



UNIVERSIDAD NACIONAL DE LA MATANZA

ESCUELA DE POSGRADO

MAESTRIA EN INFORMÁTICA

TESIS DE MAESTRÍA

**Gestión del Conocimiento:
estructuración de la información científica para la
Web Semántica**

Autor: Marcelo Claudio Perissé

:

Buenos Aires, *Septiembre 2009*

AGRADECIMIENTOS

Agradecimientos

Quisiera iniciar mis agradecimientos por la Universidad Nacional de La Matanza y el Departamento de Ingeniería e Investigaciones Tecnológicas, en el que día a día todos los miembros de la comunidad que la conforman construyen un magnífico espacio para el estudio y la investigación que conlleva al desarrollo de las personas.

A la Escuela de Posgrado y a la Maestría en Informática, tanto a su cuerpo docente, a su dirección y personal administrativo, pues en la construcción del proceso de Enseñanza-Aprendizaje, que llevan adelante cotidianamente y a la vocación puesta al servicio de la educación universitaria, han permitido que mi formación profesional en el campo de la informática sea una magnífica experiencia de vida.

A mis colegas y amigos, por su compañía y sabidurías de las cuales he tratado de nutrirme.

Y mi más profundo agradecimiento al Dr. Jorge Luis Narváez, quien ha apoyado y encausado las investigaciones y desarrollos llevados adelante para la consecución de esta Tesis.

ÍNDICE

Resumen	11
Acrónimos, siglas y abreviaturas	15
Capítulo I Introducción	17
Capítulo II La Web Semántica	25
Proyecto Web Semántica	27
Infraestructura de la Web Semántica	29
UNICODE	33
Universal Resource Identifier URI	33
Extensible Markup Language – XML	35
Namespaces	39
Aplicación de Espacios de Nombres a	
Elementos y Atributos	40
XML SCHEMA	41
Gramáticas XML orientada a datos	41
Gramáticas XML orientada a Metadatos	42
Resource Description Framework y Metadatos	45
Principios de metadatos	47
Metadatos Dublin Core	49
RDF schema	53
Ontología	58
Ontologías y OWL	58
Capas de Lógica, Reglas Y Trust	61
Digital Signature en la Web Semántica	63
La seguridad en las Redes	63
Principios de la Seguridad de la Información	63
Criptografía	65
Firma Digital	69
Aplicación de la firma digital en la Web Semántica	73
Criptografía en XML	73
Firma digital en XML	75

Capítulo III Diseño de las aplicaciones para la biblioteca digital	83
Selección del modelo	83
Object Oriented Hypermedia Design Method (OOHDM)	83
Análisis de requisitos respecto a la semántica	84
Documentación legible	85
Recuperación del esquema de un documento	85
Marcado semántico	85
Esquema semántico	86
Relaciones semánticas	87
Información semántica del elemento	88
Estructuras navegables	89
Capítulo IV Desarrollo del sistema de gestión del conocimiento	93
Modelo conceptual	93
El Hipertexto como recurso de divulgación	93
El hipertexto en la gestión de la información académica	99
Modelo Navegacional	105
Modelo de Interfaces Abstractas	109
Construcción de la Base de Datos	110
Arquitectura de la Base de Datos	110
Capítulo V Conclusiones y futuros trabajos	117
Capítulo VI Bibliografía	127

RESUMEN

Resumen

Las ideas expresadas en el presente proyecto reflejan en gran medida la necesidad de satisfacción de los requerimientos de información científica y técnica, y así generar nuevas perspectivas en un modelo que permita vincular a la Biblioteca Universitaria, y más concretamente a la hemeroteca, con los principales actores-usuarios de dichos requerimientos, como son los investigadores, los docentes, los alumnos y la comunidad de La Matanza.

La temática principal está basada en la estructuración de la información científica en Internet. El proyecto propone el desarrollo de aplicaciones informáticas, mediante la utilización de *metodologías simplificadas* basadas en: el *Rational Unified Process* (RUP) y el *Object Oriented Hypermedia Design Method* (OOHDM) como proceso para el desarrollo de aplicaciones hipermedia; aplicables a la construcción de un sistema de gestión del conocimiento para la Universidad Nacional de La Matanza, haciendo hincapié en la capacidad de generar actividades de almacenamiento y recuperación de información apropiada.

En la categorización de sitios web, se utiliza la taxonomía basada en metadatos, aplicando el modelo *Dublin Core Metadata Initiative* (DCMI) para saber qué elementos son representados, y la sintaxis *Extensible Markup Language* (XML) para saber cómo se representan dichos elementos.

Como lazo de unión hacia el aspecto del proceso de Enseñanza-Aprendizaje, las ontologías permiten un desarrollo semántico de gran profundidad que proporcionan una descripción lógica y formal de la información que almacenan, y por lo tanto puede ser interpretada por usuarios humanos y por programas informáticos. Para ello se utiliza el modelo *Resource Description Framework* (RDF) como herramienta en el

armado de este *esqueleto semántico* que sustenta a los tesauros en la representación y recuperación de información, con énfasis en la coherencia y riqueza relacional de la estructuración conceptual.

Uno de los aspectos destacables alcanzados es la especificación de las características técnicas, basadas en sintaxis *XML Encryption* y la aplicación de *XML Signature*, que determinan la viabilidad de los aplicativos de seguridad en la Web Semántica.

ACRÓNIMOS SIGLAS Y ABREVIATURAS

Acrónimos, siglas y abreviaturas

Acrónimos, siglas y abreviaturas	
RUP	Rational Unified Process
XP	eXtreme Programming
OOHDM	Object Oriented Hypermedia Design Method
DC	Dublin Core
XML	Extensible Markup Language
RDF	Resource Description Framework
UNESCO	United Nations Educational, Cientific and Cultural Organization
UML	Unified Modeling Language
W3C	Consortio World Wide Web
Scielo	Scientific Electronic Library Online
LILACS	Literatura Latino-Americana e do Caribe em Ciências da Saúde
SICyTAR	Sistema de Información de Ciencia y Tecnología Argentino
CvLAC	Currículum Vitae Latinoamericano y del Caribe
GrupLac	Directorio de Grupos y Proyectos de Investigación
SIU	Sistema de Información Universitaria
OAI-PMH	Open Archives Initiative – Protocol for Metadata Harvesting
OAI	Open Archives Initiative
HTML	HyperText Markup Language
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
NS	Namespaces
DNS	Domain Name Service
FTP	File Transfer Protocol
CGI	Common Gateway Interface
RFC	Request For Comments
IPv6	Internet Protocol Version 6

Acrónimos, siglas y abreviaturas (continuación)	
IPng	Next Generation Internet Protocol
ER	Diagramas Entidad Relación
DCMI	Dublin Core Metadata Initiative
WF	Warwick Framework
DTDs	Definiciones del Tipo de Documentos
OWL	Lenguaje de Ontologías Web
OOHDM	Object Oriented Hypermedia Design Method
SVG	Scalable Vector Graphics

CAPÍTULO I
INTRODUCCIÓN

Capítulo I - Introducción

La temática principal está basada en la estructuración de la información científica en Internet; en consecuencia el proyecto propone, basado en el estudio del modelo de gestión del conocimiento producido por la United Nations Educational, Scientific and Cultural Organization (UNESCO), la integración de los requerimientos mínimos para el desarrollo de aplicaciones informáticas intervinientes en el proceso Enseñanza-Aprendizaje.

Para ello se considera que es viable establecer una metodología simplificada, que permita el desarrollo de un sistema de gestión del conocimiento basado en la Web Semántica, haciendo hincapié en la capacidad de generar actividades de almacenamiento de información apropiada.

El desarrollo de dichas aplicaciones, se realizará mediante la utilización de metodologías simplificadas (PERISSE, 2001); las que se apoyan en la modelización del sistema a través de las técnicas de modelado *Unified Modeling Language* (UML) y en el *Object Oriented Hypermedia Design Method* (OOHDM) como proceso para el desarrollo de aplicaciones hipermedia.

Concretamente el trabajo presenta la construcción de un sistema de gestión del conocimiento universitario y la construcción de sus respectivas aplicaciones, basado en dos grandes pilares:

1. La biblioteca digital universitaria, haciendo hincapié en la capacidad de generar actividades de almacenamiento, recuperación y gestión de información apropiada y aplicada al proceso Enseñanza-Aprendizaje, de forma ordenada, distribuida y compartida.
2. El sistema curricular de los Docentes-Investigadores.

En la estructuración de la información sustentada en Internet se ha utilizado el modelo *Resource Description Framework* (RDF), pues éste se encuentra constituido como formato universal para datos en la web.

Dicho modelo relacional simple, permite vincular datos estructurados y semiestructurados que son exportados y compartidos a través de diferentes aplicaciones.

Además de la interoperabilidad de datos, el *Resource Description Framework* (RDF) provee una semántica para metadatos entendible por aplicaciones informáticas desarrolladas por la United Nations Educational, Scientific and Cultural Organization (UNESCO) y otros centros de gestión de la información científica como el Directorio y Recolector de Recursos Digitales del Ministerio de Cultura del Reino de España o la de los catálogos bibliotecarios y directorios *world-wide*.

Asimismo, el RDF, otorga una mejor precisión en la búsqueda de recursos que la obtenida por los motores de búsqueda que rastrean en el texto completo.

Como modelo para la descripción de los recursos de información, se asumió el *Dublin Core Metadata Initiative* (DCMI). Dicha iniciativa provee un estándar simple y universalmente adoptado por la comunidad científica, que permite encontrar, compartir y gestionar la información.

La sintaxis se desarrolló en *eXtensible Markup Language* (XML), debido a que es un lenguaje de etiquetado que juega un papel fundamental en el intercambio de una gran variedad de datos. Con él se describieron los distintos conjuntos de módulos ofrecidos como servicios a las demandas de los usuarios, permitiéndoles estructurar, almacenar e intercambiar información.

En esta misma sintaxis de *eXtensible Markup Language* (XML) y atendiendo a las normas de *XML Signature*, propuestas por el *Consortio World Wide Web* (W3C), se han construido los bloques de datos encriptados como forma de proteger los valores utilizados que aseguren su confidencialidad; y además, con el protocolo de la firma digital, se ha procurado garantizar la autoría e integridad del recurso.

Es importante destacar, que estos bloques serán utilizados por las computadoras y los agentes para verificar que la información adjunta, ha sido ofrecida por una fuente específica y confiable.

El proyecto se encuentra enmarcado en las políticas definidas por la *Rede ScientI*, liderada por Brasil y de la cual Argentina forma parte. Esta Red promueve un espacio público y cooperativo de interacción entre los actores de los sistemas, las comunidades de ciencia, de tecnología e innovación de sus países miembros.

Principalmente la *Rede ScientI*, se encuentra conformada por los siguientes sistemas:

- La Plataforma *Lattes*, como sistema que permite el acceso a la información científica a través de los currículum de los Docentes-Investigadores.
- El Modelo *Scientific Electronic Library Online* (SciELO), como infraestructura para la descripción de publicaciones científicas.
- La Metodología *LILACS* (Literatura Latino-Americana e do Caribe em Ciências da Saúde), para la generación de bases de datos bibliográficas.

Dentro de este mismo contexto, el proyecto también se encuentra alineado a las directivas propuestas por el Ministerio de Ciencia Tecnología e Innovación Productiva a través del *Sistema de Información de Ciencia y*

Tecnología Argentino (SICyTAR) quien administra, para Argentina, los sistemas: *Currículum Vitae Latinoamericano y del Caribe (CvLAC)*, y el *Directorio de Grupos y Proyectos de Investigación (GrupLac)*.

A su vez se asumieron las políticas promovidas por el *Sistema de Información Universitaria (SIU)*: SIU-Toba y SIU-Biblioteca, perteneciente a la Secretaría de Políticas Universitarias del Ministerio de Educación de la Argentina y del Ministerio de Ciencia Tecnología e Innovaciones Productivas; los cuales han formado parte, junto a la UNESCO y la Library of Congress, en el desarrollo de IsisMarc.

El IsisMarc es un software para bibliotecas, que permite el ingreso de información para el formato bibliográfico MARC21, sobre base de datos de tecnología *Isis*. En tanto MARC21, es un estándar internacional de catalogación que permite el intercambio de registros catalográficos entre bibliotecas.

Conjuntamente con las recomendaciones propuestas por el Ministerio de Cultura y Educación a través del Sistema de Información Universitario (SIU), se asumió la política de la *Open Archives Initiative – Protocol for Metadata Harvesting (OAI-PMH)* quien desarrolla y promueve estándares de interoperabilidad cuyo fin es facilitar la eficiencia en la disseminación de contenidos. Siendo que el trabajo de *Open Archives Initiative (OAI)* se ha ido expandiendo para promover el amplio acceso a los recursos digitales en el ámbito académico, el e-learning y la ciencia.

Es de destacar que la Investigación y Desarrollo aquí realizada incluye: el estudio de los vocabularios controlados – *Tesaurus* -, su modelo estructural – *Taxonomía* - y su representación explícita y formal como estructura conceptual de la base de conocimiento – *Ontología* -.

El presente trabajo, centra sus esfuerzos en la instrumentación de un modelo que permita la implantación de un lenguaje disciplinar que

favorezca las actividades de investigaciones pluridisciplinarias, interdisciplinarias y fundamentalmente las transdisciplinarias.

La perspectiva *Hermenéutica*, aquí abordada, del proceso Enseñanza-Aprendizaje nos permite asumir a la computadora como una herramienta ideada y desarrollada en el ámbito universitario y que puede: almacenar y procesar datos a fin de obtener información que faciliten la adquisición del conocimiento y por último automatizar reglas que nos lleven a una mejor comprensión de los hechos (ACKOFF, 1993).

