

Metadatos Dublin Core para la Web chilena

Marcela Calderón, Claudio Gutiérrez
Centro de Investigaciones de la Web CIW
Depto. de Ciencias de la Computación
Universidad de Chile
Blanco Encalada 2120, Santiago, Chile
{mcaldero,cgutierrez}@dcc.uchile.cl

Resumen

La incorporación sistemática de metadatos se está convirtiendo en una de las principales técnicas para mejorar la descripción, ubicación y recuperación de documentos en la Web. Este artículo presenta el estándar internacional de metadatos *Dublin Core*, delinea el estado del arte sobre él, discute sus ventajas y perspectivas para su utilización en Chile, y propone un plan de trabajo para difundir este estándar en Chile. El plan está centrado en una herramienta de marcado semiautomático con Dublin Core que actualmente se está desarrollando.

Palabras Claves: Web, Metadatos, Dublin Core, Chile.

1 El problema

La *Web* (*World Wide Web*) ha remecido los fundamentos de la publicación, distribución y procesamiento de la información. Desde este punto de vista, el impacto de la Web va más allá de la ciencia de la computación y la informática. Como ejemplos podemos mencionar la información, las revistas y bibliotecas en línea (*on-line*), museos y galerías en línea, comercio electrónico y servicios, gobierno electrónico y administración, sitios personales e institucionales, y últimamente, la posibilidad de acceder gigantescas bases de datos con información digital, como datos geográficos, genéticos, observaciones astronómicas. Debido a su extraordinario crecimiento diario, manejar y procesar la información de la Web se ha tornado un asunto crítico. Por ejemplo, la Web contiene hoy del orden de petabytes (1 PB = 1,000 terabytes) de información (se estima que la Biblioteca del Congreso de EE.UU. tiene 20 terabytes en libros), más de 550.000 millones de documentos individuales y más del 95% de la información está accesible al público sin necesidad de pagos o suscripciones [1], [2]. Sobre todo este cuadro, la Web muestra una extraordinaria heterogeneidad en formato, tipo y uso de diferente tipo de información como texto, sonido, gráfica, video, industrial, comercial, académico, gubernamental, científico, particular, etc.

Todas estas tendencias ponen nuevos desafíos al procesamiento de información en la Web. No es sencillo estar a la altura de la velocidad de evolución de la Web. Hace diez años el centro del problema eran aplicaciones en torno a protocolos como TCP/IP y HTTP, problemas de conectividad física, de interoperabilidad al nivel más básico de datos. Luego vino el fenómeno que puso la Web en el centro de la atención mundial, esto es, el desarrollo de aplicaciones en torno a servicios y usuarios, como comercio electrónico, servicios web, gobierno electrónico, sitios personales e institucionales, etc. La consecuencia directa de este boom fue la explosión de información en la Web.

Para los profesionales de la información, el principal desafío hoy es cómo manejar esta extraordinaria cantidad de datos que crece día a día. Estamos comenzando a ver los problemas: los motores de búsqueda a menudo no contestan lo que estamos buscando; hay dificultades para filtrar la información; la heterogeneidad de los datos y los contenidos; desde el punto de vista de quien publica el problema se ha convertido en hacer la información visible, tanto en formato como en contenido. Han habido avances en los niveles estructurales y sintácticos con el estándar *de facto* XML y sus tecnología asociadas [3]. Desafortunadamente, al nivel del significado (semántica) aún estamos muy por debajo de las necesidades. Estamos lejos de responder preguntas como “todos los museos que exhiben trabajos de Guayasamín”; o “¿Cuál es la biblioteca que tiene la mejor colección de los escritos de Gandhi?”, o “¿Cuál es la compañía que ofrece el mejor mapa de Isla de Pascua desde el punto de vista precio/resolución? Un motor de búsqueda estándar (como Google, Altavista, AlltheWeb) no puede responder tales consultas. Pero tampoco ningún agente las podría responder hoy día. Sin embargo, la información está allí: sólo necesitamos relacionarla y agregarla.

La limitación obedece a la falta de habilidad de las máquinas para entender el significado y las relaciones entre las partes de información que recolectan. Hoy día los humanos agregamos el contexto, interpretamos y damos sentido a la información que existe en la Web. En una frase: debido a las enormes dimensiones, la Web se ha convertido en una torre de Babel, no sólo al nivel del lenguaje natural, sino esencialmente al nivel del significado, contradiciendo las ideas por las cuales fue creada [4, 5]. ¿La solución? Pavimentar el camino para la construcción de agentes de software que puedan procesar información de la Web por nosotros. En esta perspectiva los metadatos juegan un rol fundamental.

