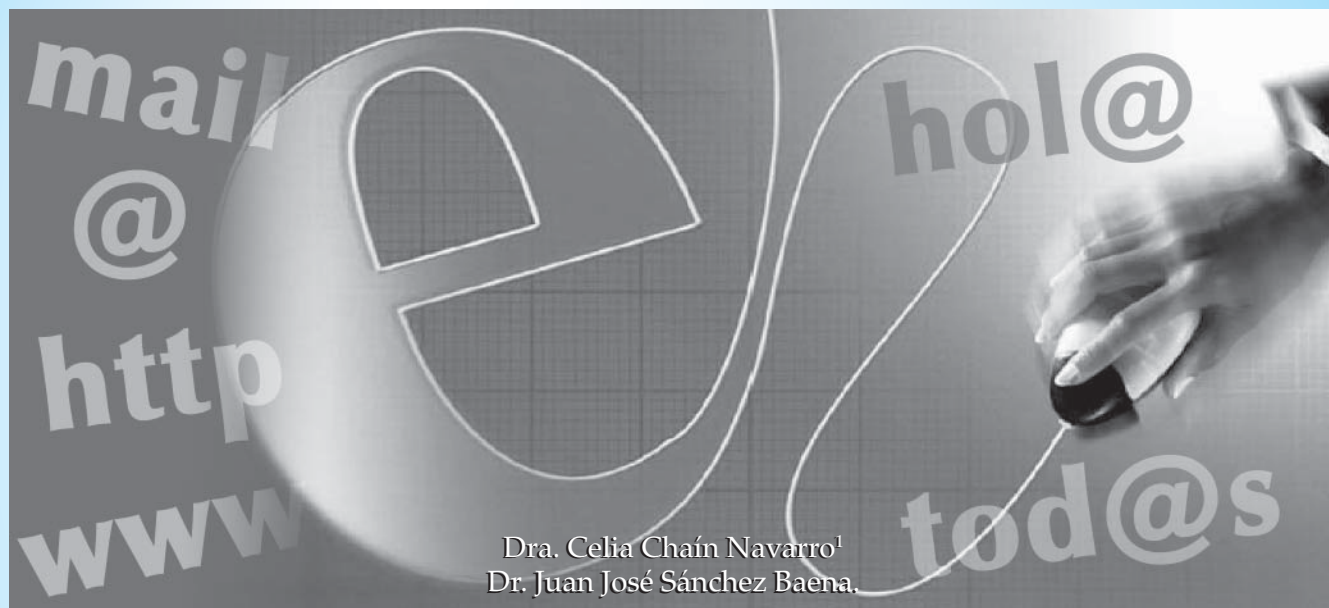


# Hacia la Web semántica como red de identidades en América Latina



Dra. Celia Chaín Navarro<sup>1</sup>  
Dr. Juan José Sánchez Baena.

## 1. Introducción

### 1.1 El concepto de “Web semántica”

La “web semántica” (Semantic web) es un concepto acuñado por el mismísimo creador de Internet, Berners-Lee<sup>2</sup> en el año 1998, y se refiere a la necesidad de estructuración semántica en la Red. Una vez que Internet ya está generada, o lo que es lo mismo, acabada ya la primera etapa (creación, desarrollo de la Red, y posibilidad de incorporación casi universal), se dio cuenta de que a partir de ese momento se hacía necesario avanzar en una segunda etapa: la de generar herramientas lingüísticas que apoyen y faciliten la ordenación, gestión y recuperación de los datos, tanto los que ya estaban incluidos en la Red como los que se podían añadir. Así, han ido apareciendo intentos variados de ordenación, estructuración, normalización, control e indización cuyo principal objetivo es facilitar al usuario la recuperación de la información que está buscando en Internet, pero que es difícil de hallar por las características intrínsecas y extrínsecas de ésta. Todos ellos se incluyen en el intento de generar esa Websemántica que un día pensó su creador como una segunda fase en la consolidación de la Red.

### 1.2 Identidad digital

Bajo el concepto de identidad digital pretendemos incluir los datos y la información (términos, y cualquier tipo de imágenes, logotipos, dibujos) que pueden servir para identificar a través de la Red a colectivos, países, sitios webs, ideologías, instituciones, etc. Resumidamente, podíamos decir que es la forma de que significado y significante se consideren relacionados en el entorno de Internet. De esta forma, es posible que identidad e identificación puedan ir unidas, y siempre dentro del ámbito indicado. Al menos desde el punto de vista de quienes identifican, aunque para los identificados ésto no sea así.

Millán, en este sentido argumenta que “la mundialización del mercado ha subrayado un hecho del que ya se tenía conciencia, aunque con un alcance limitado: que las personas que comparten una lengua constituyen de hecho una comunidad”. Y que “gran parte de la actividad basada en la lengua (información, consulta, referencia, edición, traducción, redacción, etc.) se ha ido pasando también al mundo digital”<sup>3</sup>.

<sup>1</sup> Dra. Celia Chaín Navarro, Profesora Titular de Biblioteconomía y Documentación. Facultad de Ciencias de la Información y Documentación. Universidad de Murcia. e-mail: chain@um.es.

<sup>2</sup> Dr. Juan José Sánchez Baena. Profesor Titular de Historia de América. Universidad de Murcia. e-mail: baena@um.es.

Pocos años después, tras un enorme desarrollo, y un éxito arrollador, en Internet podemos encontrar “marcas” de identificación, sean o no globalmente adquiridas o utilizadas, y referidas al propio entorno, a sus actores, o a sus usuarios. Baste recordar el famoso signo “arroba”, y cómo ha pasado de ser utilizado en las direcciones de correo electrónico para separar la denominación particular del usuario dueño de la cuenta, del servidor que la aloja, a encontrarlo en textos significando, y por lo tanto representando a ambos géneros (“hola a tod@s”).

El concepto de “sociedad de la información”, frecuentemente ligado a ideas capitalistas disfrazadas de globalización cultural, aplicado o estudiado fuera de los ambientes en los que se generó, atestigua la falta de adecuación a un mundo donde las desigualdades ya eran patentes, y es más, ahora con este nuevo parámetro, es posible que contribuyan a acentuarlas más. Sin embargo, admitido y aceptado que el modelo mítico de sociedad de la información y del conocimiento se desvanece en muchos de sus hitos, nadie puede negar la importancia y trascendencia que para cualquier ciudadano del mundo que puede conectarse a la Red tiene, y tendrá en el futuro, la posibilidad de acceder desde un solo sitio a un universo casi infinito de datos, información, ideología, formas de pensar, de rezar, etc. Evidentemente, la denominada “brecha digital” ha contribuido a aumentar las diferencias entre los que pueden y los que no pueden conectarse a la Red, pero también a hacer posible una mayor homogeneidad entre los que tienen la posibilidad de conectarse.

Con respecto a América Latina, Martín Barbero ha escrito recientemente sobre el crecimiento de sitios webs no comerciales, y con un marcado carácter educativo o cultural no formal. Destaca también como una de las características el altísimo número de jóvenes que usan la Red, por ejemplo, un 49% en México y un 55% en Brasil. Y subraya que “no obstante las precarias condiciones en las que las redes y la creatividad digitales se desarrollan en Latinoamérica, ellas representan cada día más el incontenible surgimiento y conformación de un nuevo espacio público y nuevos modos de creación cultural”<sup>4</sup>. Y termina admitiendo que el ciberespacio es una parte constitutiva del denominado espacio cultural iberoamericano en construcción. Otros autores son menos generosos en el reconocimiento de las ventajas de la Sociedad de la Información en

América Latina, en general, y de Internet en particular. Son, por ejemplo, Narváez, Trejo, Becerra, Mastri, García Canclini, Hopenhayn, o Pasquali.