Además, nos permite observar que la evaluación de la eficiencia se basa en una lógica y esta evaluación no presenta inconvenientes de ser programada en una computadora.

Esta programación puede ser automatizada mediante sistemas de inteligencia artificial basados en reglas del tipo tripleta: sujeto, predicado, objeto; como en el *lenguaje de programación lógico e interpretado* mayormente conocido bajo el nombre de Prolog, el cual se adecua al esquema del modelo *Resource Description Framework* (RDF).

CAPÍTULO II
LA WEB SEMÁNTICA

Capítulo II - La Web Semántica

Proyecto Web Semántica

La Web Semántica es un proyecto a corto, medio y largo plazo del organismo de regulación más importante del mundo en relación a Internet: el *World Wide Web Consortium* (W3C). El proyecto de la Web Semántica incluye transformaciones que ya están afectando a los ámbitos de la creación, la edición y la publicación de sitios web y que seguirán teniendo una importancia creciente en el futuro (Codina, 2002).

Sobre el World Wide Web Consortium (W3C), podemos decir que es el organismo que regula los aspectos esenciales de la Web, siendo por ende el organismo de normalización más importante de Internet.

Su director Tim Berners-Lee, el propio fundador de la Web, confía en que la Web Semántica sea el proyecto del W3C para transformar la Web en la Web de las próximas décadas (Berners-Lee, 2001); proporcionando un marco común que permita compartir y reutilizar los datos a través de aplicaciones, empresas y fronteras comunitarias.

Dicho marco común está basado en el modelo *Resource Description Framework* (RDF) e integra una variedad de aplicaciones utilizando *Extensible Markup Language* (XML) para la sintaxis y *Uniform Resource Identifier* (URI) para las denominaciones.

Según Berners-Lee (2001) entre los objetivos de la Web Semántica se encuentra la posibilidad de que se pueda sostener una interacción entre un usuario y un *agente de software*, mediante el cual el primero pueda ir expresando y perfilando sin ambigüedad puntos como los siguientes:

objetivos de la búsqueda, géneros documentales pertinentes, puntos de vista y granularidad esperada en la respuesta.

A partir de aquí, se espera que el agente de software sea capaz de elaborar una estrategia de búsqueda según su propia iniciativa; involucrando el uso de lenguajes documentales, metadatos y ontologías para responder con eficacia y rapidez al usuario. Se espera igualmente, que las computadoras puedan desarrollar tareas de gestión que requieran interpretar información y tomar decisiones adaptándolas al contexto.

Conformes a esta visión de la Web Semántica descrita por el W3C y Berners-Lee, se asumen dos aspectos relevantes:

1. Que los recursos a ser provistos en la Web Semántica deben estar organizados, estructurados y codificados según el modelo de la RDF. Esto es para que las computadoras y los actores envueltos en el proceso de Enseñanza-Aprendizaje, sean capaces de efectuar inferencias y razonar a partir de sus contenidos.
2. La colección de recursos almacenados en las Bases de Datos resultantes, deben estar integrados, historiados y disponibles. Con ello, se espera poder soportar un procesamiento sistemático y consistente de la información.

Infraestructura de la Web Semántica

La arquitectura de la Web Semántica está centrada en el modelo *Resource Description Framework* (RDF), que es un formato universal para datos en la web.

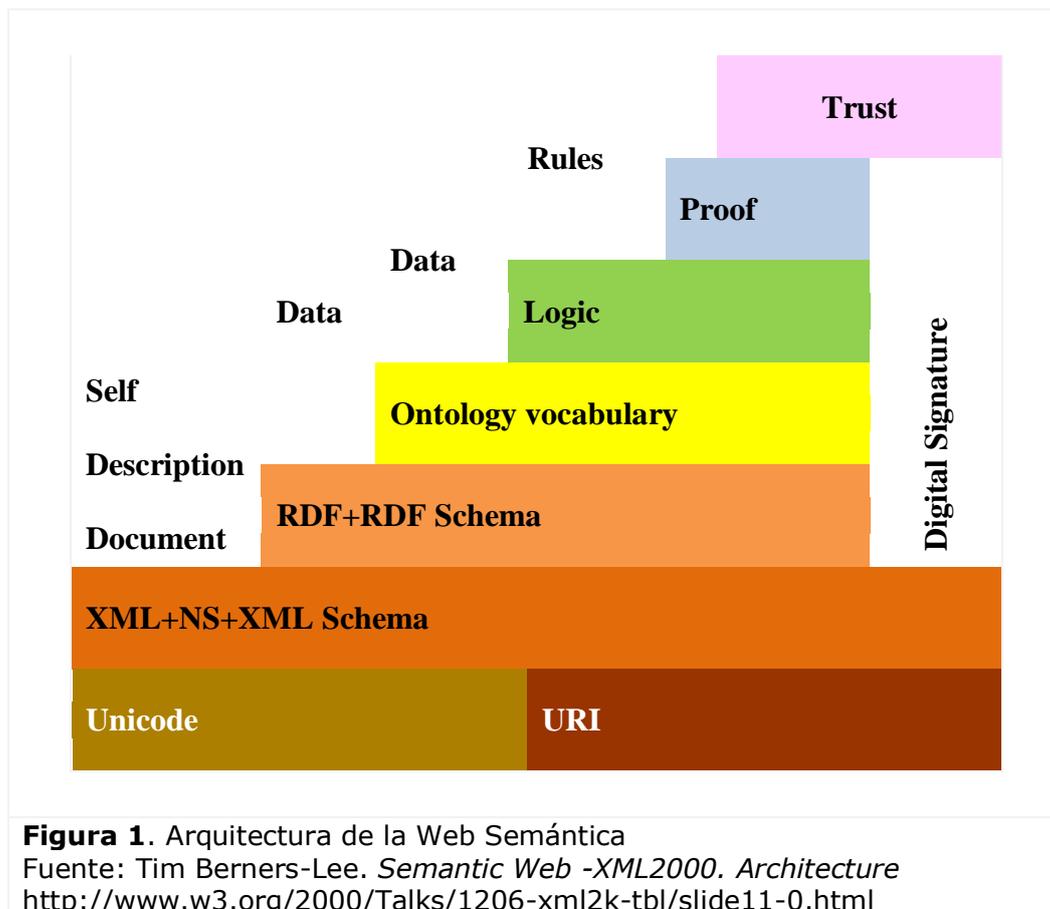
El uso de un modelo relacional simple, como el *Resource Description Framework* (RDF), permite mezclar datos estructurados y semiestructurados exportados y compartidos a través de diferentes aplicaciones.

Además de la interoperabilidad de datos, el *Resource Description Framework* (RDF) provee: una semántica para metadatos entendible por las computadoras; una mejor precisión en la exploración de recursos que la obtenida por los motores de búsqueda que rastrean el texto completo; y mejores aplicaciones. Todo ello en tanto se desarrollen los esquemas – *Schemas* - correspondientes.

En general, el RDF brinda las bases para herramientas genéricas, tanto sea para crear, gestionar o buscar datos en la web de manera entendible por las computadoras, promoviendo la transformación de la web en un repositorio de información procesable por las computadoras.

Entre la gran variedad de aplicaciones que posee el *Resource Description Framework* (RDF), este trabajo se centra en los catálogos bibliotecarios electrónicos o digitales y los directorios *Open Access*.

La Arquitectura de la Web Semántica, representada a través del diagrama desarrollado por Tim Bernes-lee, puede ser visualizada en la figura número 1.



Este modelo de gran capacidad expresiva, es el que será abordado y sobre el que guiarán las distintas actividades del proyecto. El mismo, se encuentra conformado por una metáfora de siete niveles; los que primeramente serán presentados, para luego ser descritos a mayor nivel de detalle:

- Unicode: se refiere a una codificación del texto, que permite utilizar los símbolos para expresar la información en la Web Semántica en cualquier idioma.
- URI: el *Uniform Resource Identifier* o Identificador Uniforme de Recursos, permite la localización de un recurso accesible vía Internet. Se trata de la descripción de la ubicación o *Uniform Resource Locator* (URL) – más la descripción de nombre *Uniform Resource Name* (URN).

- XML + NS + XML Schema: el *XML* ofrece un formato común para intercambio de documentos. El *Namespaces (NS)* sirve para cualificar elementos y atributos de nombres usados en *XML*, asociándolos con los espacios de nombre identificados por referencias *URI*. Por último, el *XML Schema* ofrece una plantilla para elaborar documentos estándar. De esta forma, aunque se utilicen diferentes fuentes, se crean documentos uniformes en un formato común y no propietario.
- RDF + RDF Schema: basada y apoyada en la capa anterior, esta capa define el lenguaje universal con el cual podemos expresar diferentes ideas en la Web Semántica. El *RDF* es un lenguaje simple mediante el cual definimos sentencias en el formato de una 3-upla o tripleta; compuesta por el *sujeto* o recurso al que nos referimos, el *predicado* que indica qué es lo que estamos definiendo y el *objeto* que indica el valor de lo que acabamos de definir. El modelo *Resource Description Framework (RDF)* es un modelo común que permite hacer afirmaciones sobre los recursos y que hace posible que estos recursos puedan ser nombrados por *URIs*. Por su parte el *RDF Schema* provee un vocabulario definido sobre *RDF* que permite el modelado de objetos con una semántica claramente definida. Esta capa no sólo ofrece descripción de los datos, sino también cierta información semántica. Tanto esta capa como la anterior, corresponden a las anotaciones de la información (metadatos).
- *Ontology vocabulary*: ofrece un criterio para catalogar y clasificar la información. El uso de ontologías permite describir objetos y sus relaciones con otros objetos; ya que una ontología es la especificación formal de una conceptualización de un dominio concreto del conocimiento. Esta capa permite extender la funcionalidad de la Web Semántica, agregando nuevas clases y propiedades para describir los recursos.

- Logic: además de ontologías, también se precisan reglas de inferencia. Una ontología puede expresar la siguiente regla: si un código de ciudad está asociado a un código de estado, y si una dirección es el código de ciudad, entonces esa dirección tiene el código de estado asociado. De esta forma, un programa podría deducir que una dirección de la Universidad Nacional de La Matanza, al estar en San Justo, debe estar situada en Argentina, y debería por lo tanto estar formateado según los estándares argentinos. La computadora no entiende lo que está procesando, pero puede manipular los términos de modo mucho más eficiente y beneficiar así a la inteligibilidad humana.
- Proof: significa *demostración matemática*. Se considera que una computadora alcanza la máxima fiabilidad en sus razonamientos cuando es capaz de realizar demostraciones o, lo que es lo mismo a efectos prácticos, cuando es capaz de justificar el motivo por el cual tomó (o aconsejó tomar) una decisión.
- Trust: los agentes deberían ser muy escépticos acerca de lo que leen en la Web Semántica hasta que hayan podido comprobar de forma exhaustiva las fuentes de información (Web Of Trust RDF Ontology -WOT- y FOAF).
- Digital Signature: son bloques encriptado de datos que serán utilizados por las computadoras y los agentes para verificar que la información adjunta ha sido ofrecida por una fuente específica y confiable (XML Signature).

Unicode

El conjunto de caracteres UTF8 (Transformación de la Representación Unicode) es una forma alternativa de guardar los datos Unicode y está implementado según el RFC 3629.

La idea del conjunto de caracteres UTF8 es que varios caracteres Unicode se codifican usando secuencias de bytes de longitudes diferentes:

- Usan un byte, letras básicas latinas, dígitos y signos de puntuación.
- Se guardan en una secuencia de dos bytes, la mayoría de letras europeas y del Medio Oriente: letras latinas extendidas (con tilde o diéresis), cirílico, griego, armenio, hebreo, arábigo, sirio, y otros.
- Usan secuencias de tres bytes, ideogramas koreanos, chinos, y japoneses.

Universal Resource Identifier URI

Los *Universal Resource Identifier* (URIs) fueron creados para ofrecer un mecanismo global y uniforme que permita identificar un recurso accesible determinado a través de una red. En la mayoría de los casos, las URIs que se utilizan suelen ser *Uniform Resource Locator* (URLs) (Senso Ruiz, 2003).

Los URIs son universales, es decir, tienen una sintaxis básica con la que trabajan y la que es independiente del tipo de recurso al que se refiere; bien sea para identificarlo o para dirigirse a él. Estos se dividen en tres partes:

1. Protocolo: sirve para describir el mecanismo que se debe utilizar para acceder correctamente al recurso (una página web o un servidor FTP).

2. Hostmane: es decir, el nombre principal de la dirección; en el caso del Web, el asignado mediante el *Domain Name Service* (DNS). Junto a él es posible encontrar también la ruta de acceso que es necesario seguir para alcanzar el recurso.
3. Consulta adicional: información adicional que permite realizar algún tipo de operación directamente con el recurso. Puede ser una consulta mediando un *Common Gateway Interface* (CGI) utilizando el signo "?" o una llamada a un anchor "#".

En este último punto es donde queda reflejada importancia de los URIs como elementos que permiten identificar recursos, namespaces y esquemas.

Básicamente, un URI es una cadena de caracteres que sigue las directrices especificadas en el RFC 2396 (Berners-Lee, 2001). Teniendo en cuenta que esta RFC está basada en normas que definen el comportamiento del servicio Domain Name Service (DNS), es lógico pensar que también utilizará el mismo conjunto de caracteres (el ASCII de 7-bit).

Internet2 conjuntamente con *Internet Protocol Version 6* (IPv6) o *Next Generation Internet Protocol* (IPng), se presentan como un factor relevante para dar más flexibilidad al direccionamiento.