2 Metadatos en la Web

Para automatizar el proceso de contextualizar, interpretar y dar sentido a la información que existe en la Web, uno de los pilares básicos son los estándares de *metadatos* y las *ontologías*, que permitirán que agentes de software encuentren el significado de la información en páginas Web siguiendo hiperlinks a las definiciones de términos claves y reglas para razonar acerca de ellas lógicamente.¹

El concepto de metadatos es antiguo, y de hecho, en la Web existen y circulan infinidad de tipos de metadatos, propietarios (comercio, industria, etc.), algunos con algún grado de estandarización como por ejemplo GILS (Government Information Locator Service) [13], ANZLIC (Australia New Zealand Land Information Council) [14], AGLS (Australian Government Locator Service) [15], EdNA (Education Network Australia) [16]. Finalmente hay propuestas de estándares globales entre las que destaca un proyecto minimalista, Dublin Core, del que hablaremos *in extenso* en este artículo, y el proyecto más ambicioso de la W3C, el *Resource Description Framework*, RDF [6]. Estos metadatos son sectoriales y usan una diversidad de modelos y lenguajes. Hoy día ya estamos comenzando a ver las dificultades de interoperabilidad en este nivel.

Es crucial la atención sobre un aspecto que marca una diferencia fundamental en el tratamiento clásico de metadatos y los metadatos en la Web. Clásicamente, la generación y administración de metadatos está fuertemente centralizada en especialistas y organizaciones ad hoc. Con los volúmenes de información en la Web y los porcentajes de crecimiento, este enfoque es impensable implementarlo de manera global. La única vía para generar y administrar metadatos es hacerlo de manera descentralizada (distribuida). Esto define los desafíos que se deben enfrentar y las características de los metadatos y herramientas que necesitamos.

Aunque la sintaxis no es estrictamente parte del esquema de metadatos, los datos serán inútiles a

¹ Los *metadatos* son información descriptiva acerca de un objeto o recurso, sea éste físico o electrónico. Las *ontologías* son especificaciones formales de vocabularios y conceptos compartidos para un dominio.

menos que el esquema de codificación refleje la semántica de los metadatos. La codificación permite a los metadatos ser procesados automáticamente por agentes de software. Entre los esquemas de codificación más importantes tenemos HTML, XML [3], RDF [6], MARC (MACHine Readable Cataloging) [17], MIME (Multipurpose Internet Mail Extensions) [28], Z39.50 [18], LDAP (Lightweight Directory Application Protocol) [29].

También es importante definir la forma en que los metadatos se asociarán con el recurso correspondiente desde el punto de vista físico. Esto se puede hacer: incrustando el metadato en la misma página Web que contiene los datos, por el creador de la página usando las etiquetas META de HTML; como un documento HTML separado enlazado al recurso que describe; en una base de datos enlazada a los recursos que describen sus datos.

3 Dublin Core

El estándar de metadatos de Dublin Core [20], que fue diseñado con criterios de minimalidad, es un conjunto simple que abarca quince términos (denominados elementos) para describir un amplio rango de recursos en la Web. La semántica de dichos elementos ha sido establecida a través de consensos por un grupo interdisciplinario internacional de profesionales de bibliotecología, ciencias de la computación, codificación de texto, la comunidad de museos, y otros campos relacionados. Dicho de otra forma, Dublin Core es un pequeño lenguaje para hacer declaraciones básicas acerca de recursos. Se trata de un sistema de codificación de metadatos que necesita todavía de la intervención humana. En definitiva, es un formato muy simple que incluso puede ser aplicado por catalogadores no muy expertos.

El formato Dublin Core nace para tratar de proporcionar metadatos para los recursos accesibles en la Web, y se ha gestado bajo el patrocinio de OCLC (Online Computer Library Center) [21] a partir de 1995, que ha pretendido avanzar en su nueva concepción de los catálogos a través de las páginas Web. El DCMI (*Dublin Core Metadata Initiative*) [20], la organización que impulsa Dublin Core, tiene como objetivo (1) Desarrollar estándares de metadatos para el descubrimiento de información a través de dominios; (2) Definir marcos de trabajo para la interoperación de conjuntos de metadatos; (3) Facilitar el desarrollo de conjuntos específicos de metadatos, tanto comunitarios como por disciplinas, que sean consistentes con los puntos 1 y 2.