En este momento de “imposición tecnológica” no queremos dejar de destacar que si bien los que se pueden incluir en Internet tienen en ella la posibilidad de avanzar más rápidamente, existe también esa “caída libre” al vacío para los que no tienen posibilidades de hacerlo. La parte del mundo conectada define en parámetros de desarrollo una nueva variable, que expresada en diferentes formas de medida (cantidad de ordenadores por habitante, de dominios por países, espacios webs, direcciones de correo electrónico, conexiones en hogares, etc.) viene a definir una nueva diferencia. En esa necesidad casi obligada de integración hay colectivos que siguen perdiendo fuerza, y que sin embargo en América Latina son una señal de identidad incontestable: las sociedades indígenas. La vulnerabilidad ante la globalización tecnológica es evidente, pero también solucionable.

García Canclini se pregunta “qué significa ser latinoamericano”, y comenta que la “respuesta está cambiando a comienzos del siglo XXI, se desvanecen las repuestas y surgen dudas sobre la utilidad de tomar compromisos continentales”<sup>5</sup>. Hay propuestas, críticas, sobre ese “rompimiento” subcontinental. Y dentro de esa preocupación por analizar el impacto del proceso de globalización sobre América Latina, y de proponer medidas adecuadas para enfrentarse con éxito a los desafíos que presenta dicho proceso<sup>6</sup>, aquí proponemos la posibilidad de utilizar una herramienta fácil, rápida y gratuita para generar en poco tiempo un subconjunto lingüístico con una identidad común en una parte de la Web semántica: los metadatos. No se trata de promocionar nacionalismos, ni macroculturas, sino de hacerse hueco en un pastel que, aunque cueste trabajo crearlo, lo hacemos entre todos los que formamos parte de la Red. Millán lo expresó así:

“¿es interesante, es valioso que exista un vínculo entre hispanohablantes (que épocas tal vez superadas denominaron Hispanidad)? [...] ¿puede la Red contribuir a crear este vínculo o, caso de que ya exista, a mantenerlo articulado?. El espacio de la Internet en español puede convertirse en el soporte de un ente político, pero sólo en la medida en que exista un proyecto para hacerlo”<sup>7</sup>.

Y el marco de posibilidades ha cambiado tanto que ahora no es necesario que la idea parta de las antiguas metrópolis coloniales. En el marco digital el peso viene dado por el número de espacios webs incluidos, por el número de enlaces recibidos, no por la preponderancia política ni económica de otros tiempos. Y aunque suene extraño, la tan llevada y traída, criticada y utilizada, idea de la Liga Hispanoamericana de Simón Bolívar, o las similares del Che Guevara, o de José Martí son hoy más fáciles de alcanzar en Internet que en el mundo real. Es decir, que la Red se forma día a día, la forman todos los que en ella incluyen sistemas webs, los que la navegan, los que generan herramientas para su uso, incluso los que la critican. Una página, un espacio web es más o menos importante según las conexiones que otras han generado hacia ella, según el número de visitas que tiene. En las tecnologías creadas para el desarrollo de la web semántica un espacio web será mejor cuanto más estructurada esté su información, ya que esta organización será la que permita que ciertas herramientas estructurales como las ontologías organicen sus contenidos y clasifiquen su información.

El problema radica en que las páginas en español suponían en septiembre de 2004 sólo el 9 % del total, y según la tendencia, este porcentaje está disminuyendo. Con una cifra tan pequeña, y con la desventaja de la preponderancia del inglés, está claro que la mayoría de las herramientas estarán en este idioma, lo que irremediabilmente hará que los espacios en español sean cada vez menos visibles, menos importantes. Si se empieza a facilitar el desarrollo de esta web semántica en un idioma diferente al inglés, la tendencia será diferente. Si existe un porcentaje importante de sistemas webs que contienen información semántica, la propia inercia de Internet tenderá a facilitar que se desarrollen thesauri, ontologías para ese idioma; sino seguirán dirigiéndose, como pasa en la actualidad, al inglés.

En el año 1998, Donoso, de la Universidad de Chile, ya apuntaba que "metadatos es una metodología reciente que requiere la masificación de su uso por parte de los generadores de sitios web y por los motores de búsqueda y directorios. Probablemente se requerirá una instancia nacional que se encargue de estandarizar su uso y los esquemas de recuperación que se utilicen y asegure una catalogación normalizada"<sup>8</sup>.

### 1.3 Metadatos

Los metadatos, en el ámbito de Internet, podemos definirlos como todos aquellos datos comprensibles por el ordenador, estructurados de tal forma que los podemos utilizar para localizar, identificar y describir el contenido de un espacio Web, y por lo tanto para la búsqueda y recuperación de información en la Red.

Una de las posibilidades de avanzar más rápidamente hacia esa "Web Semántica" es facilitar la recuperación de información a través de los metadatos de los espacios webs. Por ello es necesario inicialmente realizar una recogida y análisis de los metadatos contenidos en los sistemas webs. Estudios anteriores nos han permitido afirmar que la aparición y evolución de los espacios web ha sido relativamente rápida y homogénea. Sin embargo, una cuestión es la inserción en la web, y otra muy diferente es cómo se pueden indizar y clasificar por parte de los motores de búsqueda y de los agentes inteligentes. Recientes estudios nos permiten adelantar que si los gobiernos no se toman más en serio la necesidad de colaborar en la creación de la Web semántica incorporando metadatos normalizados a sus espacios webs, difícilmente podremos avanzar en la construcción de una red estructurada<sup>9</sup>, y muchos menos aspirar a que en una parte de esa "Web semántica" tenga cabida el idioma español.

Para la generación de esa "red de contenidos" son necesarias dos tipos de herramientas:

- a) las tecnológicas, programas informáticos que permitan generar, gestionar, buscar y recuperar los contenidos de esas estructuras. Ya existen buscadores de metadatos como Apache, Beagle, o Swoogle.
- b) las documentales y/o lingüísticas, o lenguajes estructurados, de naturaleza combinatoria y/o jerárquica, cuyos términos sirvan para describir los contenidos de los documentos, y que puedan usarse a la vez como esquemas en los sistemas de inteligencia artificial generados. Lenguajes de esta categoría hay pocos (por ejemplo, el que se utiliza en la base de datos Medline, en Chemical Abstracts para la formulación química, o el recientemente aparecido thesaurus multilingüe de la Red Europea de Patrimonio (HEREIN). Millan, en una arriesgada propuesta, ofrece un índice sobre el peso que tienen estas herramientas y tecnologías en determinados productos o servicios en Internet.

Producto/servicio	Índice PTL
Comercio electrónico	0,01
Industria del copyright	0,01
Información turística	0,03
Sistemas operativos	0,05
Procesadores de texto	0,10
...	
Buscadores en la Red	0,30
Gestores de información	0,50
Agentes inteligentes	0,80
...	
Programas de ayuda a la traducción	0,90
Asistentes terminológicos	0,90

Tabla 1. Peso de las tecnologías lingüísticas por producto o servicio (PTL). Fuente: MILLAN, J.A. La lengua que era un tesoro. El negocio digital del español y cómo nos quedamos sin él [web]. [Consulta 31-05-2005]. Disponible en <<http://jamillan.com/tesoro.htm>>.

Según la tabla 1, los programas que buscan, gestionan y recuperan información tienen un uso de estas tecnologías que oscila entre el 0'30 y el 0'90 sobre 1. Así que el desarrollo de éstas o, como aquí proponemos, el sólo hecho de empezar a facilitar su difusión y utilización desde los propios gobiernos, podría ser una forma de avanzar con pasos de gigante en una red mundial donde siempre ha prevalecido el inglés, y en la que, según las previsiones, continuará haciéndolo. Pero realmente aquí el problema no son los porcentajes de espacios webs en inglés, sino la preponderancia de la cultura anglosajona, con todo lo que de positivo y negativo lleva implícito. No es una cuestión de cambiar un idioma dominante por otro, ni de identificaciones políticas trasnochadas, sino de facilitar y mejorar una evolución que tendrá lugar, y que puede tener o no como identificativo cultural inicial un idioma común, un friend of a friend (FOAF)<sup>10</sup> hispano intergubernamental.

