendación— formaliza las restricciones para definir vocabularios para las aplicaciones RDF.
- Otros modelos de descripción de contenidos: el mundo de los metadatos es tan amplio y complejo como la propia web; además de las tentativas de propósito general que hemos descrito hasta aquí, existen múltiples iniciativas de metadatos con propósitos específicos, esto es, modelos de descripción del contenido web de cobertura temática, como por ejemplo, el GILS (*Government Information Locator Service*) que en consonancia con el protocolo Z39.50 sirve para describir y recuperar la información gubernamental producida por las agencias federales americanas, o el proyecto INDECS que de-

sarrolla un formato basado en RDF/XML para la información implicada en el comercio electrónico. También hay que destacar las distintas DTDs de SGML, que como el TEI, adoptan una estructura especial de aplicabilidad para distintas materias, este es el caso por ejemplo de iniciativas como CIMI (*Computer Interchange of Museum Information*) para la información de museos digitales, o el estándar de contenido para la información digital geoespacial que lleva el nombre de la institución que lo desarrolla, *Federal Geographic Data Committee* (FGDC), o EAD (*Encoded Archival Description*) para la información archivística, que en su versión 1.0 (1998) estaba basado tanto en SGML como en XML. Incluso algunas iniciativas para la normalización de la identificación de recursos web, como DOI o SICI, incluyen elementos de metadatos que permiten interpretar la información sobre derechos de autor, formato, etc. de los documentos web referidos.

#### 4. La descripción de formal de documentos electrónicos

La descripción externa de los recursos electrónicos, podría verse implícita en la descripción del contenido a través de metadatos, en tanto que éstos presumen una catalogación o descripción, también formal, de los documentos web. Sin embargo, en este apartado nos vamos a ocupar de algo más allá de la descripción formal bibliotecaria, trataremos las formas de hacer referencias «bibliográfica» relativas a la electrónica accesible en Internet. La necesidad de contar con normativa en este sentido es interesante no sólo en el ámbito bibliotecario, sino para cualquier disciplina que utilice información de la web para avalar sus trabajos o proyectos científicos.

Los manuales y guías de estilo para la descripción de documentos se están adaptando a los documentos electrónicos. Muchas de las normas tradicionales de descripción formal, como las ISO, Chicago, Harvard, etc., han incorporado pautas para aplicar a páginas web, mensajes de correo electrónicos, revistas digitales, etc. Además, universidades y otras entidades están realizando sus aportaciones al tema, ya sea ampliando a estos nuevos documentos normas ya clásicas, o bien creando estándares para afrontar la nueva realidad de los documentos electrónicos. Sin embargo, el mayor problema en la descripción formal, en la redacción de las referencias de estos documentos proviene justamente de la falta de estructuración del contenido informativo de los recursos electrónicos. A pesar de que existe un gran número de orientaciones sobre como describir la información electrónica, fundamentalmente web, la mayor parte de las mismas se basan en un reducido número de normas, como son las de la ISO, la MLA, la APA o Havard, siendo las demás adaptaciones de las anteriores. En los párrafos siguientes se detallarán las mismas, con la intención de diferenciar unos métodos de otros.

— ISO: Norma ISO 690-2. *Information and documentation - Bibliographic references - Part 2: electronic documents or parts thereof*. Norma desarrollada por el Subcomité Técnico número 46/9 de la ISO, encargado de los es-