Entre los logros y perspectivas de Dublin Core podemos mencionar:

–*Estandarización.* La adopción de Dublin Core, particularmente en el sector público de diversos países, ha hecho que la estandarización formal sea un asunto de suma importancia. En el año 2000, se adoptó la versión 1.1 del *element set* por el CEN (Comité Europeo de Normalización, [7]). En el año 2001, fue ratificada una versión modificada del *element set* como estándar ANSI, bajo los auspicios de NISO (National Information Standards Organization) de Estados Unidos.

–*DCMI y IEEE-LOM (IEEE-Learning Object Metadata)* [8]. Desde el año 2000 se han venido identificando áreas de trabajo, principios y prácticas comunes entre estas dos importantes organizaciones.

–*DC y OAI (Open Archive Initiative)* [9] Se avanza en el trabajo de asociación del protocolo del OAI [9] con Dublin Core.

–*Actividad de la Web Semántica en W3C* [10]. El lanzamiento de la actividad de la Web Semántica por el W3C, reconoce la creciente importancia de soportar la infraestructura para definir, registrar y referenciar vocabularios estructurados y ontologías en la Web. Dublin Core es parte importante de esta infraestructura, y la comunidad DCMI ha desempeñado un papel fundamental en fundar las bases de este trabajo.

3.1 Descripción técnica de Dublin Core

Dublin Core fue creado con las siguientes características en mente:

1. *Simplicidad de creación y mantención:* El conjunto de elementos de Dublin Core se ha mantenido lo más pequeño y simple posible, para permitir a los no-especialistas crear registros descriptivos simples para los recursos de información.
2. *Semántica comúnmente entendida:* Dublin Core puede ayudar al no-especialista a encontrar lo que busca, soportando un conjunto común de elementos cuya semántica es entendida y soportada universalmente. Tal convergencia en un campo común aumenta la visibilidad y accesabilidad de todos los recursos, tanto los que pertenecen a una disciplina específica como los que no.
3. *Alcance internacional:* El conjunto de elementos de Dublin Core fue desarrollado originalmente en inglés, pero han sido creadas versiones en muchos otros idiomas. Existe un grupo especial de interés en este tema que coordina esfuerzos para juntar estas versiones en un registro distribuido usando la tecnología RDF (Resource Description Framework) desarrollada por W3C (World Wide Web Consortium). La participación de representantes de casi todos los continentes, ha asegurado que el desarrollo del estándar considere la naturaleza multilingüe y multicultural del universo de la información electrónica.
4. *Extensibilidad:* Se espera que otras comunidades de expertos en metadatos crearán y administrarán conjuntos adicionales de metadatos. Los elementos de estos conjuntos podrían ser enlazados con los metadatos Dublin Core para satisfacer la necesidad de extensibilidad. Este modelo permite que diversas comunidades usen los elementos de Dublin Core para la información descriptiva básica que será usable a través de la Web, mientras admite extensiones de dominios específicos, que toman sentido dentro de campos más limitados.

3.1.1 Descripción de los elementos de Dublin Core

Los elementos básicos son quince y se agrupan de acuerdo a tres aspectos:

| CONTENIDO | PROPIEDAD INTELECTUAL | INSTANCIA |
|-------------|-----------------------|------------------|
| Título | Autor o creador | Fecha |
| Materia | Editor | Tipo del recurso |
| Descripción | Otros colaboradores | Formato |
| Fuente | Derechos | Identificador |
| Lengua | | |
| Relación | | |
| Cobertura | | |

Tabla 1: Los quince elementos de Dublin Core

Los elementos poseen nombres descriptivos que pretenden transmitir un significado semántico a los mismos. Para promover una interoperabilidad global, una descripción del valor de algunos elementos podrá ser asociada a vocabularios controlados². Cada elemento es opcional, puede repetirse,

²Un vocabulario controlado es una lista establecida de terminología estandarizada para ser usado en la indexación y recuperación de la información. Un vocabulario controlado asegura que un tema será descrito usando el mismo

aparecer en cualquier orden y tener asociados cuantos enlaces se considera necesarios. Tomando en consideración esta capacidad para describir estas fuentes con mayor precisión, han sido ideados los *calificadores* relativos a cada elemento. Los calificadores permiten aumentar la especificidad y precisión de los metadatos.