Es necesario que por parte de los gobiernos, es decir a partir de las webs institucionales, se empiece a tener en cuenta que integrar metadatos en cualquier espacio web es gratuito, cómodo, fácil, y además facilita la recuperación de la información de ese espacio en cualquier herramienta de búsqueda. Pero es que este uso acordado también puede facilitar que la comu-

nidad hispanohablante se deje oír, y tenga un sitio definido, una identidad buscada en la Red.

Los metadatos son una de las soluciones para el avance y desarrollo de la web semántica, ya que facilitan la descripción de los contenidos, proponen una estructura común, y existen ya propuestas para generar motores y sistemas de búsqueda basados en ellos.

#### 1.4 Los metadatos en la “web semántica”

El concepto de “web semántica” responde, como ya hemos anticipado, a la necesidad de facilitar el acceso a toda la información existente en Internet. Supone construir una base (o muchas y variadas entrelazadas) de conocimiento que el software sea capaz, mediante procedimientos de Inteligencia artificial, de procesar, “razonar”, combinar términos con sentido y realizar deducciones lógicas para resolver preguntas.

Con el actual formato de la web, ésto es imposible, ya que la Red es universal, y por lo tanto heterogénea y multilingüe. Se podría decir que la primera fase de implantación de la Red ha sido un gran éxito: se han generado protocolos comunes que permiten acceder desde cualquier ordenador, en cualquier sistema operativo a Internet. O lo que es lo mismo, se ha logrado normalizar la estructura de los datos. Gracias a ese “orden universal”, Internet nos llega a todos los que podemos acceder a ella. Los sistemas webs son una realidad patente y conseguida.

Aunque sólo sea para que se pueda entender más fácilmente esta propuesta, denominamos “sistema web” al conjunto de signos, programas y medios que hacen posible que accedamos en Internet, a una dirección determinada, a una información concreta. Resumidamente son los ingenios técnicos, tecnológicos, informáticos y de telecomunicaciones que permiten conectarse a la Red. Y, paralelamente, entendemos y denominamos “espacio web” al contenido de esos sistemas. Los primeros son datos que facilitan la incorporación a la Red, los segundos son también datos, pero que permiten entender la información en ellos contenida, y que además parece que son “accesibles” de forma independiente a los medios que los facilitan. Así, mientras que existen normas que homogeneizan en porcentajes muy superiores esos estándares de acceso, no lo están los que convierten, dentro de los espacios webs, esos datos en información.

Es por ello por lo que ahora falta la normalización de la segunda fase, la que se refiere a la información que contienen. A primera vista pueda parecer lo mismo, pero no es así: cualquiera de nosotros puede acceder a un espacio web en un idioma que no conoce, en ese momento estará recibiendo y visualizando datos, pero al no estar familiarizados con los signos que forman ese lenguaje no estamos recibiendo información. Los datos en Internet son “casi universales”, la información que llevan consigo no. Y esa es la segunda parte. Para que la web semántica funcione los ordenadores tienen que acceder a colecciones estructuradas de datos y a conjuntos de reglas de inferencia, que les ayuden a “entender” y a indizar los documentos. Una de las formas de facilitar esta transición es añadiendo metadatos a los espacios y sistemas webs que se generan.

### 1.5 Detractores de la web semántica

El artículo clásico sobre la falta de utilidad y perspectivas de la Web semántica, es el de Doctorow<sup>11</sup>, que la denomina meta-utopía. Sin embargo, admite la utilidad de los metadatos.

Evidentemente, si cada diseñador de espacios webs, incluye los metadatos que mejor le parecen, o no los incluye, será más difícil que la web semántica se genere pronto, y será mucho más complicado que sean útiles y eficaces las herramientas de gestión como ontologías, thesauri, topic maps, etc. Por ello, desde las W3C, y por parte de otros grupos e investigadores se está intentando concienciar sobre la necesidad de crear grupos de trabajo comunes, y por ello también aparecen cada vez más propuestas sobre la necesidad de generar metadatos y herramientas de gestión de la web semántica en español. Un pequeño repaso al número de webs oficiales que utilizaban metadatos en el año 2001, y su comparación con las que los utilizaban en 2004, pone de manifiesto el rápido incremento sufrido. Sin embargo, este análisis también deja clara la necesidad de homogeneizar, de utilizar códigos comunes, y de un liderazgo para que la web del futuro no sea solamente en inglés.

### 1.6 Metadatos como elementos para facilitar la búsqueda y recuperación de información

Muchos han sido los investigadores que se han ocupado de definir el término “metadato”<sup>12</sup>, entre otros, Myers 1960, Liston & Dolby 1982, Ribeiro 1995, Heery y el proyecto Biblink 1996, Berners-

Lee 1997, Cathro 1997, Dempsey & Heery 1997<sup>13</sup>, Iannella & Waugh 1997, Kerhervé & Gerbé 1997, Lange & Winkler 1997, Pasquinelli 1997, Chilvers & Feather 1998, Taylor 1999, Seiner 2000, y Méndez Rodríguez<sup>14</sup>, y Senso y De la Rosa, 2003<sup>15</sup>.

Para Taylor son “datos con sentido propio que proporcionan información o documentación acerca de cómo incluir información descriptiva acerca del contexto, calidad y condiciones o características de los datos”<sup>16</sup>. A pesar de que muchos de los investigadores arriba citados se empeñen en incluir los metadatos como sistemas de catalogación en Internet, es necesario realizar una precisión a este respecto, sobre todo a los lectores ajenos al mundo documental. Además de ser elementos de catalogación, es decir, de descripción externa, los metadatos incluyen descripción de contenidos, resumen, y una serie de términos que describen de forma más completa y exhaustiva un documento web que la simple catalogación de éste. Son identificadores inequívocos de recursos web que contienen mucha más información que la que habitualmente se incluyen en una ficha catalográfica.

Hablar de la descripción de contenidos webs, implica hablar del complejo mundo de los metadatos o de la metainformación. Desde mediados de los años noventa, diversos colectivos están desarrollando modelos o pseudo-estándares de metadatos, con distintos niveles de difusión, para describir sus colecciones digitales. Otros están esperando que se desarrollen más estos formatos para implementarlos de una forma más fácil y sólida en sus sistemas de información. Mientras, otros servicios tratan de solucionar la incorporación de documentos Web a través del campo 856 del formato USMARC<sup>17</sup>.

El concepto de metadato es antiguo y muy claro para los documentalistas acostumbrados a utilizar colecciones de metadatos manuales como herramientas básicas de la profesión: clasificaciones, tesauros, nomenclaturas, códigos y etiquetas. Incluso a los lenguajes de indización (vocabularios controlados) se le han llamado a veces metalenguajes. Sin embargo, otras personas que tratan con la información en sus trabajos (informáticos, estadísticos, geógrafos, y todos los que procesan datos, especialmente numéricos o factuales) se han visto en la necesidad de acudir a esta estandarización de sus parámetros y variables para poder seguir construyendo sus colecciones de datos de forma coherente, comparable e integrable con otras colecciones.

Actualmente nos podemos encontrar con innumerables sistemas de metadatos, y ante la imposibilidad de especificarlos todos, en la siguiente lista aparecen algunos de los considerados más importantes:

- El Dublin Core (DC): Basado en los formatos tradicionales para la descripción de registros en bibliotecas. Norma ISO 15836 (2003).
- El Recordkeeping Metadata Standard for Commonwealth Agencies (RMSCA): Modelo de metadatos ideado para la gestión de documentos electrónicos en archivos.
- Encoding Archival Description (EAD): Es un estándar destinado a la descripción de archivos para facilitar un acceso uniforme a los recursos electrónicos de archivos y colecciones de manuscritos de bibliotecas, sociedades y museos. Es una DTD.
- EELS Metadata Format: Es utilizado en el proyecto Engineering Virtual Library (Reino Unido) para proporcionar acceso a recursos de información electrónica de calidad que existen en Internet.
- FGDC- Content Standards for Digital Geospatial Metadata: es un formato complejo que incluye más de 300 elementos, diseñado para manejar información digital geoespacial.
- Government Information Locator Service (GILS): Fue elaborado con la finalidad de poder proporcionar al público en general, y a los empleados federales de los Estados Unidos de América, un medio para localizar información de utilidad generada por las diferentes agencias gubernamentales.
- Text Encoding Initiative (TEI) Headers: Es un formato basado en las directrices TEI, las cuales constituyen el conjunto de guías para la representación en forma electrónica de material textual que le permita a un investigador de cualquier disciplina intercambiar y utilizar este tipo de recursos, independientemente del software, hardware o aplicación utilizados. Está definido, igual que EDA, como una DTD.