En la actualidad, se encuentran URLs que no identifican siempre lo que dicen identificar; pues en ciertos casos apuntan a recursos que no existen (o que no son capaces de localizar) y además, tienen un período de vigencia muy limitado (Champin, 2000).

Si bien es cierto que muchos de estos problemas son reales, en la mayoría de las ocasiones son fácilmente subsanables si se realizara un estudio de los vínculos incluidos en un sitio Web antes de dar el alta.

A modo de conclusión, podemos afirmar que:

- Los URL describen una ubicación física.
- Los URN un nombre único
- Los URI lo describen todo; engloban a los URL y a los URN.

Extensible Markup Language - XML

El *Extensible Markup Language* (XML) es un estándar que, junto con su norma asociada *XML Schema*, permite definir tipos de documentos y los conjuntos de etiquetas necesarias para codificar tales tipos de documentos.

La idea es que una vez que los documentos están marcados o codificados con un conjunto de etiquetas XML, es posible procesarlos y explotarlos de forma automática con diversos propósitos. Incluso pueden exportarse a diversos *Sistemas de Gestión de Bases de Datos*, siempre y cuando la estructura de registros siga algún tipo de estándar (W3C XML 1.0, 2004).

Uno de estos propósitos puede ser codificar los documentos una sola vez, para poder mostrarlos a través de distintos dispositivos, tal el caso de un navegador de Internet como Mozilla o Explorer o la pantalla de una unidad telefónica móvil, siempre con el mismo conjunto de etiquetas y simplemente aplicando cada vez una hoja de estila distinta. Otros propósitos pueden ir desde la *Minería de Datos*, si es que se dispone de un conjunto de documentos suficiente, hasta la recuperación de información.

XML es, por lo tanto, un meta lenguaje de marcado que por un lado proporciona la posibilidad de codificar recursos de un modo directo con etiquetas ad-hoc y, por el otro lado, proporciona la posibilidad de definir esquemas y tipos de documentos que permiten crear instancias de documentos; cuya adecuación puede ser validada de forma automática con programas informáticos.

Con XML se pueden diseñar lenguajes de marcado estructurados y explícitos, en los cuales podemos disponer de etiquetas como: <title>, <creator>, <publisher>, <description>, <subject> o <contributor>, entre otros.

De este modo, si la Universidad precisara almacenar y procesar información sobre los recursos o currículum vitae de sus profesores, puede desarrollarse un *XML Schema* que le permita crear documentos *XML bien formados* como se muestra en la siguiente figura 2.

```
<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/metadata/dublin_core#"
  xmlns:dcq="http://purl.org/metadata/dublin_core_qualifiers#"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://unlam.edu.ar/.../xml/rdf_dc.xsd">
  ...
  -<rdf:description about="http://www.cyta.com.ar/ta0503/v5n3a2.htm">
  -<dc:title>
  -<rdf:Bag>
    <rdf:li language="spa">Infraestructura para la descripción de recursos en un
      aula digital: Marco tecnológico para la integración de los actores del proceso
      de enseñanza-aprendizaje</rdf:li>
    <rdf:li language="en">E-learning resource description framework: technological
      boundary for the teaching-learning process actors integration</rdf:li>
  </rdf:Bag>
  </dc:title>
  -<dc:creator>
  <nombreindex>Marcelo Claudio Périssé</nombreindex>
  <apellido>Périssé</apellido>
  <nombreprimero>Marcelo</nombreprimero>
  <nombresegundo>Claudio</nombresegundo>
  <url>
    http://www.sicytar.secyt.gov.ar/busqueda/prc_imp_cv_int?f_cod=00009082
    07
  </url>
  <filiacion>Departamento de Ingeniería e Investigaciones Tecnológicas,
    Universidad Nacional de La Matanza. </filiacion>
  <email>mperisse@unlam.edu.ar</email>
  ...

```

Figura 2. Vista desde un editor de XML la descripción de un recurso informático bajo el modelo RDF-DC

Los aspectos que hacen que el estándar XML sea importante para la gestión de la información, pueden agruparse en los siguientes tópicos:

1. El XML es un sistema abierto, no propietario e independiente de la plataforma (Hardware y Software).
2. El XML al ser independiente de la plataforma y el código fuente, la información contenida en él, puede ser interpretada directamente por las personas; por lo que su persistencia y conservación están razonablemente garantizadas para el futuro.
3. El XML al separar el marcado del contenido de su presentación, la información es almacenada en un único formato y puede ser reutilizada en tantos formatos como sea necesario: HTML, PDF, XLST, DOC, entre tantos otros.

Entre las aplicaciones del XML a la bibliotecología, se pueden mencionar los siguientes proyectos liderados por:

1. Library of Congress: para migrar el formato *MARC* al entorno XML y de hacer lo mismo con el formato de codificación de archivos *Encoded Archival Description (EAD)*.
2. BioMed Central: editorial a través de internet con publicaciones que se ajustan a los principios de *Open Access*.
3. Budapest Open Access Initiative: facilita, ayuda y promueve proyectos encaminados a liberar el acceso electrónico a artículos de investigación, de todas las áreas, bien a través de su archivo en un repositorio o mediante su publicación en revistas de libre acceso, es una actividad de la Fundación Soros.
4. The Public Library of Science: es una organización para el apoyo y la creación de revistas de libre acceso.

5. The Scholarly Publishing and Academic Resources Coalition: coalición de universidades y bibliotecas para difundir y contribuir, a través de la iniciativa de libre acceso, con las revistas.
6. Scientific Electronic Library Online – Scielo: es un proyecto de países latinoamericanos que facilita, a través de repositorios, la consulta y visibilidad de sus publicaciones en Internet.
7. Directory of Open Access Journals – DOAJ: proyecto internacional liderado por la Universidad de Lund en el que se han recogido más de 700 publicaciones de libre acceso de diferentes procedencias.

La gran mayoría de los proyectos actuales de bibliotecas digitales que se están desarrollando en diversos ámbitos temáticos y geográficos, se apoyan en la codificación de los metadatos de cada documento. Esta codificación se realiza con formatos como el *Dublin Core Metadata Initiative* (DCMI), expresados en sintaxis XML y bajo el modelo RDF.

Por último, podemos señalar el movimiento *Open Archives Initiative – Protocol for Metadata Harvesting* (OAI-PMH) como otra de las formas más interesantes en las cuales XML tiene incidencia en el mundo de la documentación. En este caso en el mundo de la información y documentación científica a través de los repositorios de: artículos, informes y resultados de investigación; creados por diversas instituciones de investigación y muy particularmente universidades. (Barrueco)

La *Open Archives Initiative* (OAI) desarrolla y promueve standards de interoperabilidad, cuyo fin es facilitar la eficiente disseminación de contenido. La OAI tiene sus raíces en el *Open Access* y en las iniciativas de repositorios institucionales.

A través de los años, el trabajo de OAI se ha ido expandiendo para promover el amplio acceso a los recursos digitales en el ámbito académico, el e-learning y la ciencia.

Namespaces

La noción de *namespaces* o espacios de nombres, es una parte fundamental de la infraestructura Web y particularmente en XML. Aunque el concepto es anterior a la Web, son una parte crítica de la infraestructura necesaria para desplegar sistemas de metadatos modulares en el Web (W3C, 2004).

Según la W3C, un espacio de nombres XML (Namespaces in XML) es un conjunto de nombres identificados por una referencia URI, que se utilizan en documentos XML como tipos de elemento y nombres de atributo.

Recordemos que un namespace es un vocabulario definido dentro de un *Universal Resource Identifier* (URI); considerándose que aquellas URIs, que identifican espacios de nombres son idénticas cuando son exactamente las mismas, caracter por caracter.

En lo que respecta al principio de extensibilidad, ya fue comentado en el tópico anterior que, el XML fue creado para permitir la interoperabilidad. Como cada desarrollo puede crear sus propios vocabularios en XML, esto produciría confusiones si diferentes desarrolladores optasen por los mismos nombres de elementos al representar diferentes entidades.

Por lo tanto, los namespaces fueron introducidos en XML para resolver conflictos de nombres entre elementos en un documento XML cuando los elementos se derivan de diferentes fuentes, permitiendo el uso de múltiples vocabularios en un mismo documento.

Cualquier conjunto de elementos de metadatos es un namespace limitado por las reglas y convenciones determinadas por su agencia de mantenimiento.

Las declaraciones de namespace permiten al diseñador de esquemas de metadatos definir el contexto para un término particular, asegurando que

el término tiene una única definición dentro de los límites del namespace declarado.

Usando esta infraestructura, los diseñadores del sistema de metadatos pueden seleccionar los elementos ya establecidos y más convenientes del conjunto de metadatos existentes, evitando reinventar los conjuntos de los metadatos establecidos para cada nuevo dominio.

Aplicación de Espacios de Nombres a Elementos y Atributos

Se considera que la declaración del espacio de nombres se aplica al elemento en que está especificada y a todos los elementos pertenecientes al contenido de ese elemento, siendo que se pueden declarar varios prefijos de espacios de nombres como atributos de un mismo elemento, como se muestra en este ejemplo:

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc = "http://purl.org/metadata/dublin_core#"
  xmlns:xsi = "http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation = "http://www.cyta.com.ar/xml/rdf_dc.xsd">
  <rdf:description about = "http://www.cyta.com.ar/ta0101/tapa.htm">
    <dc:title>
      <rdf:Bag>
        <rdf:li language = "es">
          La competitividad internacional lleva a los ejecutivos a las
          Universidades</rdf:li>
        <rdf:li language = "en">
          The international competitiveness takes to the executives to the
          Universities</rdf:li>
      </rdf:Bag>
    </dc:title>
```

XML Schema

Según la W3C (W3C Esquema XML, 2001), las gramáticas XML pueden clasificarse en dos ejes diferentes:

- **Orientada a datos:** juegos de etiquetas para Interfaz de Usuario (UI); interpretación especializada como por ejemplo *Scalable Vector Graphics* (SVG) o cualquier formato de almacenamiento de datos genéricos.
- **Orientada a Metadatos:** cuando el contenido marcado es un metadato. Ejemplos: para expresar datos que se procesan con *Extensible Style Language* (XSL), con Resource Description Framework (RDF) o lenguajes de Esquema.

Gramáticas XML orientada a datos

Una definición informal podría ser cualquier cosa para la que tenga sentido la pregunta ¿Hay un equivalente textual para todos los componentes de datos de medios enriquecidos? La schemata centrada en datos incluye los aspectos tanto de interacción como de comportamiento de una aplicación XML.

Esto nos lleva a que uno de los aspectos más importantes de la Web Semántica es que los documentos deben ser igualmente *accesibles* (entendidos), por la totalidad de su público objeto, independientemente del dispositivo que se utilice para acceder a él.

Entonces, consideramos que un esquema XML es *accesible* si posibilita, y de hecho promueve activamente, la creación de documentos accesibles. Por lo tanto, en la creación de los documentos, debe procurarse ser independiente del tipo de dispositivo y exportar sus semánticas tanto como sea posible.

Ambos aspectos son importantes, pues sin el conocimiento del significado de los elementos y atributos XML, será poco probable que las aplicaciones alternativas del usuario puedan hacer algo inteligente contando solo con partes del documento.

Consecuentemente, si se cuenta con aserciones de semánticas que puedan ser leídas por la computadora, los recursos podrán ser *accesibles*; permitiendo a otros usuarios mapear desde su lenguaje a otros existentes o viceversa.

Gramáticas XML orientada a Metadatos

La razón de ser de un esquema es definir una clase de documentos XML, y por tanto el término *documento instancia* se usa frecuentemente para describir un documento que se conforma con un esquema en particular.

Es importante destacar que, ni las instancias ni los esquemas necesitan existir como documentos per se, ya que pueden existir como flujos de bytes enviados entre aplicaciones o como campos en un registro de una base de datos.

En la figura 3 Se muestra un fichero en sintaxis XML y en el que se describen, a través del modelo Dublin Core (DC), los artículos publicados en una publicación seriada de Información Científico-Técnica (ITC).

En este ejemplo cada recurso es descripto por 15 elementos: language, title, creator, subject, description, publisher, contributor, type, format, identifier, identifier, source, language, coverage y date.

```

<?xml version="1.0" encoding="utf-8" ?>

  <dc:DC xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/metadata/dublin_core#"
  xmlns:dcq="http://purl.org/metadata/dublin_core_qualifiers#">

<DC>

<dc:title language="ES">La competitividad internacional lleva a los ejecutivos a las
Universidades</dc:title>
<dc:title language="en">The international competitiveness takes to the executives to the
Universities</dc:title>
<dc:creator>Romanni, Gianni</dc:creator>
<dc:description language="ES">Presenta cual es la importancia que tienen en la actualidad,
para los profesionales que actúan en el área de la administración, los programas de
Magister.</dc:description>
<dc:subject>Emprendedurismo</dc:subject>
<dc:subject>Posgrados</dc:subject>
<dc:subject>Educación superior</dc:subject>
<dc:subject>Capacitación</dc:subject>
<dc:subject>Posgrado</dc:subject>
<dc:publisher>Técnica Administrativa issn 1666-1680</dc:publisher>
<dc:contributor />
<dc:type>Technical note</dc:type>
<dc:format>htm</dc:format>
<dc:identifier>010100</dc:identifier>
<dc:source>www.cyta.com.ar/ta0101/tapa.htm</dc:source>
<dc:language>ES</dc:language>
<dc:coverage>ADMG</dc:coverage>
<dc:date>2001-09-15</dc:date>

</DC>

```

Figura 3. Fichero XML

El esquema del recurso consiste de un elemento schema y una variedad de subelementos, desde los más notables son: element (elemento), complexType (tipo complejo) y simpleType (tipo simple), que determinan la aparición de elementos y su contenido en el documento instancia dc.xml, tal se puede apreciar en la figura 4.