3.1.2 Descripción de los calificadores (Qualifiers)

En la determinación de los calificadores, se dio preferencia a vocabularios, notaciones y términos actualmente mantenidos por agencias establecidas. Se espera que los implementadores desarrollen calificadores adicionales para sus propios dominios específicos. Sin embargo, los calificadores que están de acuerdo a los principios de calificación de Dublin Core serán más reusables por otras comunidades.

El DCMI reconoce dos grandes clases de calificadores:

1. *Refinación de elementos*: Estos calificadores hacen que el significado de un elemento sea más estrecho o más específico. Un elemento refinado comparte el significado del elemento no calificado, pero con un alcance más restrictivo.
2. *Esquema de codificación*: Estos calificadores identifican esquemas que ayudan en la interpretación del valor de un elemento. Estos esquemas incluyen vocabularios controlados y notaciones formales o reglas de parseo.

En la Tabla 2 se presenta un resumen de los elementos y sus calificadores respectivos.

3.1.3 Ejemplo

```
<meta name    = "DC.Title"
      content = "Canto General">
<meta name    = "DC.Title.Alternative"
      content = "Canto Gral.">

<meta name    = "DC.Creator"
      content = "Neruda, Pablo">

<meta name    = "DC.Subject"
      scheme   = "LCC"
      content = "PQ8097.N4 C35">

<meta name    = "DC.Subject"
      scheme   = "DCC"
      content = "861.6">

<meta name    = "DC.Description.TableofContents"
      content = "La lámpara en la tierra.--Alturas de Macchu
                Picchu.--Los conquistadores.--Los libertadores.
                --La arena traicionada.--América, no invoco tu
                nombre en vano.--Canto general de Chile.--La
```

término cada vez que es indexado, esto permite encontrar de manera más fácil, toda la información sobre un tópico específico, durante el proceso de búsqueda.

| Elemento | Refinación | Codificación |
|---------------------|---|---|
| Título | Alternativa | - |
| Autor o creador | - | - |
| Materia | - | LCSH MeSH DDC LCC UDC |
| Descripción | Tabla de Contenidos Resumen | - |
| Editor | - | - |
| Otros colaboradores | - | - |
| Fecha | Creado Válido Disponible Publicado Modificado | DCMI Period W3C-DTF |
| Tipo del recurso | - | DCMI Type Vocabulary |
| Formato | Extensión Medio | IMT |
| Identificador | - | URI |
| Fuente | - | URI |
| Lengua | - | ISO 639-2 RFC 1766 |
| Relación | Es la versión de Tiene la versión Es sustituido por Sustituye Es requerido por Requiere Es parte de Tiene parte Es referido por Referencias Es formato de Tiene el formato | URI |
| Cobertura | Espacial | DCMI Point ISO 3166 DCMI Box TGN |
| | Temporal | DCMI Period W3C-DTF |
| Derechos | - | - |

Tabla 2: Calificadores de Dublin Core

```
tierra se llama Juan.--Que despierte el leñador.
--El fugitivo.--Las flores de Puitaqui.--Los
ríos del canto.--Coral de Año Nuevo para la
patria en tinieblas.--El gran océano.--Yo soy.">
```

```
<meta name      = "DC.Date.Issued"
      scheme    = "W3C-DTF">
      content    = "1950">
```

3.2 DC en el mundo real: Perspectivas

Son muchos los organismos e instituciones que tratan de definir y delimitar los formatos adecuados para poder insertar o publicar un documento en la Web. Estos formatos se deben emplear en el momento en que se introduce la información para su localización posterior. Así, organismos bibliotecarios nacionales, internacionales y centros de normalización tratan de definir estos formatos que reportarán un gran acceso a la información bibliotecaria, bibliográfica y otras. Para una lista completa de los proyectos que usan Dublin Core ver [11].

La aplicación práctica del formato Dublin Core se ha proyectado en varias grandes bibliotecas como la del Congreso de Washington, la Biblioteca Nacional de Australia y la Biblioteca Nacional de Nueva Zelanda, que han tratado de integrar los elementos del Dublin Core en sus catálogos. Otra aplicación práctica de interés, es la realizada por la unión de catálogos en línea de OCLC o el denominado WorldCat, que emplea como interfaz de acceso los elementos establecidos por Dublin Core.