### 1.6.1 Dublin Core (DC):

DC es el formato de metainformación más divulgado y más citado a nivel general. El DC tiene sus

comienzos en el año 1995 cuando la OCLC (Online Computer Library Center) y el NCSA (National Center for Supercomputing Applications) convocaron a 52 expertos en el campo de la biblioteconomía, ciencias de la computación, codificadores de textos y áreas afines con el fin de crear un formato normalizado que promoviese un estándar, de propósito general y sencillo, para describir los recursos Web de cualquier materia.

Originalmente se concibió como un conjunto de etiquetas que deberían ser generadas por el autor del documento HTML con la finalidad de facilitar su identificación y posterior recuperación en Internet. No obstante, este modelo ha llamado la atención de diversas comunidades de profesionales interesados en la descripción de recursos en museos, bibliotecas, organizaciones comerciales<sup>18</sup>, etc.

Los objetivos básicos que dieron pie al primer conjunto de metadatos, el Dublin Core Metadata Initiative (DCMI)<sup>19</sup>, son resumidos por Senso<sup>20</sup>:

- Establecer un sistema normalizado para la descripción de documentos distribuidos en el World Wide Web, cualquiera que sea su formato: HTML, PDF, PS, SGML...
- Facilitar a autores y editores de estos documentos la incorporación de elementos que identifiquen y describan sus aportaciones.
- Corregir las deficiencias de los sistemas de recuperación basados en texto completo.
- Evitar el uso de formatos complejos (MARC, TEI, etc.).
- Crear una serie de elementos:
  - Fáciles de crear y actualizar.
  - Comprensibles para cualquier usuario.
  - Normalizados.
  - Ampliables manteniendo la compatibilidad con definiciones anteriores.
  - A los que se pueda aplicar diferentes niveles de especificidad y control mediante el uso de cualificadores y subtipos.

Y que esos elementos:

- faculten la integración haciendo uso de sistemas de clasificación, indización y control de autoridades con los que la comunidad bibliotecaria se encuentra familiarizada, y

- Tengan carácter internacional en su creación, desarrollo e implementación.
- Estén referidos a las características intrínsecas de los documentos.
- Sean opcionales.
- Sean repetibles tantas veces como sea necesario.

En las primeras versiones Dublin Core estaba formado por 13 elementos (o campos) que permitían identificar las características básicas de un recurso electrónico. A pesar de que un año más tarde se le añadieron 2 elementos más, el modelo de referencia sobre el que se ha trabajado ha sido siempre el mismo. En este momento, año 2005, cuenta con 17 elementos.

Elementos o Campos	Descripción
1. Título (“Title”)	Título dado al recurso por su creador o editor
2. Autor (“Creator”)	La persona u organización responsable del contenido intelectual del recurso.
3. Materias y Palabras clave (“Subject”)	El tema del recurso: palabras clave o frases que describen el tema o el contenido del recurso, incluyendo vocabularios controlados o cuadros de clasificación.
4. Descripción (“Description”)	La descripción textual del contenido del recurso, incluyendo resúmenes en el caso de objetos de tipo documento o descripciones en el caso de recursos visuales
5. Editor (“Publisher”)	La entidad responsable de que el recurso esté disponible en su forma presente, tal como un editor, un departamento de universidad o entidad corporativa
6. Contribuciones (“Contributor”)	Personas u organizaciones en adición a aquellas especificadas en elemento “Autor” que ha realizado contribuciones intelectuales significativas al recurso.
7. Fecha (“Date”)	La fecha en la cual el recurso se hizo accesible en su forma presente.
8. Tipo de recurso (“Type”)	La categoría del recurso, tal como sede web, novela, poema, informe, informe técnico, ensayo, diccionario...
9. Formato (“Format”)	La representación de datos del recurso, tal como “text/html”, ASCII, “Postscript”, Aplicaciones ejecutables o imágenes.
10. Identificador del recurso (“Identifier”)	Cadena de número para identificar unívocamente al recurso. Ejemplos para recursos en red incluyen URL y URN .
11. Fuente (“Source”)	El trabajo, ya sea impreso o electrónico, del que deriva el recurso, si es aplicable.
12. Idioma (“Language”)	Lengua (s) del contenido intelectual del recurso
13. Relaciones (“Relation”)	Relaciones con otros recursos. La intención de este elemento es proporcionar un medio para expresar relaciones entre recursos formalmente relacionados, pero que existen como recursos discretos cada uno de ellos.
14. Cobertura (“Coverage”)	Las características espaciales y temporales del recurso. Especificaciones en el desarrollo del estándar
15. Propiedad Intelectual (“Rights”)	La intención de este elemento es proporcionar un enlace (una URL u otro URI disponible si procede) hacia una información de copyright, una afirmación de gestión de derechos, o tal vez, un servicio que proporcione esta información de forma dinámica.
16. Orígen (“Provenance”)	Sirve para reflejar los cambios de custodia o propiedad del recurso desde su creación. Sólo los cambios importantes para su autenticidad, integridad o interpretación.
17. Derechos de propiedad (“RightsHolder”)	Persona u organización que es propietaria o que gestiona los derechos de uso del recurso.

Tabla 2. Elementos del formato Dublín Core.

De forma general, la entrada de un elemento descriptivo dentro de HTML tiene la siguiente sintaxis:

```
<META NAME="NombreElemento"CONTENT="VALOR">
```

donde "NombreElemento" y "Valor" se sustituirán respectivamente por cada uno de los 17 elementos y su valor.

El logro más importante, desde el punto de vista de la normalización, tuvo lugar el 10 de septiembre de 2001, cuando la ANSI norteamericana aprobó definitivamente el documento que reconocía el conjunto de etiquetas DC como norma internacional con la denominación ANSI/NISO Z39.85-2001<sup>21</sup>.

Por su parte, el 26 de febrero de 2003 fue aprobada la norma ISO 15836, reguladora del conjunto de metadatos DC<sup>22</sup>.

### **Dublin Core en España y América Latina**

En agosto del año 2004 se celebró un taller sobre Web semántica en Argentina, precisamente destinado a dar a conocer este concepto, y con la idea de promocionar o concienciar sobre la necesidad de creación de trabajos e investigaciones sobre la web semántica en español<sup>23</sup>. Previamente, y en el mismo año se había celebrado otro en la capital de España. También en noviembre de ese año tuvieron lugar las 4as. Jornadas Iberoamericanas de Ingeniería del software e ingeniería del conocimiento en Brasil, en las que aparte de las comunicaciones se propuso la realización de un taller sobre ontologías y web semántica, que al final no llegó a celebrarse.

Productos desarrollados para el idioma español son, por el momento pocos, de ellos son destacables, por ejemplo, el gestor de thesauri, TemaTres, y el generador automático de metadatos del grupo de metadatos de la Universidad de Chile

En el entorno español, la mirada de apatía y/o escepticismo con la que la mayoría de profesionales de la información se ha fijado en todos los sistemas de metadatos en general, y en Dublin Core en particular, ha frenado su posible uso y propagación.