```

<?xml version="1.0" encoding="utf-8" ?>
- <xsd:schema xmlns:xsd="http://www.w3c.org/2001/XMLSchema"
  targetNamespace="http://www.cyta.com.ar/xml/rdf_dc.xsd">
- <xsd:annotation>
  <xsd:documentation xml:lang="es">Esquema de artículos de la revista Técnica
    Administrativa ISSN 1666-1680</xsd:documentation>
  </xsd:annotation>
- <xsd:complexType name="rdf:description">
- <xsd:sequence>
  <xsd:element name="dc:title" type="xsd:string" />
  <xsd:element name="dc:creator" type="xsd:string" />
  <xsd:element name="dc:subject" type="xsd:string" />
  <xsd:element name="dc:description" type="xsd:string" />
  <xsd:element name="publisher" type="xsd:string" default="Ciencia y Técnica
    Administrativa" />
  <xsd:element name="dc:contributor" type="xsd:string" />
  <xsd:element name="dc:type" type="listaDeTipos" />
  <xsd:element name="dc:format" type="xsd:string" />
  <xsd:element name="dc:identifier" type="xsd:decimal" use="required" />
  <xsd:element name="dc:source" type="xsd:string" />
  <xsd:element name="dc:language" type="xsd:string" />
  <xsd:element name="dc:relation" type="xsd:string" />
  <xsd:element name="dc:coverage" type="xsd:string" />
  <xsd:element name="dc:date" type="xsd:date" />
  </xsd:sequence>
</xsd:complexType>
- <xsd:simpleType name="listaDeTipos">
- <xsd:restriction base="xsd:string">
  <xsd:enumeration value="Research article" />
  <xsd:enumeration value="Review article" />
  <xsd:enumeration value="Technical note" />
  </xsd:restriction>
</xsd:simpleType>
</xsd:schema>

```

Figura 4. Vista del esquema correspondiente al modelo RDF-DC

Por lo tanto, un esquema XML debe contener la definición de la estructura válida para un tipo de documento XML, es decir:

- Los elementos que pueden aparecer en el documento.
- Los atributos que pueden utilizarse junto a cada elemento.
- Cómo se pueden anidar los elementos padres e hijos.
- El orden en el que deben aparecer los elementos hijos de un mismo padre.

- El número permitido de elementos hijos.
- Si un elemento puede ser vacío o no.
- Tipos de datos para elementos y atributos.
- Valores por defecto y fijos para elementos y atributos

Las ventajas de XML Schemas son:

- Precisión en la definición de tipos de datos mediante formatos y facetas.
- Los esquemas se definen como documentos XML con extensión XSD.
- Los documentos XML que se basen en esquemas, contendrán una referencia al respectivo archivo XSD.

Resource Description Framework y Metadatos

El objetivo general de *Resource Description Framework* (RDF) es definir un mecanismo para describir recursos que no cree ninguna asunción sobre un dominio de aplicación particular, ni defina (a priori) la semántica de algún dominio de aplicación. La definición del mecanismo debe ser neutral con respecto al dominio; sin embargo, el mecanismo debe ser adecuado para describir la información sobre cualquier dominio.

El fundamento del RDF, es contar con un modelo para representar propiedades designadas y valores de propiedades. El modelo RDF se basa en principios perfectamente establecidos, por varias comunidades, de representación de datos.

Las propiedades representadas en RDF pueden hacernos recordar a atributos de recursos, y en este sentido corresponden con los tradicionales pares de atributo-valor. Éstas propiedades en *Resource Description Framework* (RDF) representan también la relación entre recursos y por lo tanto, un modelo RDF puede asemejarse a un *Diagrama Entidad Relación*

(ER). En la terminología del diseño orientado a objetos, los recursos se corresponden con objetos y las propiedades se corresponden con objetos específicos y variables de una categoría.

La sintaxis XML aplicada al modelo RDF puede hacer posible especificar la semántica para las bases de datos en XML de una forma normalizada e interoperable. También es importante entender que esta sintaxis XML es sólo una sintaxis posible para RDF y que pueden surgir formas alternativas para representar el mismo modelo de datos RDF.

El modelo de datos *Resource Description Framework* proporciona un marco abstracto y conceptual para definir y utilizar metadatos.

La representación del modelo de datos, se usa para evaluar la equivalencia en significado; por lo tanto, dos expresiones RDF son equivalentes si y sólo si, sus representaciones del modelo de datos son las mismas. Esta definición de equivalencia permite algunas variaciones sintácticas en expresiones sin alterar el significado.

El modelo de datos básico consiste en tres tipos de objetos:

- Recursos: todas las cosas descritas por expresiones RDF se denominan recursos.
- Propiedades: una propiedad es un aspecto específico, característica, atributo o relación utilizada para describir un recurso.
- Sentencias: declaraciones o enunciados.

Estas tres partes individuales de una sentencia se denominan, respectivamente: sujeto, predicado y objeto.

El *Resource Description Framework* es la propuesta del Consorcio Web W3C, en el proyecto de la *Web Semántica*, para definir metadatos en el Web y la base para el procesamiento de metadatos, proporcionando

interoperabilidad semántica entre las aplicaciones que intercambian información entendible por máquina (Díaz Ortuño, 2003).

El RDF es simplemente un modelo de datos que permite crear metadatos legibles y entendibles por máquina. La interoperabilidad semántica de sistemas de metadatos implica significados compartidos y gramáticas compartidas.

Principios de metadatos

Los metadatos abarcan tanto los datos asociados con un sistema de información, como los asociados a un objeto de información con la finalidad de descripción, administración, señalar los requisitos legales, la funcionalidad técnica, el uso, la utilización y la preservación.

Los principios comunes a todos los dominios de metadatos y que podrían servir de base en el diseño de cualquier esquema de metadatos o aplicación (DUVAL, 2002), como ya habíamos adelantado en el tópico de namespace, son:

1. Modularidad: debe ser alcanzable una flexibilidad en la arquitectura de metadatos Web, que permita a los diseñadores mezclar una variedad de módulos semánticos en un esquema compuesto, dentro de un fundamento sintáctico común. En una arquitectura modular de metadatos, los elementos de datos de esquemas diferentes así como los vocabularios y otros elementos constructivos pueden combinarse de una manera interoperable sintáctica y semánticamente. Pueden ensamblarse conjuntos modulares basados en esquemas de metadatos preestablecidos que incluyan la funcionalidad de cada componente y que reúnan los requisitos específicos de una determinada aplicación.

2. Extensibilidad: las arquitecturas de metadatos deben acomodar la noción de un esquema base con elementos adicionales que ajusten una aplicación, dada a las necesidades locales o las necesidades específicas del dominio sin comprometer excesivamente la interoperabilidad proporcionada por el esquema base. Los sistemas de los metadatos deben permitir extensiones para que puedan acomodarse a las necesidades particulares de una determinada aplicación. Algunos elementos probablemente serán comunes a la mayoría de los esquemas de metadatos (por ejemplo, el concepto de creador o identificador de un recurso de información) mientras que otros serán específicos a aplicaciones particulares o dominios.
3. Refinamiento: los dominios de aplicación se diferirán de acuerdo con el nivel de detalle necesario o deseable. El diseño de estándares de metadatos debería permitir, a los diseñadores de esquemas, el nivel de detalle apropiado a la aplicación. La inclusión de calificadores refinan el significado de un elemento como autor, compositor o escultor que son ejemplos de tipos particulares de un término más general como es el de creador. El uso de tesauros, es una aproximación sumamente importante que mejora la precisión de las descripciones y permite aprovechar la inversión intelectual realizada por muchos dominios, en la mejora del acceso por materias a los recursos.
4. Plurilingüismo: es esencial adoptar arquitecturas de metadatos que respeten la diversidad lingüística y cultural. Si los recursos están disponibles a los usuarios en sus idiomas nativos, en los conjuntos de caracteres apropiados, y con los metadatos apropiados a la gestión de los recursos, la Web alcanzará su potencial como sistema de información global.

Metadatos Dublin Core

Según la *Iniciativa de Metadatos Dublin Core* (DCMI), organismo responsable del mantenimiento constante del conjunto de elementos de metadatos Dublin Core, ésta iniciativa incluye todos los términos que son matizaciones, esquemas de codificación y términos de vocabulario controlado destinados a facilitar la localización y la recuperación de recursos.

La Iniciativa Dublin Core inició sus actividades en 1995 en un encuentro en Dublin, Ohio (USA) en el que participaron el National Center for Supercomputing Applications (NCSA) y el On Line Library Computer Center (OCLC), junto con representantes de la Internet Engineering Task Force (IETF) y en el que bibliotecarios, proveedores de contenido y expertos en lenguajes de marcado pretendieron desarrollar estándares para describir los recursos de información y facilitar su recuperación. Siendo que en febrero del año 2003, la *Dublin Core metadata element set* se convirtió en norma ISO 15836/2003, de la *International Organisation for Standardisation*.

Es así como nació un pequeño conjunto de descriptores, que en principio fueron pensados para que fuera el propio autor el que los incluyera en el documento o recurso. Pero rápidamente adquirieron alcance global, pues también se interesaron en ellos numerosos y variados proveedores de información pertenecientes a diferentes sectores como el de las artes, las ciencias, la educación, los negocios y la administración.

Hoy, los metadatos Dublin Core se han convertido en uno de los estándares más extendidos para la recuperación de información en la World Wide Web y es en la actualidad un vocabulario muy utilizado no solamente en el ámbito bibliotecario y documental, sino también en otros muchos sectores.

Este conjunto de metadatos, puede ser aplicado en el encabezado de los archivos con formato HyperText Markup Language (HTML), como se puede observar en la figura número 5.

```
<!doctype HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<html>
<head>
  <!--Dublin Core -->
  <meta name="dc:title" content="Cultura organizacional y nivel de innovación ">
  <meta name="dc:title" content="Organizational culture and degree of innovations">
  <meta name="dc:creator" content="Camio, María Isabel">
  <meta name="dc:creator" content="Rébori, Alfredo José">
  <meta name="dc:creator" content="Romero, María del Carmen">
  <meta name="dc:subject" content="cultura">
  <meta name="dc:subject" content="innovación">
  <meta name="dc:subject" content="empresas">
  <meta name="dc:subject" content="software">
  <meta name="dc:subject" content="culture">
  <meta name="dc:subject" content="innovation">
  <meta name="dc:subject" content="companies">
  <meta name="dc:description" content="Se trata de ...">
  <meta name="dc:description" content="This is an exploratory...">
  <meta name="dc:publisher" content="Técnica Administrativa issn 1666-1680">
  <meta name="dc:format" content="htm">
  <meta name="dc:identifier" content="080102">
  <meta name="dc:source" content="www.cyta.com.ar/ta0801/v8n1a2.htm">
  <meta name="dc:language" content="es">
  <meta name="dc:relation" content="adm">
</--/Dublin Core -->
```

Figura 5. Conjunto de metadatos DC aplicados en el encabezado de un recurso en HTML

También, puede ser aplicado en lenguajes estructurados como XML (ver figura 6) y conjuntamente con otros lenguajes de descripción como RDF.