La descripción formal para metadatos de Dublin Core con calificadores en RDF (Resource Description Framework) y la forma de extenderlos ha sido uno de los puntos más importantes de trabajo. El trabajo de RDF se realiza bajo los auspicios del Consorcio de la Web W3C, la organización que trata la estandarización en la Web. Esta estructura de metadatos pretende lograr una interoperabilidad de intercambio entre máquinas.

Entre los trabajos a futuro, podemos mencionar el trabajo en extensiones y perfiles de aplicación de comunidades educacionales, de gobierno y de bibliotecas, la expresión de metadatos de Dublin Core en XML, el enlace de agentes (referir datos estructurados externos como alternativa o complemento al valor literal de un elemento), los calificadores CCP, calificadores para los elementos Creador, Otros colaboradores y Editor (*Creator, Contributor, Publisher*). Por otra parte, se trabaja en la aceptación de Dublin Core versión 1.1 como estándar ISO, mediante un proceso de vía rápida, y se está elaborando una nueva guía para el usuario.

4 Metadatos en la Web chilena

La Web chilena comparte las mismas tendencias internacionales respecto a la presencia de metadatos: su uso sistemático y estandarizado es casi inexistente. También en Chile, hoy en día (2002), los metadatos Dublin Core son escasamente utilizados. De hecho, SISIB (Sistema de Servicios de Información y Bibliotecas de la Universidad de Chile SISIB [23]) los usa para marcar páginas de unidades internas de SISIB y para marcar obras que estén disponibles por Internet en la biblioteca digital. En este momento los metadatos son usados sólo para marcar con la perspectiva que surja un buscador que los utilice. Aunque RENIB (Red Nacional de Información Bibliotecas [24]) y DIBAM (Dirección de Bibliotecas y Museos [25]) no lo usan por ahora, Dublin Core es la tendencia actual en el mundo de las bibliotecas, por lo que es de esperar que en un futuro cercano lo adopten.

Pensamos que esta falta de Dublin Core y de metadatos globales en general en la Web chilena se debe principalmente al desconocimiento que existe de este formato y estas tecnologías. Se presenta aquí una nueva versión de las “dos culturas”: por una parte, los profesionales del área de bibliotecología, catalogación y servicios de información no están hoy día directamente involucrados en la creación y recuperación de información en la Web; por otra parte, los ingenieros y técnicos que trabajan en la Web, creando páginas, enlazando bases de datos, etc., aún no tienen familiaridad con los conceptos y tecnologías asociadas a los metadatos. Tampoco existe trabajo en conjunto entre los generadores de páginas Web y los bibliotecarios dentro de una misma organización (ver al respecto [31]). Finalmente, no existen herramientas que acerquen o faciliten al usuario común (no experto) el uso de estas tecnologías. Con esto nos referimos tanto al marcado con metadatos como a la búsqueda y recuperación de información marcada con metadatos.

4.1 Dublin Core para la Web chilena

Hemos argumentado suficientemente respecto de las ventajas de incorporar metadatos estándares para la documentación en la Web chilena. Se puede diferir en las metodologías para implantarlos, en las técnicas para codificarlos. Creemos que una de las maneras más eficientes de acercar estas tecnologías al grueso de los usuarios de la Web chilena, es incorporando metadatos livianos (i.e. simples, breves, fáciles de crear y usar). Dublin Core es una muy buena elección en esa dirección. Desgraciadamente, aún no existen herramientas que aprovechen estos metadatos masivamente para la recuperación de información. Entre los buscadores que usan Dublin Core mencionamos anteriormente el proyecto de la OCLC [30]. Pero sin duda la dificultad mayor es la falta de herramientas adecuadas que faciliten el marcado. En esta dirección hay algunas experiencias (marcado asistido, generadores automáticos de metadatos Dublin Core, por ejemplo DC-dot [22]).