Salvo contadas excepciones (Puerto Rico), ninguna biblioteca española ni latinoamericana usa DC para catalogar recursos electrónicos, y son pocos

los proyectos –públicos o privados– que se valen de este sistema para realizar la descripción de cualquier objeto electrónico, si bien es cierto que sí usan otros mecanismos, como RDF (en varios proyectos realizados por la editorial Anaya, por ejemplo) o variaciones sobre la base SGML (como algunos trabajos del Instituto Cervantes y el proyecto COVAX).

Uno de los trabajos más avanzados, desde el punto de vista de la tecnología utilizada, que usa DC es SARAC<sup>24</sup> (Servicio de Acceso a Recursos de Alta Calidad). Idea fraguada entre la Red IRIS y la Universidad de Granada, que trata de ofrecer un buscador de recursos caracterizados por su alta calidad informativa. Para lograr ese valor añadido se sirven de un conjunto de metaetiquetas propias y otras de DC que, enmarcadas dentro de RDF con el fin de lograr la interoperatividad con otros trabajos, proporcionan descripciones muy exhaustivas de recursos electrónicos seleccionados por expertos nacionales de diferentes áreas de conocimiento.

Al contrario de lo que ha sucedido en el ámbito hispanohablante, el conjunto de etiquetas DC se ha extendido de forma vertiginosa por varios países, destacando especialmente las áreas geográficas de Oceanía y los países nórdicos<sup>25</sup>.

Según Senso<sup>26</sup> al igual que sucede con la etiqueta META del HTML, DC tiene que combatir contra una serie de factores muy importantes:

- Editores de texto. Ninguno de los editores profesionales que se comercializan en la actualidad permite la inclusión de metadatos en las páginas HTML. Esto hace que los interesados en incluir este tipo de información en sus documentos tengan que acudir a programas complementarios o, en el mejor de los casos, plantillas que se encarguen de la conversión de la información en un formato de metadatos determinado.
- Curva de aprendizaje. La complejidad que tienen las diferentes etiquetas DC, la aparición de los cualificadores, la dificultad para introducir este sistema dentro de RDF, etc. Un análisis de los campos DC propuestos en 1995 y los actuales revela que la complejidad para implementarlos se ha multiplicado por tres. Por el contrario, siguen sin existir guías de referencia que expliquen el uso correcto de las etiquetas y de los cualifica-



dores o las diversas y complejas casuísticas que se pueden dar dentro del mundo electrónico.

- Instituciones implicadas en la introducción de metadatos. Han sido pocos los autores que hayan propuesto soluciones suficientemente eficaces para este problema. Entre otras cosas, y muy probablemente, porque no se trata sólo de un aspecto técnico o de un problema tecnológico, como se apuntaba anteriormente, sino de concienciación sobre la necesidad de las cosas.

## 2. Metadatos en los espacios webs latinoamericanos

Para poder comprobar el estado actual de inclusión de metadatos en el mundo hispanoparlante en América Latina, hemos recogido las metaetiquetas de contenido de los espacios webs de los gobiernos latinoamericanos, incluido el del estado libre asociado de Puerto Rico. Para que la muestra fuera relevante, ya que sólo se analizaba una de las muchas miles de webs por cada país, se han seleccionado las correspondientes a dos instituciones de importancia crucial: las bibliotecas nacionales y los archivos nacionales. Estas instituciones contienen los repositorios documentales más importantes para la conservación, creación y difusión de la identidad nacional y de la historia de cada país, y al mismo tiempo en su conjunto pueden representar otras identidades transnacionales, subcontinentales, culturales, lingüísticas ó étnicas.

Para que las webs seleccionadas fueran todas las posibles dentro de su clase, y conscientes de que a pesar de que se utiliza el mismo idioma, las denominaciones han cambiado, se han buscado los archivos de categoría nacional que en algunos países se denominan "Archivo General de la Nación", o "Archivo Nacional de la República". De esta forma, la muestra de metadatos en sistemas webs aparece más completa, y cabe la posibilidad de establecer mayor número de conclusiones. En la revisión de estos sistemas hemos podido comprobar que también existen, de forma paralela, las denominadas "bibliotecas digitales" o virtuales, separadas de las tradicionales, y con un apoyo estatal importante. La mayoría ofrece textos literarios y políticos completos, de gran interés, pero en su patrimonio no atesoran como sus homólogas nacionales cientos de años de historia documentada, ni tampoco están en todos los países de la muestra.

En la evolución de los sistemas webs se han definido ya las primeras etapas<sup>27</sup>. En lo relativo a los espacios gubernamentales está ya admitido que la presencia o ausencia de dominios propios (<http://www.dominio.xx>) permite analizar la mayor o menor evolución de éstos, y la necesidad de control, modificaciones y actualización de sus contenidos. Todos tienen espacio y dominio propio, o lo que es lo mismo, está superada la primera etapa, que exige presencia independiente en la web.

En la tabla 3 se puede observar la distribución, según elementos (o campos) utilizados, de los metadatos en las webs gubernamentales (obviamente, los países que sus webs no contienen estos elementos no se han incluido). Las palabras-clave (o "keywords") son las más utilizadas, ya que aparecen en 8 de los 16 sistemas webs. Como suele suceder en otras de características similares, no utilizan términos comunes, y los que más se repiten son "gobierno" (6 casos), información (4), ministerios (3) y trámites (2). Ni siquiera todas las que tienen los mismos elementos, como por ejemplo las palabras-clave, incluyen el nombre del país que contiene la web. Sólo en 6 de las 8 aparece el nombre de la nación. Incompresiblemente en Paraguay, el nombre que aparece es el de Bolivia.

Del resto, los elementos que más suelen aparecer son, por orden la descripción, 6 veces (y que suele ser tan somera que apenas se puede denominar así), el autor, lenguaje, copyright, y sorprendentemente la denominada "generator". Los metadatos del sistema web del gobierno guatemalteco, aparecen sólo en inglés, y un contenido se repite en 3 campos distintos, el nombre de la empresa que los ha realizado: "eZsystems", en cambio no incluyen información tan básica como el nombre del país, de su presidente, de las regiones o municipalidades que lo forman.

Hay otros sistemas que aunque no siguen íntegramente la norma ISO sobre metadatos, incluyen unas metaetiquetas muy completas, especialmente si los comparamos con el resto. Son los de Chile, Colombia, México y Perú.

Con respecto a las instituciones documentales de rango nacional como bibliotecas y archivos, esta primera etapa, ya está superada en casi la totalidad de los espacios analizados, con 3 excepciones (Archivo Nacional de Nicaragua, y Paraguay, que no

tienen página propia y el Archivo de Panamá que no tiene presencia en Internet).

Sin embargo, en el marco de la web semántica existen todavía muchos problemas, algunos incluso son fáciles de solucionar. La diferencia entre contenido recuperable o no, dentro del entorno antes citado, es la presencia o ausencia de etiquetas documentales. Pues bien, podemos decir que no son precisamente muy abundantes, ni muy ricas, con excepciones, y que a menudo contienen errores que dificultan la recuperación de los contenidos.

Recogidos los metadatos de archivos y bibliotecas nacionales del ámbito hispano (tabla 4), la estadística refleja que sólo 12 instituciones de todas las analizadas, un 38% aporta etiquetas documentales. A saber, Costa Rica, Cuba, Ecuador, España, El Salvador, Guatemala, Honduras, Nicaragua, Panamá, Puerto Rico y Uruguay. Entre éstas sólo 7 incluyen los elementos, o campos, más frecuentes (keywords y description). De ellas, 4 son Archivos Nacionales y 3 Bibliotecas. Sólo uno de los países analizados tiene metadatos en las dos instituciones, Puerto Rico, pero éstos están repetidos, es decir que no tienen entidad propia, ni elementos específicos para el Archivo y la Biblioteca por separado. Los mejores y más completos son los puertorriqueños, con 12 etiquetas diferentes y 8 con formato Dublin Core, y los pertenecientes al Archivo Nacional de la República de Cuba. En el diseño cubano se han incluido los términos con acento y sin él, muy conscientes de que son uno de los problemas con los que se suelen encontrar los sistemas de navegación y los de búsqueda de datos. La que contiene más etiquetas es la Biblioteca Nacional de Panamá. En la parte negativa, existen instituciones como el Archivo Nacional de Ecuador, que incomprensiblemente incorpora todos los elementos referidos a estética y peluquería, y la ya citada repetición en las dos instituciones documentales nacionales de Puerto Rico.