- tándares en materia de presentación, identificación y descripción de documentos. Como todas las normas ISO establece una serie de elementos y un sistema de ordenación de los mismos. Los tipos de documentos a los que afecta son: monografías electrónicas, bases de datos y programas de ordenador, así como partes de los mismos y contribuciones a estos tipos de documentos; además incluye publicaciones electrónicas seriadas; boletines electrónicos, listas de discusión y mensajes electrónicos. En esta norma se pretende que se deje constancia del formato del documento descrito, de la accesibilidad del mismo y de las fechas en las que fue creado y consultado.
- *MLA*: las normas de la *Modern Language Association* se han convertido en un estándar de facto, debido a la presencia en todo el mundo de esta institución y de su actividad en pro de la investigación. Aunque fueron concebidas para documentos impresos ya han incorporado modelos para describir formalmente bases de datos en línea y documentos accesibles a través de Internet. Para las páginas web indica aspectos importantes como las fechas de actualización y consulta y la URL del documento, aunque la disposición de los elementos y su tipografía son diferentes a las indicadas en la norma ISO 690-2.
  - *APA*: las pautas para citar documentos elaboradas por la *American Psychological Association* también son una norma de hecho, ya que se han consolidado gracias a la profusión con que están empleando. Además de las formas de citar documentos impresos la APA da pautas para formatos de referencias de documentos electrónicos, entre los que se encuentran: mensajes de correo electrónico, sitios web, documentos específicos dentro de un web, artículos o resúmenes de bases de datos, webs citados en un texto. Una vez más, los elementos contenidos en la referencia son los mismos que en los estilos anteriores, aunque difiere la ordenación de los mismos.
  - *Harvard*: este estilo también puede considerarse clásico, ya que es uno de los más empleados, sobre todo en el ámbito de la investigación anglosajona, aunque su sistema es utilizado por escritores e investigadores de todo el mundo. Este método también normaliza la descripción formal de documentos electrónicos, entre los que incluye estas categorías: CD-ROMs, URLs, E-Journals, imágenes en línea, documentos web y correo electrónico y listas de discusión. Como es habitual en este método el autor irá siempre vinculado de forma directa a la fecha de producción del documento. Los restantes datos que aparecen son: título, sistema de acceso, edición, lugar de publicación, disponibilidad y fecha de consulta.
  - Otros estilos de citas: muchas universidades han adaptado las normas anteriores o bien otras para establecer sus propias pautas. Un buen ejemplo es la Universidad de Columbia, cuya guía está basada en MLA, Chicago y APA e incluye modelos de referencias de una variada tipología de documentos electrónicos. Otros buenos ejemplos de estas adaptaciones son las Universidades de Laval, Alberta, Bourthmouth, New Jersey y Northumbria. Las bibliotecas de estas universidades han establecido una serie de in-

dicaciones para informar a sus usuarios cómo deben citar recursos electrónicos. Otras iniciativas parten de opiniones personales, como la adaptación que la biblioteca médica del *North Memorial Health Care* hace de la obra de Kate Turabian.

## 5. Conclusiones y reflexiones finales

A tenor de lo que se ha expuesto a lo largo de esta breve comunicación, las conclusiones parecen evidentes: en la Red proliferan los proyectos y tentativas de normalización de la información web para poder recuperarla, compartirla y hacerla compatible. Esta plétora de estándares o proyectos de normalización en desarrollo, denota el estado incipiente del tema de la ubicación, identificación y descripción de documentos web, que dista mucho de contar con un estándares sólidos, fiables e internacionalmente reconocidos.

- En la mayoría de los casos, se trata simplemente de iniciativas más o menos serias de normalización o de estándares que pretenden serlo *de facto* para una comunidad especializada de usuarios. El proceso de estandarización formal evoluciona demasiado lentamente para ajustarse a las necesidades del entorno de información de la web, y en la mayoría de los casos, la falta de masa crítica, dificulta aún más el éxito de estos proyectos.
- Es fundamental en el desarrollo y evolución de estas normas el papel que juegan instituciones encargadas de la estandarización como la ISO o la ANSI/NISO, o las encargadas del desarrollo de Internet y la web como el W3C o el IETF; así como las empresas desarrolladoras de software, que deben asimilar estas tendencias.
- Los objetivos de todas estas iniciativas normativas son los mismos, identificar y recuperar el conocimiento albergado en la web, por ello, y ante la diversidad de este nuevo entorno informativo, los estándares deben desarrollarse de manera paralela hacia normas flexibles de estructura que permitan desarrollar un «web semántico». Los problemas que tratan de paliar las normas tanto de localización como de descripción, son los mismos, básicamente la inconsistencia de la información web. XML, basado SGML (norma ISO 8879-1986) puede marcar la pauta, en su madurez, de una nueva generación de un web flexible y adaptable, pero estandarizado.

## Bibliografía general

- AHRONHEIM, Judy . *Judy and Magda's List of Metadata Initiatives*. [en línea]. 2 nov. 1997. Disponible en Internet <<http://www-personal.umich.edu/~jaheim/alcts/bibaccess.htm>> [Consulta: 20 julio 2000]
- BERNERS LEE, Tim. *Tejiendo la red: el inventor del World Wide Web nos descubre su origen*. Madrid: Siglo veintiuno, 2000