Hay dos aspectos íntimamente relacionados con la difusión del estándar Dublin Core para la Web chilena: la difusión de la necesidad de contar con metadatos, y la construcción de herramientas que faciliten este proceso. Algunos avances en esa dirección:

1. *Herramienta de marcado.* Hoy en día no existe ninguna herramienta que genere metadatos de Dublin Core automáticamente a partir de documentos escritos en castellano. Hacemos notar que existen dificultades técnicas que hacen que herramientas diseñadas para otros idiomas no sirvan en castellano: (1) El procesamiento sintáctico previo al análisis semántico (Extracción automática de sustantivos relevantes, eliminación automática de artículos, adverbios, etc.), (2) Lematización (obtención de raíces y formas normales de los vocabularios), y (3) Procesamiento semántico (los vocabularios, ontologías, etc. están en idiomas diferentes). Otro defecto no menor es que las herramientas tampoco tienen una interfaz en castellano no permitiendo el uso generalizado por usuarios no expertos.³ Esta inconveniencia es un obstáculo importante para la masificación del marcado con metadatos ya que discrimina a los usuarios por los idiomas que manejan.⁴

Se está implementando una herramienta de marcado de dominio público cuyo objetivo es masificar el uso de metadatos para recursos en castellano, al menos en Chile. Como filosofía se dará más importancia a la facilidad de uso que a la especialización, es decir, la generación de metadatos estará enfocada al usuario normal no experto, más que al usuario técnico o

³Con *usuario no experto* en este contexto nos referimos al usuario no experto en disciplinas referentes a la computación, metadatos y bibliotecología.

⁴Como ejemplo, señalemos que DC-dot [22], una de las mejores herramientas en inglés disponibles, al ser probada con páginas del Departamento de Ciencias de la Computación chilenos no dio resultados siquiera mínimamente aceptables (por ejemplo en el elemento *Materia* retorna artículos y adverbios como palabras claves).

especialista en alguna disciplina específica. Esto se logra dando flexibilidad al momento de escoger calificadores.⁵

Las herramientas de marcado existentes hoy en día, como DC-dot, funcionan extrayendo información de las etiquetas conocidas de HTML (tales como Title, META, etc) para usarla como contenido de los metadatos Dublin Core. La herramienta que actualmente está en desarrollo en el DCC, también extrae información de la etiquetas de HTML, por medio de un parseador y va generando los metadatos con esta información. Cabe destacar que el elemento más importante, y que representa la mayor dificultad en el desarrollo de la aplicación, es el elemento *Materia*. Su importancia radica en que se espera que sea el criterio más usado al momento de realizar una búsqueda y su complejidad radica en la necesidad de extraer información a partir de lenguaje natural en castellano para generar palabras claves y eliminar palabras irrelevantes (palabras vacías, artículos, adverbios, etc).

2. *Difusión*. Se creó una página sobre metadatos que puede orientar al principiante en el tema y en la búsqueda de información que se ha hecho inmanejable estos últimos años.

Desarrollo de un tutorial acerca de Dublin Core donde se entrega una descripción técnica completa del *Element Set* y de los calificadores (traducción al castellano) [26, 27].

3. *Traducción calificadores*. En el sitio oficial de Dublin Core existen traducciones del *Element Set* y de los calificadores (*Qualifiers*) a distintos idiomas. Está la traducción al castellano del *Element Set* pero no la de los calificadores. Se tradujeron los calificadores al idioma castellano; esta traducción está terminada y será enviada a DCMI para su publicación.
4. La perspectiva de utilizar un buscador chileno (por ejemplo *TodoCl*) para incorporar opciones de búsqueda con metadatos Dublin Core. Hay dificultades que resolver, desde indexación de metadatos Dublin Core hasta el tema de confianza (por ejemplo, metadatos generados con fines deshonestos).

5 Conclusiones

Esperamos que las motivaciones de este trabajo se hayan cumplido: presentar el tema de metadatos en la Web, motivar su uso, presentar Dublin Core en particular. Los autores están trabajando en la herramienta de marcado automático que debiera estar lista para fines de 2002, y entrelazar este proyecto con otros asociados a la Web Semántica en curso, como por ejemplo [33], y otros del Centro de Investigación de la Web [34].

Agradecimientos. Financiado por Núcleo Científico Milenio, Centro de Investigación de la Web, Grant P01-029-F, Mideplan, Chile.

⁵Un usuario experto tiende a describir sus recursos de manera estandarizada, es decir, utilizaría calificadores normalizados que pertenecen a vocabularios controlados muy específicos y rígidos. Exigirle al usuario no experto que describa sus recursos de esta manera le resta facilidad de uso a la herramienta, ya que lograría que el usuario se perdiera entre términos técnicos específicos. Un ejemplo de esto son los calificadores que recomienda Dublin Core para la codificación del elemento *Materia*. Estos calificadores pertenecen a sistemas de clasificación, *thesaurus* y vocabularios controlados muy rígidos y específicos, como por ejemplo LCSH (Library of Congress Subject Headings [32]). Por otro lado, sería deseable poder incluir la opción de utilizar calificadores normalizados para que el usuario experto los utilice para marcar de manera más estándar sus recursos, pero éstos no compliquen al usuario no experto.