También se encuentran términos mal escritos como "Bibliota" en la Biblioteca Nacional de Uruguay, con acentos que distorsionen el término, campos mal definidos (`<meta name="" content="Archivo Nacional de la República de Cuba">`), vacías, y con nombres nuevos no normalizados (`<meta name="Biblioteca" content="Guatemala">`; `<meta name="Classification" content="Una lista completa`

de Museos, Archivos y Bibliotecas en Nicaragua, América Central.">).

De las 12 instituciones que contienen estos campos con información semántica válida, tan sólo 7 han añadido el campo "keywords", que contiene las palabras clave que describen el contenido informativo del espacio web. Son, el Archivo Nacional de Cuba, Biblioteca Nacional de Panamá, Biblioteca Nacional de Uruguay, Archivo Nacional de Nicaragua y las dos de Puerto Rico. El Archivo de Ecuador también contiene "keywords", pero son sorprendentemente, como ya se ha señalado, términos como "estilismo, cabello, look, belleza, peluca, postizos, permanente, tinte, rubia, morena, joven, elegancia, moda".

El resto de los campos que más aparecen son los referidos al programa informático con el que se genera el sistema web, "Generator", 5 veces, "Description" 3, "Author" otras 3. El resto contiene elementos que sólo aparecen una vez, como "Originator", "ProgId", "Designer", "Publisher", "Audience", "Classification", "Microsoft Theme" y "Content-Language". De todas las metaetiquetas, sólo Puerto Rico ha elegido el formato Dublin Core, y en algunas se han repetido dos veces (una con el nombre sin formato, y otra con el DC, como por ejemplo, título y descripción).

Con respecto a los términos incluidos en el elemento "Keywords", tampoco hay palabras comunes, con alguna excepción. El máximo de instituciones que coinciden en un término son dos, y casi siempre está el Archivo Nacional de Cuba entre ellas. De hecho, los términos con mayor frecuencia de aparición son los que se pueden incluir en la raíz "archiv" y "document" (en 6 y 5 ocasiones, respectivamente), todos ellos pertenecientes a la citada institución cubana y a las puertorriqueñas, seguido de "informac", contenidos en los respectivos sistemas de Cuba y Panamá, "investig" (en Cuba y Puerto Rico), y con 3 términos distintos las raíces "catalog", y "nacion".

En cuanto al número de términos incluidos en el campo "Keywords", Cuba tienen 76 en la web de su Archivo, Puerto Rico 38 en ambas, Panamá en su Biblioteca 13, Nicaragua 8 en su Archivo y 3 en Uruguay en su Biblioteca. Ecuador llega a incluir 20 términos, pero ninguno relacionado con la institución que los contiene.

El elemento "Author" es el que incluye más variables, ya que se interpreta como la persona o grupo que ha diseñado el sistema web. Por ejemplo, la Biblioteca Nacional de Uruguay lleva el nombre del diseñador, igual que el Archivo Nacional de Nicaragua.

### 3. Conclusiones

Tras un análisis detallado de los espacios webs generados por los gobiernos, bibliotecas y archivos nacionales latinoamericanos podemos afirmar que todavía queda mucho por avanzar en la construcción de la web semántica. Se ha superado, y en ocasiones con una estética propia de los paisajes latinoamericanos, la primera fase, la de presencia en la Web. Se ha consumado esa integración aparente en la que muchos creen, pero que no es ni mucho menos la fase más importante en la actualidad. De lo que ahora se trata es de manifestar esa presencia a través de contenidos, aprovechar el número y los elementos comunes para forzar una web semántica donde el español no sea un idioma secundario, sino uno de los que dan paso y facilitan la reflexión cultural y científica.

Para lograrlo se intuyen ciertos problemas, la mayoría de ellos relacionados con el método utilizado para sumergirse en Internet. Un estudio inicial parece sugerir que los espacios webs son propiedad de las empresas que los construyen más que de los gobiernos que los encargan. Éste es un escollo fácil de salvar siempre que sean los gobiernos los que controlen sus propios espacios. No parece excesivamente entendible que aparezca el autor (la empresa, "eZsystems" en Guatemala), y no se incluya el campo título, o como en el caso de México que el propio autor se genere su etiqueta, y no se acuerde de que la web es gubernamental y pública. Pero aún es más curioso lo que ocurre con dos países como Bolivia y Paraguay. En un ejemplo paradigmático, los metadatos del primero han sido repetidos íntegramente en los del segundo. Así, las palabras-clave que se encuentran en el gobierno de Paraguay son "Portal del Gobierno de Bolivia, Ministerios Bolivia, información". Dicho de otra forma, jamás se encontrará en buscador alguno nada cuando se haga la consulta sobre Ministros de Paraguay, ya que este contenido aparece, como se puede ver entrecomillado es "Ministros Bolivia".

En estudios anteriores sobre esta evolución en los organismos de las Administraciones Públicas españolas hemos podido constatar que el paso a esa segunda etapa es breve, de hecho es sólo cuestión de acordar términos comunes y añadirlos (aunque hay veces que sólo se incluyen, pero que no se acuerdan). Por ello incluimos aquí esta reflexión, para seguir destacando que en la rapidez, autonomía y capacidad de adaptación está la posibilidad de integración, de identificación, de búsqueda de igualdades en un espacio virtual que se hace cada día. La diferencia con la evolución histórica precedente está en que en los espacios virtuales todo se desarrolla mucho más rápidamente, todo avanza a un ritmo frenético, y sólo la capacidad de cambio, de actualización define parámetros de adaptación válidos. La otra gran diferencia es que una vez incluidos en el mundo de los espacios webs, el avance no exige, como en tiempos anteriores, invertir grandes cantidades de dinero, sino avanzar en el desarrollo de herramientas lingüísticas y documentales de apoyo a las tecnologías existentes.

Si bien América Latina, y los pueblos que la forman, son "económicamente" sólo emergente, cultural, artística y filosóficamente sentaron hace mucho tiempo bases difíciles de remover y de ignorar. Bases que ahora sólo es cuestión de saber promocionar y exponer en este mundo virtual tan apasionante como siniestro ("e-exclusion"), y tan real como huidizo. Y sobre todo es una realidad que aunque pretendamos ignorarla va a definir ese futuro cercano en el que queramos, o no, nos van a incluir. La diferencia está en que ahora podemos definir cómo queremos que sea esa "e-inclusión", cómo aparecer, si unidos o separados, como entidad cultural o como entidades históricas. Y ésto es lo que puede determinar si la identidad digital es decidida por otros, o va a ser poco a poco bordada sobre realidades apuntadas por sus propios protagonistas.

### Notas

- 2 La primera vez que se utilizó el término fue en BERNERS-LEE, T. Semantic Web Road map. Sept. 1998, aunque es más conocido el artículo que publicó en el año 2001: BERNERS-LEE, T.; HENDLER, J. & LASSILA, O. The semantic web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American, May 17, 2001.