```

<?xml version="1.0" encoding="utf-8" ?>
<dc:DC xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/metadata/dublin_core#"
  xmlns:dcq="http://purl.org/metadata/dublin_core_qualifiers#"
  xsi:schemaLocation="http://www.cyta.com.ar/xml/rdf_dc.xsd">
...
- <dc>
  <dc:title language="ES">Cultura organizacional y nivel de innovación: estudio
    comparado de casos en empresas de software</dc:title>
  <dc:title language="en">Organizational culture and degree of innovation:
    comparative case study in software companies</dc:title>
  <dc:creator>Camio, María Isabel</dc:creator>
  <dc:creator>Rébori, Alfredo José</dc:creator>
  <dc:creator>Romero, María del Carmen</dc:creator>
  <dc:description language="es">Se trata de una indagación...</dc:description>
  <dc:description language="en"> This is an exploratory...</dc:description>
  <dc:subject>cultura</dc:subject>
  <dc:subject>innovación</dc:subject>
  <dc:subject>empresas</dc:subject>
  <dc:subject>software</dc:subject>
  <dc:subject>culture</dc:subject>
  <dc:subject>innovation</dc:subject>
  <dc:subject>companies</dc:subject>
  <dc:publisher>Técnica Administrativa issn 1666-1680</dc:publisher>
  <dc:type>Research article</dc:type>
  <dc:format>htm</dc:format>
  <dc:identifier>080102</dc:identifier>
  <dc:source>www.cyta.com.ar/ta0801/v8n1a2.htm</dc:source>
  <dc:language>es</dc:language>
  <dc:relation>adm</dc:relation>
  <dc:coverage>admng</dc:coverage>
  <dc:date>2009-01-15</dc:date>
  </dc>

```

Figura 6. metadatos DC aplicados en un recurso en XML

El conjunto de elementos *Dublin Core* pueden clasificarse en 3 grupos, los que indican la clase o el ámbito de la información que contienen, a saber:

1. Elementos relacionados principalmente con el contenido del recurso:
 - a. Title (título)
 - b. Subject (tema)
 - c. Description (descripción)

- d. Source (fuente)
 - e. Lenguaje (lenguaje)
 - f. Relation (relación)
 - g. Coverage (cobertura)
2. Elementos relacionados principalmente con el recurso cuando es visto como una propiedad intelectual:
- a. Creator (autor)
 - b. Publisher (editor)
 - c. Contributor (colaboradores)
 - d. Rights (derechos)
3. Elementos relacionados principalmente con la instanciación del recurso:
- a. Date (fecha)
 - b. Type (tipo de recurso)
 - c. Format (formato)
 - d. Identifier (identificador)

La *Iniciativa de Metadatos Dublin Core* ya indicaba, en sus primeros momentos de desarrollo, que para promover una interoperabilidad global, podría asociarse una descripción del valor de algunos elementos a vocabularios controlados y asumía que serían desarrollados otros vocabularios controlados para asegurar esta interoperabilidad en dominios específicos.

Esto es lo que se ha hecho con los elementos DC Format y DC Type mediante la creación de una norma ISO 11179 sobre registros de metadatos y la recomendación DCMI Type Vocabulary, respectivamente.

Dublin Core permite emplear calificadores opcionales para cada elemento que posibilitan indicar la normativa empleada, en caso de haber usado normas de descripción bibliográfica usuales. Estos calificadores permiten aumentar la especificidad y precisión de los metadatos, aunque pueden también introducir cierta complejidad que disminuiría la compatibilidad con otras aplicaciones que usen Dublin Core.

Dentro de las ventajas de usar metadatos Dublin Core, podemos destacar que éstos representan un estándar muy extendido a nivel internacional, que son aplicables en distintas disciplinas y además su tienen una gran facilidad de uso. De hecho, se ha convertido en un modelo de descripción de datos semántico a través de RDF.

Además, Dublin Core se usa junto con otros perfiles de aplicación como el protocolo Z39.50 para el intercambio entre sistemas bibliotecarios, quien acordó incluir los 15 elementos de los metadatos Dublin Core como formato aceptado y, de esta forma, se pueden recuperar los elementos DC especificándolos en las búsquedas.

RDF Schema

Resource Description Framework (RDF) y el *lenguaje de RDF Schema* se han basado en la investigación sobre metadatos en el ámbito de las Bibliotecas Digitales.

En particular, RDF adopta una aproximación modular a los metadatos que puede considerarse una implementación del Warwick Framework (WF) y representa una evolución de dicho modelo, permitiendo que cada vocabulario de metadatos se represente en una sintaxis diferente.

En RDF, todos los vocabularios se expresan dentro de un modelo individual bien definido, dando lugar a una mezcla específica más refinada de los vocabularios que se pueden procesar por máquina. Esto permite

crear metadatos, en los que las sentencias pueden trazarse en múltiples vocabularios, que se gestionan de forma descentralizada por comunidades independientes de expertos.

La especificación del Esquema RDF, proporciona un sistema entendible por máquina para definir esquemas para vocabularios específicos; como en el caso del Dublin Core. Esto posibilita a los diseñadores, especificar clases de tipos de recursos y propiedades que permiten dar a conocer las descripciones de esas clases, las relaciones entre aquellas propiedades y clases y las restricciones en las combinaciones permitidas de: clases, propiedades y valores.

Los Esquemas RDF, podrían contrastarse con las Definiciones del Tipo de Documentos (DTDs) de XML y los Esquemas XML; En donde a diferencia de una DTD o Esquema XML, que dan restricciones específicas en la estructura de un documento XML, un Esquema RDF proporciona información sobre la interpretación de una sentencia dada en un modelo de datos RDF.

Además, mientras un Esquema XML puede utilizarse para validar la sintaxis de una expresión RDF/XML; un esquema sintáctico sólo, no es suficiente para los objetivos de RDF. También, los Esquemas RDF pueden especificar restricciones que deben seguirse por estos modelos de datos. Se espera que resultados de trabajos futuros en torno a estos dos esquemas (XML Y RDF), podrán facilitar la sencilla combinación de reglas sintácticas y semánticas para ambos.

Por lo tanto el RDF Schema es una extensión de RDF, que proporciona primitivas adicionales que enriquecen el modelo básico; proporcionando un vocabulario para RDF, que se asume tiene una cierta semántica. Permitiendo a los diseñadores poder especificar, a través del RDF Schema, una jerarquía explícita de clases de recursos y propiedades que describen

estas clases, junto con las restricciones sobre las combinaciones permitidas de clases, propiedades y valores (Díaz Ortuño, 2003).

A continuación se describen algunas características de *RDF Schema* (RDFS), que fueron aplicadas en el proyecto:

Clases:

- `rdfs:Resource`: cualquier cosa descrita por una sentencia RDF se considera una instancia de la clase
- `rdf:Property`: es la clase de todas las propiedades utilizadas en la caracterización de las instancias de `rdfs:Resource`.
- `rdfs:Class`: se usa para definir conceptos y cada concepto debe ser una instancia de `rdfs:Class`.

Propiedades:

- `rdf:type`: modela interrelaciones del tipo "instancia de" entre recursos y clases; siendo que un recurso puede ser una instancia de más de una clase.
- `Rdfs:subClassOf`: modela la jerarquía de clases, donde una clase puede ser subclase de otras subclases.
- `rdfs:subPropertyOf` : si una propiedad P2 es una subpropiedad de otra propiedad P1, y si un recurso R tiene una propiedad P2 con valor V, esto implica que el recurso R también tiene la propiedad P1 con valor V.

Restricciones y clases:

- `rdfs:ConstraintResource`: define la clase de todas las restricciones.
- `rdfs:ConstraintProperty`: es un subconjunto de `rdfs:ConstraintResource`.
- `rdfs:range` y `rdfs:domain`: son dos instancias de `rdfs:ConstraintResource`, que se usan para restringir el rango y el dominio de las propiedades. No siendo posible expresar más de una restricción de rango sobre una propiedad. Destacando que en dominios sí se permite, y se interpreta como una unión de dominios.

Antes de pasar al próximo tópico, podríamos resumir lo visto desde el inicio del capítulo diciendo que: cuando escribimos una frase en lenguaje natural, utilizamos palabras que encierran la intención de transmitir un significado inequívoco.

Ese significado, atribuido por el emisor, es fundamental para que los receptores puedan entender el enunciado desde los mismos parámetros del emisor y, en el caso de aplicaciones de RDF, es crucial para establecer que el procesamiento correcto se dé como se esperaba.

Con ello queremos decir que es esencial que tanto el emisor como el receptor, de una sentencia o declaración, entiendan el mismo significado para los términos utilizados. Significados tales como: *creator*, *contributor*, *publisher*, *subject*; si ello no ocurriera, el resultado será una gran confusión.

Para ello, en RDF el significado se expresa a través de un esquema, que conforma una especie de diccionario, en el que se documentan las

definiciones y restricciones de uso de las propiedades que se utilizarán en una sentencia RDF y le otorgará significados específicos.

Con RDF se pueden utilizar una gran variedad de formas de esquemas, incluyendo una forma específica definida en un documento aparte como el RDF Schema, que tiene algunas características específicas para ayudar a la automatización de tareas usando RDF (W3C-RDF).

Para evitar confusiones entre definiciones independientes y posiblemente conflictivas del mismo término, RDF utiliza la facilidad de los namespace de XML.

Estos namespaces o espacios de nombre, son una forma de asociar el uso específico de una palabra en el contexto del esquema (diccionario) en que se puede encontrar una definición determinada. En RDF, cada predicado utilizado en una sentencia debe ser identificado con un sólo namespace, o esquema.

Por último, cabe señalar que el ámbito de aplicación del RDF no solamente es aplicado en el marco del área de la Documentación; este marco es muy amplio y presenta aplicaciones muy diversas, como por ejemplo la estructuración de la información económica y financiera. La Asociación Española de Contabilidad y Administración de Empresas, ha emitido el documento *Extensible Business Reporting Language (XBRL)*; éste es un Estándar para el Intercambio Electrónico de Información Económica y Financiera. Concretamente el XBRL es una especificación abierta que utiliza un conjunto de etiquetas basadas en el lenguaje de marcas XML para describir la información contenida en los estados financieros.

Ontología

El significado del término ontología se refiere a la descripción formal de los conceptos y relaciones que existen dentro de un dominio. Esto significa que una ontología se relaciona con un vocabulario y un lenguaje específico a diferencia de la filosofía, que aborda dicha cuestión desde la existencia más que desde el lenguaje.

Una ontología representa un acuerdo, el cual no precisa necesariamente abarcar toda la conceptualización de un determinado dominio, pudiendo comprender parte de él, o sea, puede ofrecer una visión para el dominio.

De esta forma, una ontología actúa como una *aparcería* (contrato entre partícipes para un mismo objetivo) que permite una comunicación segura, efectiva y eficiente entre ellos, dentro del contexto del dominio de la información.

Cuando esto ocurre, le permite a un agente de software que esté comprometido con una determinada ontología, ser capaz de interpretar semánticamente los ítems de información comprendidos en esa ontología y comunicarse con otros agentes que también se encuentren comprometidos con esa misma ontología. Es así que una ontología establece una comunidad de usuarios en Internet.

Ontologías y OWL

Las ontologías constituyen bases de conocimiento estructuradas en las cuales se modelan: conceptos, instancias, sus atributos y relaciones, y éstos tienen contrapartes directas en el dominio que se intenta modelar, las cuales conforman la conceptualización antes mencionada.

Por lo tanto, las ontologías son una especificación formal de contenido semántico, conformando un conjunto de conceptos que puede ser utilizado

por agentes para dialogar en un lenguaje común ya que modelan un dominio de manera rigurosa.

Esto nos lleva a que la definición de la semántica de una ontología, se sustenta sobre formalismos basados en la lógica de quienes brindan los axiomas y las reglas de inferencia necesarios para derivar información a partir de la ontología.

La definición de una ontología puede realizarse de diversas maneras, debido a la existencia de varios lenguajes de ontologías. Ante este abanico de posibilidades, la adopción de estándares maduros bien establecidos y aceptados, resulta sumamente importante.

En este sentido, la utilización del *Lenguaje de Ontologías Web* (OWL) como lenguaje para la especificación de la ontología resulta muy conveniente, debido a la gran disponibilidad de herramientas y utilidades existentes y en desarrollo hoy en día.

El *Lenguaje de Ontologías Web* (W3C OWL, 2004), está diseñado para ser usado en aplicaciones que necesitan procesar el contenido de la información, en lugar de únicamente representar información para los humanos.

OWL facilita un mejor mecanismo de interpretación de contenido Web que los mecanismos admitidos por XML, XML Schema, RDF y RDF Schema, proporcionando un vocabulario adicional junto con una semántica formal que permiten definir recursos interrelacionados con semántica explícita.

Sin duda, el principal atractivo de la utilización de una ontología es la posibilidad de razonar y derivar conclusiones a partir de ella. Para ello, es necesario contar con un mecanismo de razonamiento capaz de procesar la ontología, validarla (chequear potenciales inconsistencias) y responder consultas en base a la semántica de la misma.

En este sentido se definen tres sublenguajes de OWL: OWL Lite, OWL DL y OWL Full con su respectivo nivel de expresividad creciente que tiene para la construcción de un razonador completo para ontologías.

El lenguaje a adoptar en este trabajo es el *OWL DL*, debido a que brinda la mejor combinación entre poder expresivo y complejidad de razonamiento.

Una parte esencial en la arquitectura de la Web Semántica, es la capacidad de los productores de contenidos para conformar y compartir la información ontológica, para poder definir la semántica formal del contenido en la Web (Díaz Ortuño, 2003).

Desde el punto de vista de la aplicación, solo podrá representarse correctamente aquello que tenga una descripción formal de los conceptos y de las relaciones existentes entre esos conceptos.

Por lo tanto, las ontologías se usarán para proporcionar una especificación concreta de los nombres y de los significados de los términos y a través de ella, se espera establecer una terminología compartida entre miembros de cada una de las comunidades de interés.

Una vez asumidos estos compromisos ontológicos compartidos, los agentes inteligentes podrán alcanzar un conocimiento compartido parcial y de esta forma operar entre sí.

Con respecto a la representación basada en *Frames* (Marcos) se vio como un paradigma adecuado para la construcción de ontologías, percibiendo que el formalismo RDF Schema puede verse como una manifestación de la representación basada en marcos para la Web.

Los *Frames* son la forma de organizar la información referente a un objeto, en la cual se describen las propiedades del objeto en forma similar a un registro de una base de datos, donde se reservan los campos, para cada propiedad.

Dicha similitud entre los Frames con los registros de las bases de datos, sumado su íntimo parentesco con la programación orientada a objetos, fortaleció la decisión de adoptar el uso del *Unified Modeling Language* (UML) para el diseño y de *Object Oriented Hypermedia Design Method* (OOHDM), como metodología para el desarrollo de las aplicaciones.

Capas de Lógica, Reglas y Trust

La llamada *Capa Lógica*, debe verse como una capa de requerimiento de la Web Semántica e implementada a través de las capas funcionales de XML y RDF (DANGER MERCADERES, 2007).

Lenguajes como el *Lenguaje de Ontologías Web*, permiten definir además del vocabulario ciertos axiomas para deducir nuevas informaciones desde la información explícita, pero no admiten la definición de reglas generales sobre las propiedades. Tal capacidad debería ser provista por la capa de reglas; sin embargo, aún no contamos con el conocimiento de un consenso en cómo ella debe ser.