Referencias

- [1] M. K. Bergman, *The Deep Web: Surfacing Hidden Value* The Journal of Electronic Publishing, University of Michigan Press, Vol. 7, 2001., <http://www.press.umich.edu/jep/07-01/bergman.html>
- [2] M. Lesk, *How Much Information Is There In the World?*, 2000, <http://www.lesk.com/mlesk/ksg97/ksg.html>
- [3] *Extensible Markup Language (XML) 1.0*, W3C Recommendation, 10 February 1998, <http://www.w3.org/TR/1998/REC-xml-19980210.html>
- [4] Tim Berners-Lee, <http://www.w3.org/People/Berners-Lee/>
- [5] Tim Berners-Lee, *Weaving the Web*, Harper San Francisco, 2000.
- [6] Resource Description Framework (RDF), <http://www.w3.org/RDF/>
- [7] Metadata - Dublin Core (MMI-DC) Workshop . <http://www.cenorm.be/iss/Workshop/MMI-DC/>
- [8] Memorandum of Understanding between the Dublin Core Metadata Initiative and the IEEE Learning Technology Standards Committee 2000-12-06. <http://dublincore.org/documents/2000/12/06/dcmi-ieee-mou/>
- [9] Open Archives Initiative, <http://www.openarchives.org/>
- [10] Semantic Web Activity Home Page, <http://www.w3.org/2001/sw/>
- [11] Proyecto Dublin Core, <http://www.dublincore.org/projects/>
- [12] On Line Computer Library Center, <http://www.olcl.org>
- [13] Global Information Locator Service (GILS), <http://www.gils.net/>
- [14] Australia New Zealand Land Information Council (ANZLIC), <http://www.anzlic.org.au/>
- [15] Australian Government Locator Service (AGLS), http://www.naa.gov.au/recordkeeping/gov_online/agls/summary.html
- [16] Education Network Australia (EdNA), <http://www.edna.edu.au/>
- [17] MACHine Readable Cataloging (MARC), <http://www.loc.gov/marc/>
- [18] ANSI/NISO Z39.50 -1995 (Z39.50), Information Retrieval : Application Service Definition & Protocol Specification, http://www.niso.org/standards/standard_detail.cfm?std_id=465
- [19] National Information Standards Organization (NISO), <http://www.niso.org/>
- [20] Dublin Core Metadata Initiative, <http://www.dublincore.org>
- [21] Online Computer Library Center (OCLC), <http://www.oclc.org/>
- [22] Dublin Core Metadata Editor (DC-dot), <http://www.ukoln.ac.uk/metadata/dcdot/>
- [23] Sistema de Servicios de Información y Bibliotecas (SISIB), <http://www.sisib.uchile.cl/>
- [24] Red Nacional de Información Bibliográfica, <http://www.renib.cl/>
- [25] Dirección de Bibliotecas Archivos y Museos, <http://www.dibam.cl/>
- [26] Metadatos Dublin Core, <http://www.dcc.uchile.cl/~mcaldero/cursodc/apunte.htm>
- [27] Página de Metadatos, <http://www.dcc.uchile.cl/~cgutierr/metadatos> (metadatos)
- [28] Multipurpose Internet Mail Extensions, MIME, <http://www.oac.uci.edu/indiv/ehood/MIME/MIME.html>
- [29] Lightweight Directory Access Protocol, <http://www.umich.edu/~dirsvcs/ldap/>, <http://www.openldap.org/>
- [30] WorlCat, <http://www.oclc.org/oclc/man/6928fsdb/worldcat.htm>
- [31] Thomas H. Davenport, *La "I" de la "TI"*, Management en Administración de la Información, Escuela de Ing. U. Católica. Depto Ingeniería Industrial y de Sistemas, El Diario, Ediciones Financieras, 2000, pp. 2-5.
- [32] Library of Congress Subject Headings, <http://www.loc.gov/cds/lcsh.html>
- [33] Depmark, <http://anakena.dcc.uchile.cl:8085/depmark/index.jsp>
- [34] Centro de Investigación de la Web, <http://www.ciw.cl>