- BURNARD, Lou and LIGHT, Richard. «Three SGML metadata formats: TEI, EAD, and CIMI». En: *Work Package 1 of Telematics for Libraries project BIBLINK (LB4034)*. [en línea] rev. 14 may 1998. Disponible en Internet <<http://www.ukoln.ac.uk/metadata/BIBLINK/wp1/sgml/>> [Consulta: 20 julio 2000]
- CAUSTON, Laurie. *Identifying and describing web resources* [en línea]. 17 nov. 1998. Disponible en Internet <<http://www.elpub.org/html/webres.html>> [Consulta: 4 julio 2000]
- CONNOLLY, Dan. *Naming and addressing: URIs, URLs, Ö* [en línea]. 8 mar. 2000. Disponible en Internet <<http://www.w3.org/Addressing>> [Consulta: 4 julio 2000]
- CRANE, N.B. *Bibliographic formats for citing electronic information* [en línea]. 29 oct. 1997. Disponible en Internet <<http://www.uvm.edu/~ncrane/estyles>> [Consulta: 4 julio 2000]
- CROUSE, Maurice. *Citing electronic information in history papers* [en línea]. 7 sep. 1999. Disponible en Internet <<http://www.people.memphis.edu/~mcrouse/elcite.html>> [Consulta: 4 julio 2000]
- DEMPSEY, Lorcan and HEERY, Rachel. «Metadata: a current view of practice and issues». *Journal of Documentation*, March 1998, vol. 54, n. 2, p. 145-172
- eLIB Standards Guideliness*. Lorcan Dempsey, et al. [en línea]. Version 2.0, 27 oct. 1998. Disponible en Internet <<http://www.ukoln.ac.uk/services/elib/papers/other/standards/version2/>> [Consulta: 20 julio 2000]
- ESTIVILL, Assumpció; URBANA, Cristóbal. «Cómo citar recursos electrónicos» [en línea]. 30 may. 1997. Disponible en Internet <<http://www.ub.es/div5/biblio/citae-e.htm>> [Consulta: 4 julio 2000]
- FLETCHER, Gordon; GREENHILL, Anita. *Academic referencing of Internet-based resources* [en línea]. Abr. 1997. Disponible en Internet <<http://www.spaceless.com/WWWVL/refs.html>> [Consulta: 4 julio 2000]
- HAIGH, Susan. *Glossaire des normes, des protocoles et des formats liés à la bibliothèque numérique* [en línea]. 6 may. 1998. Disponible en Internet: <<http://www.nlc-bnc.ca/pubs/netnotes/fnotes54.htm>> [Consulta: 20 julio 2000]
- HEERY, Rachel. «D1.1 Metadata Formats». En: *WP Study of Metadata*, issue 1.0. [en línea] 23 dic. 1996. Disponible en Internet <<http://hosted.ukoln.ac.uk/biblink/wp1/d1.1.rtf>> [Consulta: 20 julio 2000]
- IFLA. *Citing guides form electronic documents* [en línea]. 30 sep. 1999. Disponible en Internet <<http://www.ifla.org/I/training/citation/citing.htm>> [Consulta: 4 julio 2000]
- IFLA. *Digital libraries: metadata resources* [en línea]. 19 may. 2000. Disponible en Internet <<http://www.ifla.org/II/metadata.htm>> [Consulta: 20 julio 2000]
- LI, X.; CRANE, N.B. *Electronic styles: a handbook for citing electronic information*. Medford, NJ: Information today, 1996.
- LAGONZE, Carl. «The Warwick Framework: A container architecture for diverse sets of metadata». *D-Lib Magazine* [en línea] July/August 1996. Disponible en Internet <http://www.dlib.org/dlib/july96/lagonze/07lagonze.html> [Consulta: 20 julio 2000]
- LYNCH, Clifford. «Identifiers and their role in networked information applications». *A Bi-monthly Newsletter of Research Library Issues and Actions* [en línea]. Issue 195, 29 dic. 1997. Disponible en Internet <<http://www.arl.org/newsltr/194/identifier.html>> [Consulta: 4 julio 2000]
- MÉNDEZ RODRÍGUEZ, Eva M.ª «RDF: un modelo de metadatos flexible para las bibliotecas digitales del próximo milenio». En: *Jornades Catalanes de Documentació*. Barcelona: Col.legi Oficial de Bibliotecaris-Documentalistes de Catalunya, 1999, p. 487-498. Disponible también en Internet <<http://www.bib.uc3m.es/~mendez/publicaciones/7jc/rdp.htm>>
- PASKIN, Norman. *Information identifiers* [en línea]. 9 abr. 1998. Disponible en Internet <<http://www.elsevier.nl/inca/homepage/about/infoident>> [Consulta: 4 julio 2000]