A continuación se describirán brevemente las alternativas propuestas, a pesar de que ninguna ha recibido especial atención por la falta de una formalización semántica completa:

1. Enfoque basado en XML: este enfoque define un vocabulario XML para crear varios tipos de programas lógicos y una base de conocimiento subyacente. El representante más importante de este enfoque es RuleML Boley, que define una sintaxis para crear una jerarquía de lenguajes de reglas que permiten, a su vez, construir lenguajes cada vez más sofisticados. Hasta ahora, incluye 12 sublenguajes. Su principal limitación es su completitud parcial, en la parte sintáctica y total en la semántica.

2. Enfoque basado en RDFS con soporte para OWL: este enfoque se basa en permitir formar reglas sobre la base de definiciones de los grafos RDF. Algunos de los sistemas, como Euler Roo y CWM (Berners-Lee, 2000), utilizan la "Notación³" por su simplicidad para la escritura y comprensión de las descripciones. Otros, como el Racer Haarslev & Moller, operan directamente con los conocidos grafos de las tripletas de RDF traduciendo éstos a cláusulas en la lógica de Horn y compilados como programas Prolog, que deben llamar a programas externos que permitan manejar los constructores de OWL que no puedan ser manipulados como cláusulas de Horn. Sin embargo, al igual que en el caso anterior, este enfoque no ha recibido especial atención por la falta de formalización semántica.
3. Enfoque basado en OWL: las propuestas basadas en OWL, son más recientes y tratan de combinar OWL y reglas (Grosz, 2003) o extender el lenguaje OWL para adicionar a OWL reglas de cláusulas de Horn, definiendo el lenguaje Semantic Web Rule Language (SWRL). Ambas soluciones corren el riesgo de trabajar sobre conjuntos de datos y propiedades indecidibles, lo que les impide asegurar la terminación de sus algoritmos (Horrocks, 2003a).

La capa de reglas aún se encuentra en un estado muy incipiente, en gran parte porque capas anteriores han completado sus estándares muy recientemente.

Lo mismo ocurre con la *capa Trust* o de la verdad, de la que se desconocen trabajos de formalización y descripción detallada.

El objetivo de esta capa es mantener y promover hacia almacenes de datos, todo conocimiento que se reconozca como verdadero. Ella debe

considerar los problemas relacionados con la autenticidad y autoría de los conocimientos, que se abordan en la capa de Firma Digital.

En consideración a estos aspectos es que el trabajo presenta una descripción más extensa sobre firma digital que el resto de las capas, la que se pasa a describir en el próximo ítem.

Digital Signature en la Web Semántica

La seguridad en las Redes

La seguridad en redes se refiere al intercambio de datos vía redes de forma segura, asegurando que sus datos están protegidos y no podrán ser comprendidos ni modificados por personas no autorizadas.

Para que un sistema de información basado en redes sea considerado seguro, es necesario el uso de una serie de recursos como:

- La *criptografía*, que asigna un nivel determinado de privacidad.
- La *firma digital*, que permite inferir que la información no ha sido modificada.
- El *certificado digital*, que posibilita la identificación de personas.

Principios de la Seguridad de la Información

Los recursos de criptografía, firma digital y certificado digital, serán los responsables por la ejecución de la política de seguridad implantada en el sistema.

Estos recursos pueden presentarse de forma independiente e interactúan para garantizar la seguridad, cumplimentando con los siguientes principios de seguridad en la información:

- Confidencialidad: los componentes del sistema serán accesibles sólo por aquellos usuarios autorizados.
- Integridad: los componentes del sistema sólo pueden ser creados y modificados por los usuarios autorizados.
- Disponibilidad: los usuarios deben tener disponibles todos los componentes del sistema cuando así lo deseen.

Adicionalmente se debe considerar:

- Autenticidad: busca asegurar la validez de la información en tiempo, forma y distribución. Asimismo, se garantiza el origen de la información, validando el emisor para evitar suplantación de identidades.
- Auditabilidad: define que todos los eventos de un sistema deben poder ser registrados para su control posterior.
- Protección a la duplicación: consiste en asegurar que una transacción sólo se realiza una vez, a menos que se especifique lo contrario. Impedir que se grabe una transacción para luego reproducirla, con el objeto de simular múltiples peticiones del mismo remitente original.
- No repudio: se refiere a evitar que una entidad que haya enviado o recibido información alegue ante terceros que no la envió o recibió.
- Legalidad: referido al cumplimiento de las leyes, normas, reglamentaciones o disposiciones a las que está sujeta la empresa.
- Confiabilidad de la Información: es decir, que la información generada sea adecuada para sustentar la toma de decisiones y la ejecución de las misiones y funciones.

Todos los principios presentados tienen funciones específicas, con el objetivo de instalar barreras para la restricción selectiva del acceso a la información del sistema.

Para que estas barreras sean selectivas (no definitivas) puedan ofrecer a los usuarios autorizados, los servicios disponibles en la red, éstas deben ser implementadas según estándares como la norma IRAM-ISO IEC 17799 (ISO 17799:2000).

Criptografía

La criptografía es un método eficaz que ofrece protección, tanto al almacenamiento como al transporte de la información por la red.

Mediante este método que agrupa a un conjunto de técnicas, se puede codificar datos en información aparentemente sin sentido o inentendible, con el fin de que las personas no autorizadas no puedan entenderlo y así no acceder a la información que fuese cifrada. Para ello se aplican funciones matemáticas y un código especial denominado *llave*, sin la cual la información no puede ser descifrada.

Esta *llave* enmarcada en los métodos de encriptado, asegurará las transmisiones seguras en la red y en entornos abiertos.

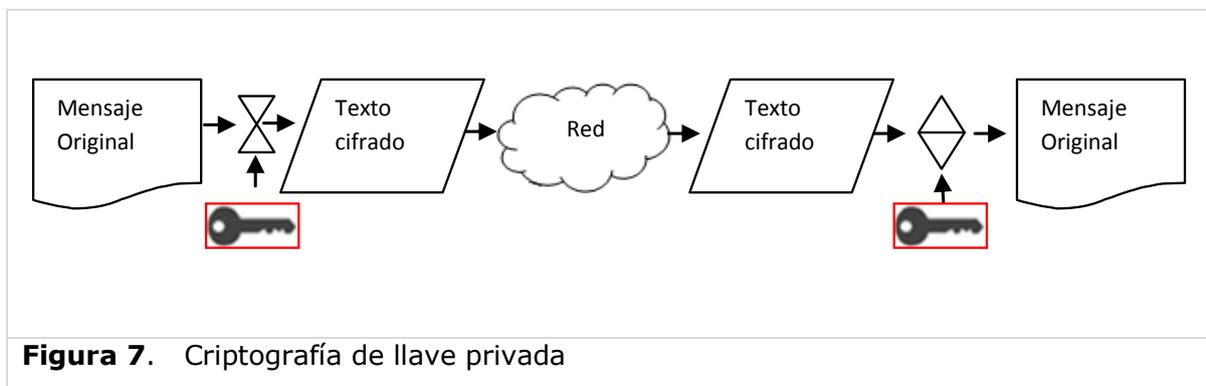
Hay dos tipos de llaves: la llave privada o *encriptado simétrico*, y la llave pública o *encriptado asimétrico*.

La Criptografía de llave privada ocurre cuando existe una única llave que encripta y desencripta los datos, para lo cual la llave es secreta y solamente compartida entre el emisor y el destinatario.

Ésta, llave secreta, debe ser una llave relativamente pequeña y rápida; siendo su mayor problema el de su distribución, lo que hace que la eficiencia de este método dependa principalmente de la seguridad en la conexión para su distribución.

El procedimiento comprende tres pasos (ver figura 7):

1. El emisor cifra un mensaje utilizando un algoritmo criptográfico y una llave secreta, para transformar un mensaje original en un texto cifrado.
2. El texto cifrado se envía al destinatario, asumiendo que la distribución se realiza por una red insegura.
3. El receptor destinatario, para descifrar el mensaje utiliza el algoritmo y la llave secreta correspondiente; lo que permite transformar el texto cifrado en el mensaje original.



La seguridad del sistema radica en mantener en secreto la llave y no el algoritmo; siendo que el algoritmo frecuentemente utilizado es el *Advanced Encryption Standard (AES)*, con llaves de 128 bits, 192 bits y 256 bits.

Entre las técnicas de cifrado simétrica podemos citar:

1. Cifrado simétrico por bloques: códigos Feistel entre los que encontramos el DEA (DES), Triple DES y variantes, G-DES, DESX; Blowfish, RC2, RC5, IDEA, FEAL y variantes

2. Cifrado simétrico continuo: Vernam, RC4, SEAL
3. Funciones unidireccionales (hash) de uso criptográfico: Havies-Meyer (basada en DES), MD2, MD5, SHA1, RIPE-MD, Haval
4. Variantes especiales de firma digital: Firma a ciegas, Firmas con verificador designado, Firmas de grupo, Firmas no reusables, Firmas sin propagación de fallo, Firmas no repudiables, Sello de tiempo.

El método de Criptografía de llaves públicas o criptosistema asimétrico, está basado en el concepto de dos llaves: una pública y otra privada. La llave pública es utilizada para encriptar mensajes y se encuentra públicamente disponible; en tanto que la llave privada se utiliza para descifrar mensajes y se debe mantener en secreto.

El factor crítico de este procedimiento depende del resguardo de la llave privada.

El procedimiento comprende dos pasos (ver figura 8):

1. El emisor cifra el mensaje utilizando la llave pública del destinatario y envía el mensaje.
2. El receptor recibe el mensaje y descifra el mensaje con su llave privada.

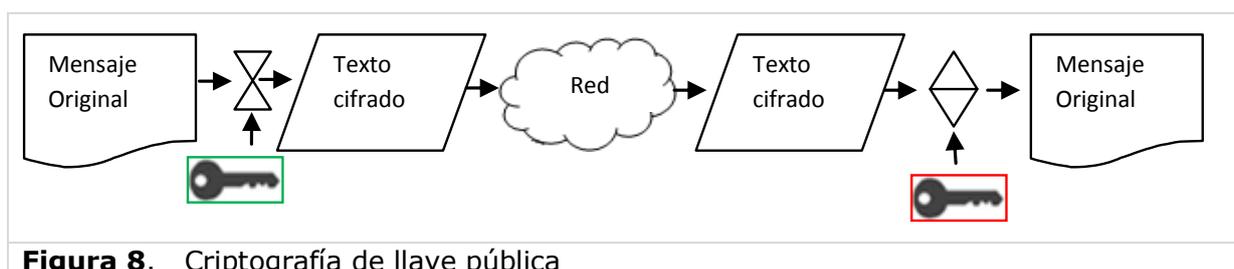


Figura 8. Criptografía de llave pública

Con este método el emisor no puede descifrar o descifrar el mensaje, ya que no cuenta con la llave secreta del receptor; lo que asegura que cualquier intruso que pudiera interceptar el mensaje tampoco podrá descifrar el mensaje, por más que cuente con la llave pública.

Generalmente la criptografía de llave pública es adecuada en ambientes inseguros, que se caractericen por contar, con un gran número de usuarios. Para asignarle mayor nivel de seguridad, este sistema, puede ser complementado con la técnica de llaves privadas.

El algoritmo Diseñado en 1977 por Rivest, Shamir y Adleman (RSA) es el método de criptografía que satisface los requisitos exigidos para que la llave pública sea eficaz; este algoritmo usa claves de de 512, 768, 1024 o 2048 (típico 1024) bits formadas por la multiplicación de dos números primos de 512 bits, lo que garantiza la seguridad del proceso, que se basa en la dificultad de factorizar números primos muy grandes.

Entre las técnicas de cifrado asimétrico o firmas digitales podemos citar:

- Basado en factorización: RSA y el Protocolo de firmado Rabin
- Basado en logaritmos discretos: Diffie-Hellman, diseñado en 1977 por Diffie y Hellman. Usa claves de 512, 1024 es utilizado mayoritariamente para negociar claves; necesita autenticación adicional (man-in-the-middle) y no soporta firmas digitales, STS (DH con acuerdo de claves autenticado) y ElGamal, diseñado en 1984 por Taher ElGamal puede realizar cifrado y firma basado en la dificultad de calcular logaritmos discretos y DSA (DSS), diseñado por el NIST (National Institute of Standards and Technology) inicialmente se utilizaban claves de 512 y posteriormente se incrementó a 1024 para mayor seguridad; es una variante de Schnorr y ElGamal - Sólo soporta firmas digitales

Firma Digital

La Firma Digital es una modalidad de firma electrónica, resultado de una operación matemática que utiliza algoritmos de criptografía asimétrica y permite inferir, con seguridad, el origen y la integridad del documento.

Para crear una firma digital se deben cumplimentar los siguientes pasos (ver figura 9):

1. Calcular del resumen del mensaje: el valor de *hash* del mensaje, el cual comúnmente es denominado *resumen del mensaje*, es calculado por la aplicación de un algoritmo criptográfico de hashing, como: MD2, MD4, MD5 o SHA1.
2. Calcular de la firma digital: la información obtenida en el primer paso (resumen del mensaje) es cifrada con la llave privada de la persona que firma el mensaje y así el valor que es obtenido, conforma la firma digital. Los algoritmos, frecuentemente utilizados, para el cálculo de la firma digital de un determinado resumen son: RSA, basado en la teoría de los números; DSA, basado en la teoría de los logaritmos discretos; o el ECDSA, basado en la teoría de las curvas elípticas.

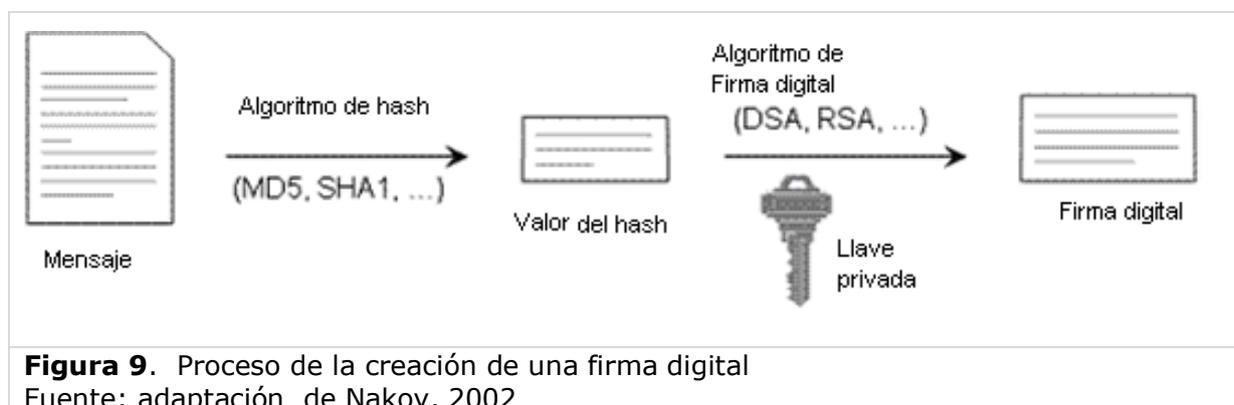
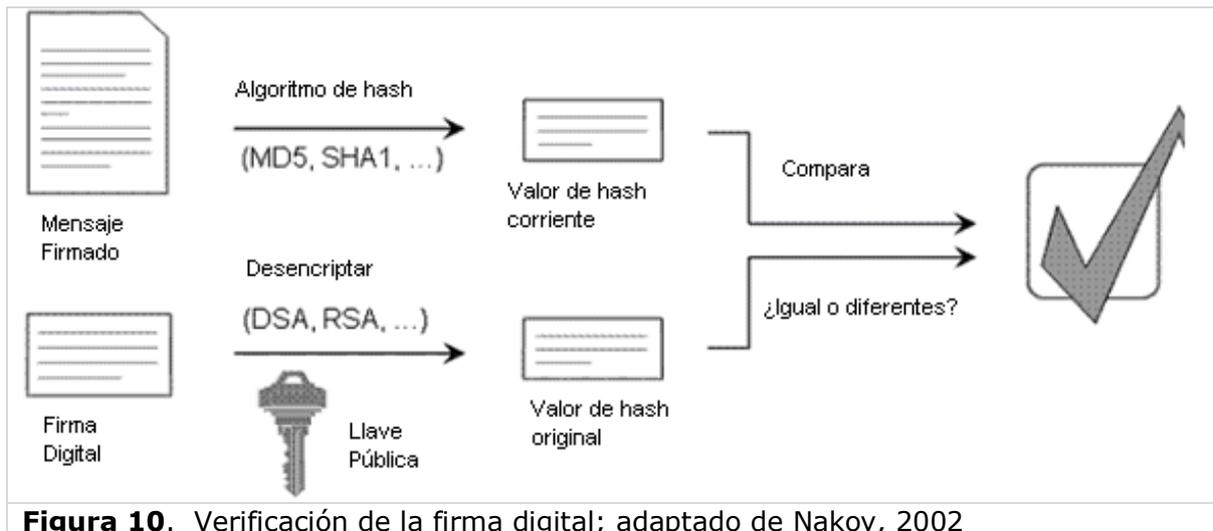


Figura 9. Proceso de la creación de una firma digital
Fuente: adaptación de Nakov, 2002

La tecnología de la firma digital permite, al receptor de un mensaje con una firma digital, verificar su integridad. El proceso de verificación busca determinar si un mensaje fue firmado por la llave privada que corresponde a una dada llave pública. La verificación de la firma digital no puede determinar si un mensaje fue firmado por una determinada entidad, para ello se precisaría obtener la llave pública de alguna manera segura a través de procedimientos estipulados.

Para verificar la integridad de una firma digital se debe (ver figura 10):

1. Calcular el valor corriente del hash: calcular un valor de hash del mensaje firmado; para este cálculo es usado el mismo algoritmo tal cual fue aplicado en el proceso de generación de la firma digital. El valor obtenido es denominado valor de hash corriente, pues él es creado a partir del estado actual del mensaje.
2. Calcular el valor original del hash: la firma digital es descifrada con el mismo algoritmo utilizado durante la generación de la firma digital. El descifrado es realizado con la llave pública asociada a la llave privada utilizada durante la firma del mensaje. Como resultado, se obtiene el valor original de hash que fue calculado del mensaje original durante el primer paso de la creación de la firma digital (el valor original del resumen del mensaje – valor de hash).
3. Comparar el valor corriente original de hash: comparar el valor corriente del hash obtenido en el primer paso con el valor original del hash obtenido en el segundo paso. Si los dos valores son idénticos, prueba que el mensaje fue firmado con la llave privada que corresponde a la llave pública usada en la certificación.



Estos últimos dos procesos de encriptado, clave pública y firma digital, pueden combinarse para obtener intimidad y autenticación al mismo tiempo. Esto se logra si se firma primero el mensaje con la clave secreta y se encripta después el mensaje firmado con la clave pública del destinatario.

Si el destinatario sigue estos pasos en sentido contrario al desencriptar primero el mensaje con su propia clave secreta y comprobar después la firma con la clave pública del remitente (PGP MIT).

Este proceso que provee una solución de seguridad, permitiendo asegurar la identidad de una o de todas las partes comprometidas en una transacción, es el *Certificado Digital*.

Por lo tanto, el *Certificado Digital* es un documento conteniendo datos de identificación de la persona o institución que desea comprobar su propia identidad, como así también confirmar la identidad de terceros.

Los certificados digitales, vinculan un par de llaves electrónicas que pueden ser usadas para encriptar y firmar información digital.

El certificado digital puede ser firmado por alguien en quien el origen deposita su confianza y es denominada autoridad de certificación (Certification Authority), la que funciona como un registro electrónico.

Concretamente el certificado digital funciona de la siguiente manera (ver figura 11):

1. Se debe localizar la llave pública de la persona con quien se desea comunicar.
2. Se debe obtener una garantía de que la llave pública encontrada sea proveniente del destinatario.

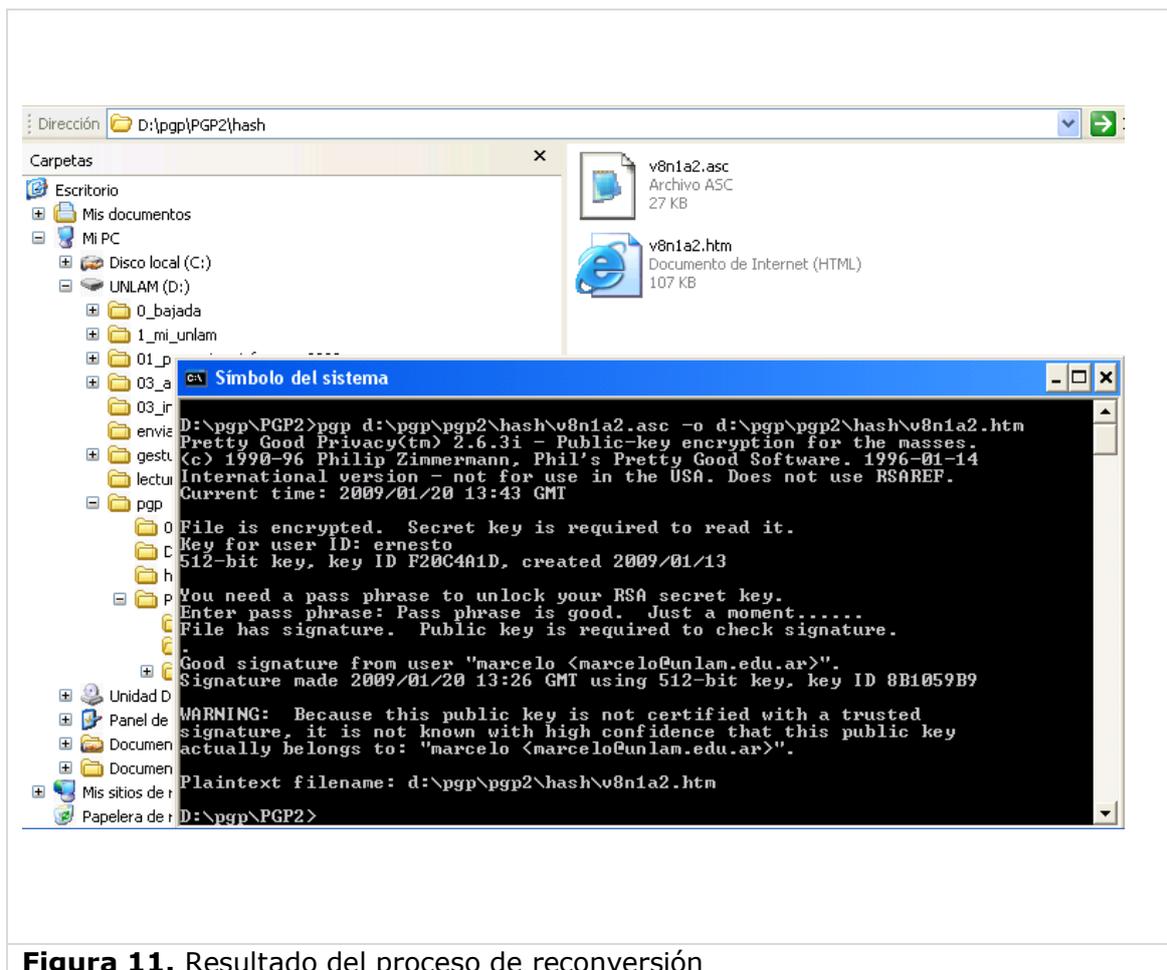


Figura 11. Resultado del proceso de reconversión

Aplicación de la firma digital en la Web Semántica

Ya vimos que la Web semántica es un conjunto de iniciativas destinadas a convertir la World Wide Web en una gran base de datos capaz de soportar un procesamiento sistemático y consistente de la Información y que estructurando la información en XML se permite a los agentes de software leer e intercambiar los datos.

Esto se constituye en un primer paso hacia la producción de una Web de semántica legible por los computadores.

Pero esta utilidad en la estructuración de la información está limitada si no pueden verificarse su autenticidad y fidelidad, y para ello necesita Firmar-XML.

Criptografía en XML

XML Encryption es un lenguaje cuya función principal es asegurar la confidencialidad de partes de documentos XML, realizado a través de la encriptación parcial o total del documento transportado.

El algoritmo Rivest, Shamir y Adleman (RSA) en Extensible Markup Language (XML) no es utilizado para encriptar documentos, sí para encriptar la llave de sesión, que queda junto al documento XML en un elemento separado y localizado en otro archivo. El algoritmo utilizado en la criptografía del documento es el Advanced Encryption Standard (AES) con claves 128 o 256.

A continuación en la figura 12 se muestra un ejemplo aplicado a un recurso científico, del tipo artículo.