- 3 MILLAN, J.A. La lengua que era un tesoro. El negocio digital del español y cómo nos quedamos sin él [web]. [31-05-2005]. Disponible en <<http://jamillan.com/tesoro.htm>>.
- 4 La Sociedad de la información en América Latina. Revista Telos [web], Monográfico oct.-dic. 2004, nº 61.[Consulta 15-10-2005]. Disponible en <<http://campusred.net/telos>>.
- 5 Ibidem.
- 6 Sobre este tema, Albornoz ha hecho una buena síntesis en Telos 2004.
- 7 MILLAN, J.A. Internet y el español. Madrid: Fundación Retevisión, 2001. p. 26.
- 8 DONOSO, R. Metadata: Herramienta para la recuperación de información en Internet. Taller de Tecnología de redes Internet para América Latina y el Caribe [web], INET'98. [Consulta: 27-06-2005]. Disponible en <<http://biblioteca.udea.edu.co/~hlopera/metadata.html>>.
- 9 Sobre este tema puede consultarse la bibliografía reciente: CHAÍN NAVARRO, C. y TOMAS CARRIÓN, P. Análisis comparativo del desarrollo de la administración local española en Internet durante el año 2004 en las 10 ciudades más pobladas. Anales de Documentación [Web], 2005, 8, p. 37-49; CHAÍN NAVARRO, C.; BARBA MERINO, M.L. y GARRIDO LOVA, J. Recuperación de Información en los Sistemas Web de las Administraciones Públicas Españolas: Análisis de los metadatos documentales. ICOM'2004 (La Habana). [cd-rom].
- 10 ANTON BRAVO, A. FOAF: Friend of a friend o cómo construir redes sociales en la web semántica. Web semántica hoy [web] 2004. [Consulta 15-10-2005] Disponible en <<http://www.wshoy.sidar.org/index.php?2004/12/03/17foaf-fiend-of-a-frieno-o-como-construir-redes-sociales-en-la-web-semántica>>.
- 11 DOCTOROW, C. Metacrap: Putting the torch to seven straw-men of the meta-utopia [web] versión 1.3 26 August 2001. <<http://www.well.com/~doctorow>>
- 12 Sobre este tema CHAÍN NAVARRO, C.; BARBA MERINO, M.L. y GARRIDO LOVA, J. Recuperación de Información en los Sistemas Web de las Administraciones Públicas Españolas: Análisis de los metadatos documentales. ICOM'2004 (La Habana)[cd-rom].
- 13 DEMPSEY, L; HEERY, R. Metadata: a current view of practice and issues. Journal of Documentation, 1998, 54,2, p.145-172.
- 14 MENDEZ RODRÍGUEZ, E.M. El documento como objeto de información: metadatos y esquemas de organización de la información en la Web. En: ZAPICO ALONSO, F. (ed.). Contenidos digitales; creación, descripción y evaluación de recursos informativos. Mérida: Junta de Extremadura, 2002.
- 15 SENSO, J.A. Y DE LA ROSA PINERO, R. Dublin Core Metadata Initiative: Norma internacional para la descripción de recursos electrónicos (ISO 15836:2003). Boletín de la ANABAD, 2003, p. 25-56. Y SENSO, J.A. Y DE LA ROSA PINERO, R. El concepto de metadato. Algo más que descripción de recursos electrónicos. Ciencias de la Informação, 2003, 32,2, p. 95-106.
- 16 Ibidem.
- 17 Los datos en este campo pueden ser una URI, URL, o un protocolo.
- 18 DCMI. Dublin Core Metadata Initiative - Home Page [web] 10 octubre 1998. [Disponible en <<http://purl.org/DC/index.htm>>]
- 19 WEIBEL, Stuart. Metadata: the foundations of resource description. D lib magazine [web], julio 1995. [Disponible en <<http://www.diglib.org/dlib/July95/07weibel.html>>]
- 20 SENSO, J.A. Y DE LA ROSA PINERO. Dublin Core Metadata Initiative: Norma internacional para la descripción de recursos electrónicos (ISO 15836:2003). Boletín de la ANABAD, 2003, p. 25-56.
- 21 American National Standards Institute (ANSI/NISO). ANSI/NISO Z39.85-2001. The Dublin Core Metadata Element Set. [Web]. [Consulta 17/11/2004]. Disponible en <<http://www.niso.org/standards/resources/Z39-85.pdf>>.
- 22 Internacional Standard Organisation (ISO). ISO 15836/2003 The Dublin Core Metadata Element Set. [Web]. [Visitado 17/11/2004]. Disponible en <<http://www.niso.org/international/SC4/n515.pdf>>].
- 23 La web semántica en América Latina [web]. 2004. [Consulta 04-06-2005]. Disponible en <<http://www.w3.org/2001/sw/Europe/events/200408-eze>>.
- 24 RedIris. Proyecto SARAC [web]. [Consulta 12-10-2005]. Disponible en <<http://sarac.rediris.es>>.
- 25 Se puede acceder a una lista actualizada de estos proyectos en: <http://dublincore.org/projects/> >.
- 26 SENSO, J.A. Y DE LA ROSA PINERO. Dublin Core Metadata Initiative: Norma internacional para la descripción de recursos electrónicos (ISO 15836:2003). Boletín de la ANABAD, 2003, p. 25-56.
- 27 CHAIN NAVARRO, C. y MUÑOZ CAÑAVATE, A. Análisis comparativo regional del desarrollo de la Administración Local española en Internet (1997-2002). Revista de Bibliotecología, 2004, 36, 8, p. 96-117; CHAÍN NAVARRO, C. & MUÑOZ CAÑAVATE, A. Hacia el cibergobierno: los ayuntamientos de la Región de Murcia en Internet (1997-2002). Anales de

Documentación [web], 2004, 7, p. 55-68. Disponible en <<http://www.um.es/fccd/anales/ad07/ad0704.pdf>>; MUÑOZ CAÑAVATE, A. & CHAÍN NAVARRO, C. The World Wide Web as an information system in Spain's Regional Administrations (1997-2000). *Government Information Quarterly*, 2004, 2, p. 199-218.

- 28 MUÑOZ CAÑAVATE, A. & CHAÍN NAVARRO, C. La administración local española en Internet: estudio cuantitativo de la evolución de los sistemas de información web de los ayuntamientos (1997-2002). *Ciencias de la Información*, 2004, 35, 1, p. 43-55.

## Bibliografía

BALAGUER PRESTES, R. *Internet: un nuevo espacio psicosocial*. Montevideo: Trilce, 2003.

BECERRA, M. *La Sociedad de la Información: proyecto, convergencia, divergencia*. Buenos Aires: Norma editorial, 2003.

BUSTAMANTE, E. (coord.). *Comunicación y cultura en la era digital. Industrias, mercados y diversidad en España*. Barcelona: Gedisa, 2002.

CAÑIZALES, A. ¿Políticas públicas en tiempos de una comunicación globalizada?. *Anuario ININCO*, 14. Caracas: Universidad Central de Venezuela, 2002. p. 143-159.

CASTELLS, M. *La era de la información: economía, sociedad y cultura*. México: Siglo XXI, 2002. 3 vols.

CEPAL. *Declaración de Bávaro: informe de la Conferencia Regional de América Latina y el Caribe para la Comisión Mundial sobre la Sociedad de la Información*. *Diálogos de la Comunicación*, 2003, 67, p. 77-84.

CHAIN NAVARRO, C. *Técnicas y métodos de recuperación de información*. Murcia: DM, 2004.

CHAÍN NAVARRO, C. Propuesta de un modelo de análisis de contenido de los sistemas webs de las administraciones locales españolas. *Revista Española de Documentación Científica*, 2005 (en prensa).

CHAÍN NAVARRO, C. & MUÑOZ CAÑAVATE, A. The World Wide Web as an information system in Spain's Regional Administrations (1997-2000). *Government Information Quarterly*, 2004, 2, p. 41-63.

CHAÍN NAVARRO, C. & MUÑOZ CAÑAVATE, A. La administración local española en Internet: estudio cuantitativo de la evolución de los sistemas de información webs de los ayuntamientos (1997-2002). *Ciencias de la Información*, 2004, 35, 1, p. 43-55.

DEMPSEY, L & HEERY, R. Metadata: A Current View of Practice and Signes. *Journal of Documentation*, marzo 1998, vol. 54, nº 2, pp. 145-172.

DONOSO, R. Metadata: Herramienta para la recuperación de información en Internet. Taller de Tecnología de redes Internet para América Latina y el Caribe, INET'98 [web]. [27-06-2005]. Disponible en <<http://biblioteca.udea.edu.co/~hlopera/metadata.html>>.