- POWELL, Andy. «Unique identifiers in a digital world». [en línea]. *Ariadne*, vol.8, 8 abr. 1997. Disponible en Internet <<http://www.ariadne.ac.uk/issue8/unique-identifiers>> [Consulta: 4 julio 2000]
- TONG, Josie. *Citation style guides for Internet and electronic sources* [en línea]. 14 feb 2000. Disponible en Internet <[http://www.library.ualberta.ca/library\\_html/help/pathfinders/style/style.html](http://www.library.ualberta.ca/library_html/help/pathfinders/style/style.html)> [Consulta: 4 julio 2000]
- TURABIAN, Kate L. «A manual for writers of term papers, theses, and dissertations». 5th ed. Chicago: University of Chicago Press, 1987.
- WORLD WIDE WEB CONSORTIUM. *Resource Description Framework (RDF): Model and Syntax Specification. W3C Recommendation, 22 February 1999* [en línea]. Ora Lassila and Ralph R. Swich, eds. 22 feb. 1999. Disponible en Internet <<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>> [Consulta: 20 julio 2000]
- WORLD WIDE WEB CONSORTIUM. *Resource Description Framework (RDF) Schema Specification 1.0 W3C Candidate Recommendation 27 March 2000* [en línea]. Dan Brickley, R.V. Guha, eds. 27 mar. 2000. Disponible en Internet <<http://www.w3.org/TR/2000/CR-rdf-schema-20000327/>> [Consulta: 20 julio 2000]

### Sitios Web relacionados

American Psychological Association (**APA**): [Http://www.apa.org/journals/webref.html](http://www.apa.org/journals/webref.html)  
[Http://owl.english.purdue.edu/Files/34.html](http://owl.english.purdue.edu/Files/34.html)

Dublin Core Metadata (**DC**): [Http://purl.org/DC/](http://purl.org/DC/)

Digital Object Identifier System (**DOI**): [Http://www.doi.org/](http://www.doi.org/)

Encoded Archival Description (**EAD**): [Http://lcweb.loc.gov/ead/](http://lcweb.loc.gov/ead/)

Content Standards for Digital Geospatial Metadata (**FGDC**): [Http://www.fgdc.gov/Metadata](http://www.fgdc.gov/Metadata)

Government Information Locator Service (**GILS**): [Http://www.gils.net](http://www.gils.net)

Handle System (**HS**): [Http://www.handle.net](http://www.handle.net)

ISO 690-2 (**ISO**): [Http://www.nlc-bnc.ca/iso/tc46sc9/standard/690-2e.htm](http://www.nlc-bnc.ca/iso/tc46sc9/standard/690-2e.htm)

Modern Language Association (**MLA**): [Http://www.mla.org/publications/stylemanual\\_index.htm](http://www.mla.org/publications/stylemanual_index.htm)

Platform for Privacy Preferences (**P3P**): [Http://www.w3.org/P3P/](http://www.w3.org/P3P/)

Platform for Internet Content Selection (**PICS**): [Http://www.w3.org/PICS/](http://www.w3.org/PICS/)

Publisher Item Identifier (**PII**): [Http://www.elsevier.nl/inca/homepage/about/pii/](http://www.elsevier.nl/inca/homepage/about/pii/)

Persistent Uniform Resource Locator (**PURL**): [Http://purl.org/](http://purl.org/)

Resource Description Framework (**RDF**): [Http://www.w3.org/RDF](http://www.w3.org/RDF)

Serial Item and Contribution Identifier (**SICI**): [Http://sunsite.berkeley.edu/SICI/](http://sunsite.berkeley.edu/SICI/)

Text Encoding Initiative (**TEI**): [Http://www-tei.uic.edu/orgs/tei/](http://www-tei.uic.edu/orgs/tei/)  
[Http://etext.virginia.edu/TEI.html](http://etext.virginia.edu/TEI.html)

Uniform Resource Identifiers Working Group (**URI**): [Http://www.ics.uci.edu/pub/ietf/uri/](http://www.ics.uci.edu/pub/ietf/uri/)

Uniform Resource Name Working Group (**URN**): [Http://www.ietf.org/html.charters/urn-charter.html](http://www.ietf.org/html.charters/urn-charter.html)

eXtensible Markup Language (**XML**): [Http://www.w3.org/XML](http://www.w3.org/XML)