```

<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:rdf = http://www.w3.org/1999/02/22-rdf-syntax-ns#   xmlns:dc =
"http://purl.org/metadata/dublin_core#"
  xmlns:dcq = "http://purl.org/metadata/dublin_core_qualifiers#">
  <rdf:description about = "http://www.cyta.com.ar/ta0702/v7n2a2.htm">
    <dc:title>
      <rdf:Bag>
        <rdf:li>Evidencias conceptuales sobre intangibles: una revisión doctrinal</rdf:li>
        <rdf:li>Conceptual evidences on intangibles: a doctrinal revision</rdf:li>
      </rdf:Bag>
    </dc:title>

    ....

    <EncryptedData xmlns = "http://www.w3.org/2001/04/xmlenc#" Type =
"http://www.w3.org/2001/04/xmlenc#">
      <CipherData>
        <CipherValue>Â°►J-=9Â¨gÃ¿'Ã 9iNÃ¢W</CipherValue>
      </CipherData>
    </EncryptedData>
  </rdf:li>
  <rdf:li>
    <EncryptedData xmlns = "http://www.w3.org/2001/04/xmlenc#" Type =
"http://www.w3.org/2001/04/xmlenc#">
      <CipherData>
        <CipherValue>23e0199ee9495df1542a4312591</CipherValue>
      </CipherData>
    </EncryptedData>
  </rdf:li>
  </rdf:Bag>
</dc:contributor.referee>
<dc:type>Research article</dc:type>
<dc:format>htm</dc:format>
  </rdf:description>
</rdf:RDF>

```

Figura 12. Ejemplo de un documento XML encriptado

Tal como se puede apreciar en la figura número 12, dentro del elemento EncryptedData se encuentran los elementos encriptados. Su estructura es de fácil comprensión y visualización, permitiendo la criptografía de un elemento en particular o la del documento completo.

En el tag de la criptografía, la llave de sesión tiene que ser generada, encriptada y debe ser incluida información sobre ella.

Firma digital en XML

XML Signature tiene por objetivo principal, asegurar la integridad de partes de los documentos XML transportados.

También, proporciona la autenticación de mensajes y servicios de autenticación de firma, para datos en el XML que incluye la firma o en cualquier otra parte.

Lo que hace principalmente *XML Signature*, es asociar claves con los datos de consulta.

XML Signature representa un sistema, que a través de una firma digital, permite ofrecer autenticidad de los datos.

El estándar XML permite que estos documentos puedan ser firmados digitalmente, poseyendo las características de autenticación e integridad de los datos, sin olvidar que los mensajes no serán repudiados.

La firma digital en documentos XML tiene la ventaja de poder ser firmadas partes específicas del documento o, si es necesario, la firma del documento completo.

Pero si ocurriera que la firma digital fuese realizada sobre una parte del documento, no se podría garantizar la integridad total del documento, dado que algunas partes del mismo podrían ser alteradas.

En la figura 13 se muestra la estructura de una firma digital en XML, aplicando el código hash.