GARCIA CANCLINI, N. *Latinoamericanos buscando lugar en este siglo*. Buenos Aires: Paidós, 2002.

GARCIA CANCLINI, N. y MONETA, C.J. (coords.). *Las industrias culturales en la integración latinoamericana*. Buenos Aires: EUDEBA/SELA, 1999.

GUY, M. POWELL, A. & DAY, M. Improving the Quality of Metadata in Eprint Archives. *Ariadne* [web], 2004, 36. [Consulta: 27-06-2005]. Disponible en <<http://www.ariadne.ac.uk/issue38/guy.html>>.

HAMEL, R.E. El español como lengua de las ciencias frente a la globalización del inglés. Conferencia dictada en la Universidad Autónoma Metropolitana-Iztapalapa. México, mayo 2003.

HEERY, R. & JOHNSTON, P. Metadata schema registries in the partially Semantic Web: the CORES Experience. *Proceedings of the 2003 Dublin Core Conference* [web]. Seattle (Wa). 28 sept.-2 oct. 2003. [Consulta: 27-06-2005]. Disponible en <<http://purl.oclc.org/dc2003/03heery.pdf>>.

HERZOG, R.; HOFFMAN, B. y SCHULT, M. Internet und politik in Lateiamerika [web] 2002. [Consulta: 15-10-2005]. Disponible en <<http://www.rz.uni-hamburg.de/Iik/nikt/comparaciones.pdf>>.

LOZADA, M. El ciberciudadano: representaciones, redes y resistencias en Venezuela y América latina. En MATO, D. *Políticas de ciudadanía y sociedad civil en tiempos de globalización* [web]. Caracas: Universidad Central de Venezuela, 2004. [Consulta 10-10-2005]. Disponible en <<http://www.globalcult.org.ve/pub/Rocky2.htm>>.

MARTÍNEZ, G. Internet y ciudadanía global: proceso de producción de representaciones sociales de la ciudadanía en tiempos de globalización. En MATO, D. *Políticas de ciudadanía y sociedad civil en tiempos de globalización*. Caracas: Universidad Central de Venezuela, 2004. [Consulta 10-10-2005]. Disponible en <<http://www.globalcult.org.ve/pub/Rocky2.htm>>.

MASTRINI, G. y BOLAÑO, C. (ed.). *Globalización y monopolios de la comunicación de América Latina*. Buenos Aires: Biblos, 2000.

- MÉNDEZ RODRÍGUEZ, E.M. Metadatos y recuperación de información. Madrid: Trea, 2000, 220 p.
- MILLAN, J.A. Internet y el español. Madrid: Fundación Retevisión, 2001. 289 p.
- MILLAN, J.A. La lengua que era un tesoro. El negocio digital del español y cómo nos quedamos sin él [web]. [Consulta 31-05-2005]. Disponible en <<http://jamillan.com/tesoro.htm>>.
- MILLAN, J.A. El español en la sociedad digital: una propuesta. III Congreso internacional de la Lengua española. El español en la sociedad de la información [web]. [Consulta 31-05-2005]. Disponible en <[http://cvc.cervantes.es/obref/congresos/valladolid/mesas\\_redondas...](http://cvc.cervantes.es/obref/congresos/valladolid/mesas_redondas...)>.
- MUÑOZ CAÑAVATE, A. y CHAIN NAVARRO, C. Análisis comparativo regional del desarrollo de la Administración Local española en Internet (1997-2002). Revista de Bibliotecología, 2004, 36, 18 p. 96-116.
- NARVAEZ, A. Puentes tecnológicos, abismos sociales. Manizales: Universidad, 2002.
- NISO. Understanding Metadata. Bethesda (MD): NISO Press, 2004.
- ROMANO, S.S. La cultura digital. Navegantes de Internet, personalidades interactivas y agrupamientos virtuales. Buenos Aires: Lugar Editorial, 2000.
- SENSO, J.A. Y DE LA ROSA PINERO, R. Dublin Core Metadata Initiative: Norma internacional para la descripción de recursos electrónicos (ISO 15836: 2003). Boletín de la ANABAD, 2003, p. 25-56.
- Sociedad de la información en América Latina. Revista Telos [web], Monográfico oct.-dic. 2004, n° 61 [Consulta 15-10-2005]. Disponible en <<http://campusred.net/telos>>.
- TREMBLAY, G. Panam. Industrias culturales y diálogo de civilizaciones en América. Montreal: Presses de l'Université Laval, 2003.
- TREJO DELARBE, R. La Internet como bien público. Diálogos de la Comunicación, 2003, 67, p. 29-39.
- W3C. Namespaces in XML [web]. Tim Bray, Dave Hollander y Andrew Layman (eds.). 14 de enero de 1999. [Consulta 16-04-2004]. Disponible en <<http://www.w3.org/TR/REC-xml-names>>.
- W3C. Resource Description Framework (RDF): Model and Syntax Specification. W3C Recommendation, 22 February 1999 [web]. [Consulta 16-04-2004]. Disponible en <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>.
- W3C. XML Information Set. W3C Recommendation 24 Octubre 2001 [web], Jonh Cowan, Richard Tobin, eds. W3C, 24 de octubre del 2001, [Consulta 16-04-2004]. Disponible en <<http://www.w3.org/TR/2001/REC-xml-infoiset-20011024>>.
- W3C. RDF/XML Syntax Specification (Revised). W3C Working Draft 25 March 2002. Davve Beckett, ed. W3C, 25 de marzo del 2002. [Consulta 16-04-2004]. Disponible en <http://www.w3.org/TR/2002/WD-rdf-syntax-grammar-20020325>.
- W3C. Feature Synopsis for OWL Lite and OWL W3C Working Draft 29 July 200 [web]. Deborah L. McGuinness, Frank van Harmelen, (eds.). W3C, 29 de julio del 2002. [Consulta 16-04-2004]. Disponible en <<http://www.w3.org/TR/2002/WD-owl-features-20020729>>.
- W3C. OWL Web Ontology Language 1.0 Reference. W3C Working Draft 29 July 2002 [web]. Mike Dean, Dan Connolly, Frank van Harmelen, James Hendler, Ian Herrick's, Deborah L. McGuinness, Peter F. Patel-Schneider y Lynn Andres Stein, eds. W3C, 29 de julio del 2002. [Consulta 16-04-2004]. Disponible en <<http://www.w3.org/TR/2002/WD-ol-ref-20020729>>.
- WEB semántica en América Latina, La [web]. Curso celebrado en Buenos Aires, 21 y 22 agosto 2004. [Consulta 04-06-2005]. Disponible en <<http://www.w3.org/2001/sw/Europe/events/200408-eze>>.
- WEIBEL, S. Metadata: the foundations of resource description. D-Lib Magazine [web]. 1995. [Consulta 16-04-2004]. Disponible en [www.dlib.org/dlib/July95/07weibel.html](http://www.dlib.org/dlib/July95/07weibel.html)>.
- WEIBEL, S. Dublin Core Metadata Element Set: reference description [web]. 1997. [Consulta 16-04-2004]. Disponible en <[http://www.purl.org/metadata/dubli\\_core\\_elements](http://www.purl.org/metadata/dubli_core_elements)>.
- WEIBEL, S. Approval of initial Dublin Core interoperability qualifiers [web]. [Consulta 16-04-2004]. Disponible en <[www.mailbase.ac.k/lists/dc-general/2000-04/0010.html](http://www.mailbase.ac.k/lists/dc-general/2000-04/0010.html)>.
- WEIBEL, S. y HAKELA, J. DC5: The Helsinki Metadata Workshop: A Report on the Workshop and Subsequent Developments. D-Lib Magazine [web]. 1998. [Consulta 16-04-2004]. Disponible en <<http://www.dlib.org/dlib/february98/02weibel.html>>.
- WOLTON, D. Internet, ¿y después?. Barcelona: Gedisa, 2002.