```

<Signature Id="EjemploXMLSignature"
xmlns="http://www.w3.org/2000/09/xmldsig#">
  <SignedInfo>
    <CanonicalizationMethod
      Algorithm="http://www.w3.org/TR/2001/REC-xml-c14n-20010315"/>
    <SignatureMethod
      Algorithm="http://www.w3.org/2000/09/xmldsig#dsa-sha1"/>
    <Reference
      URI="http://www.w3.org/TR/2000/REC-xhtml1-20000126/">
      <DigestMethod
        Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/>
      <Digest Value>j6lwx3rvEPO0vKtMup4NbeVu8nk=</Digest Value>
    </Reference>
  </SignedInfo>
  <Signature Value>MC0CFrVLtRlk=...</Signature Value>
  <KeyInfo>
    <KeyValue>
      <DSAKeyValue>
        <p>...</p><q>...</q><g>...</g><y>...</y>
      </DSAKeyValue>
    </KeyValue>
  </KeyInfo>
</Signature>

```

Figura 13. Ejemplo de la estructura de una firma digital usando XML

El elemento `Signature`, encapsula la firma digital y contiene tres sub-elementos: `SignedInfo`, `SignatureValue` y `KeyInfo`.

El elemento `SignedInfo`, contiene información sobre qué es lo que se firma y cómo se firma, es decir, contiene la información necesaria para crear y validar la firma.

Este elemento contiene dos algoritmos, en donde por un lado está el `<CanonicalizationMethod>` que es el algoritmo de transformación de `SignedInfo` antes de realizar la firma digital; y por otro lado, estaría el método de firma `<SignatureMethod>`, que sería el algoritmo utilizado para calcular el valor de la firma digital.

El elemento `<CanonicalizationMethod>` es el encargado de indicar el algoritmo para canonizar el elemento `SignedInfo`, que tendrá lugar durante la creación de la firma.

El `<SignatureMethod>`, es el encargado de indicar el algoritmo para generar la firma a partir de la canonización de `SignedInfo`.

También se incluye en el elemento `SignedInfo` las referencias a los objetos que se van a firmar (`<Reference>`) que incluye además `<DigestMethod>` y `<DigestValue>`.

Recordemos que la validación de una firma requiere dos procesos, que son: la validación de la firma y la validación de los resultados de las referencias.

El resultado obtenido se indicará en el elemento `SignatureValue`. Cada elemento `<Reference>`, incluye una referencia al objeto que se firmará.

Al mismo tiempo incluye el resultado de `<DigestValue>` que es el valor resultante.

El elemento `<SignatureValue>`, contiene el resultado de la firma digital que se ha aplicado sobre el elemento `SignedInfo`.

El resultado de esta firma está codificado y contiene un atributo que es único con el que se identificará la firma en procesos posteriores de validación.

El elemento `<KeyInfo>`, es un elemento opcional que indica la clave que ha de utilizarse para validar la firma.

El elemento `<KeyValue>`, especifica la clave para validar la firma digital.

Resumiendo, tenemos que:

- El elemento SignedInfo contiene lo que se firma.
- El elemento SignatureValue, contiene la firma; es decir, contiene el elemento SignedInfo en forma canonizada, resumida y encriptada con la clave pública del firmante.
- El elemento KeyInfo, contiene el certificado de la clave pública del firmante.

Para ello se asumieron los modelos propuestos en las normas de XML Signature y XML Criptografiado por la W3G, en base a ellos se han construido los bloques de datos encriptados como forma de proteger los valores utilizados que aseguren su confidencialidad y la firma digital que permite garantizar autoría (autenticidad) e integridad (fidelidad) del recurso.

Siendo que estos bloques serán utilizados, por las computadoras y los agentes, para verificar que la información adjunta ha sido ofrecida por una fuente específica y confiable.

A continuación se muestra en la figura 14, un ejemplo aplicado a un recurso científico en el que se destaca el tag de la criptografía EncryptedData, conteniendo el valor del elemento dc:contributor.referee criptografiado.

Para dicho ejemplo se han utilizado las funciones, en la Base de Datos MySQL, de SHA1 y AES para el encriptamiento mediante el modelo de llave única.

```

<dc:contributor>Universidade Estadual da Paraíba </dc:contributor>
  <dc:contributor.referee>
    <rdf:Bag>
      <rdf:li>
        <EncryptedData xmlns = "http://www.w3.org/2001/04/xmlenc#"
Type = "http://www.w3.org/2001/04/xmlenc#">
          <CipherData>
            <CipherValue>Â°▶J-=9Â¨gÃ¿'Ã 9iNÃçW</CipherValue>
          </CipherData>
        </EncryptedData>
      </rdf:li>
      <rdf:li>
        <EncryptedData xmlns = "http://www.w3.org/2001/04/xmlenc#"
Type = "http://www.w3.org/2001/04/xmlenc#">
          <CipherData>
            <CipherValue>23e0199ee9495df1542a4312591</CipherValue>
          </CipherData>
        </EncryptedData>
      </rdf:li>
    </rdf:Bag>
  </dc:contributor.referee>
  <dc:type>Research article</dc:type>
  <dc:format>htm</dc:format>
  <dc:coverage>ADMG</dc:coverage>
  <dc:date>2008-04-15</dc:date>
  <dc:date>2008-04-15</dc:date>
</rdf:description>
</rdf:RDF>

```

Figura 14. Aplicación del hash al campo del evaluador en el modelo DC en XML

CAPÍTULO III
DISEÑO DE LAS APLICACIONES PARA
LA BIBLIOTECA DIGITAL

Capítulo III - Diseño de las aplicaciones para la biblioteca digital

Selección del modelo

Para el diseño de las aplicaciones de la biblioteca digital, hemos basado la selección de la metodología en procurar un diseño evolutivo y creativo (HENRY C., LUCAS Jr.).

En donde la propuesta de diseño evolutivo, debe permitir a los usuarios ejercer influencia sobre el modelo de sistema, y a su vez crear un compromiso psicológico hacia él y hacia el sentido de propiedad del sistema.

Es por ello que la decisión recayó sobre la metodología *Object Oriented Hypermedia Design Method* (OOHDM); donde su principal propiedad, en función de los requerimientos del proyecto, radica en la identificación de tres niveles, similares a la arquitectura Cliente/Servidor Multicapas – ANSI/SPARC y que son:

- El dominio de aplicación.
- La necesidad de proveer acceso de navegación simple a grandes cantidades de datos multimediales.
- La de construir interfaces web fáciles de usar.

Object Oriented Hypermedia Design Method (OOHDM)

El Object Oriented Hypermedia Design Method (OOHDM), propone el desarrollo de aplicaciones hipermedia a través de un proceso compuesto por cuatro fases: diseño conceptual, diseño de navegación, diseño de interfaces abstractas e implementación (ver figura 15).

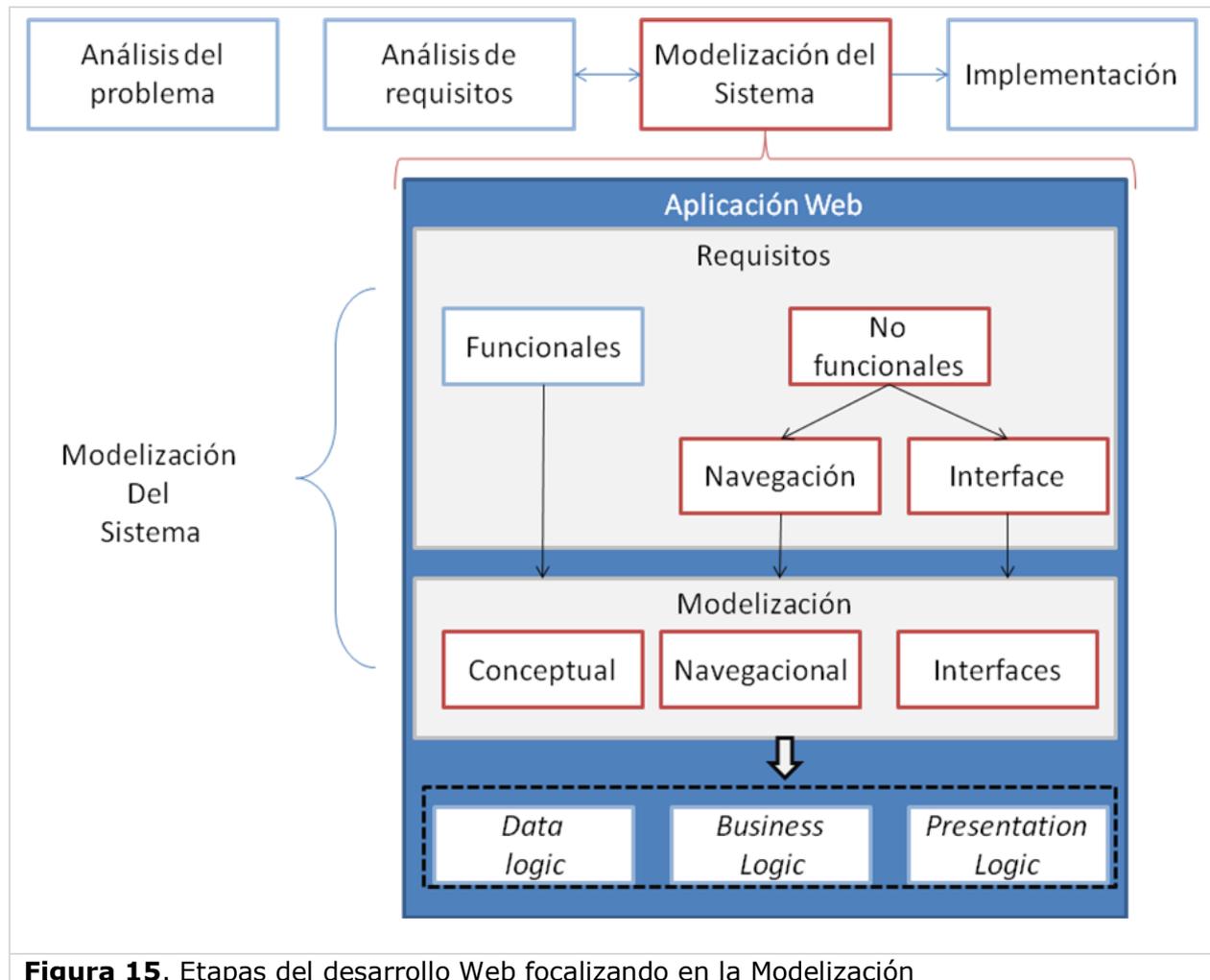


Figura 15. Etapas del desarrollo Web focalizando en la Modelización

El aspecto más importante al cual nos vamos a referir en el próximo tópico es el de documentar y exportar las semánticas.

Análisis de requisitos respecto a la semántica

Es necesario en este contexto que todas las personas deban entender el diseño y mapear sus elementos. Además, los partícipes deben poder hacer fácilmente aseveraciones sobre los elementos y no tener que hacer suposiciones sobre el propósito de los mismos.

A continuación veremos ciertos requisitos del sistema y el planteo respectivo, a cada uno de ellos, de algunas técnicas apropiadas para su consideración.

Documentación legible

Garantizar documentación legible para todos los actores partícipes; todos deben poder leer y entender una especificación técnica, incluso una que esté dirigida a un determinado tipo de usuario en particular.

Técnicas: aplicando por ejemplo el atributo `title` a las etiquetas de imágenes en archivos HTML *accesibles*. Los archivos XML deben contar con su respectivo esquema.

Recuperación del esquema de un documento

Proporcionar un mecanismo inteligible para obtener el esquema de un documento; los programas deben poder recuperar automáticamente la documentación en un idioma determinado.

Técnica: utilizar como esquema el lenguaje *XML Schema*, haciendo referencia a él a través del atributo `xsi:schemaLocation`.

Ejemplo:

```
<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/metadata/dublin_core#"
  xmlns:dcq="http://purl.org/metadata/dublin_core_qualifiers#"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.cyta.com.ar/xml/rdf_dc.xsd">
  <rdf:description about="http://www.cyta.com.ar/ta0101/tapa.htm">
  ...
```

Marcado semántico

Proporcionar definiciones explícitas legibles por los partícipes en el marcado semántico. Para que el diseñador del esquema y los autores de documentos utilicen las mismas semánticas, éstas requerirán ser

documentadas; siendo que toda mejora en el proceso de documentación, redundará en una mejor comprensión compartida.

Ejemplo:

```
<?xml version="1.0" encoding="utf-8" ?>
<xsd:schema xmlns:xsd="http://www.w3c.org/2001/XMLSchema"
  targetNamespace="http://www.cyta.com.ar/xml/rdf_dc.xsd">
  <xsd:annotation>
  <xsd:documentation xml:lang="es-AR">Esquema de articulos de la
    revista Técnica Administrativa ISSN 1666-680
  /xsd:documentation>
  </xsd:annotation>
  <xsd:complexType name="rdf:description">
  <xsd:sequence>
  <xsd:element name="dc:title" type="xsd:string" />
    <xsd:annotation>
      El nombre que se ha dado al recurso
    </xsd:annotation>
  </element>
```

Esquema semántico

Utilizar el esquema para proporcionar documentación/anotación explícita de las semánticas de los elementos. El Esquema permite a los diseñadores de lenguajes añadir explícita documentación para los elementos y atributos, y de esta manera se hace más entendible el lenguaje.

Técnica: el ejemplo escribe claramente la necesidad de usar el elemento de encabezado.

```
<xsd:element name="dc:subject" type="xsd:string" />
  <xsd:annotation>
  <xsd:documentation xml:lang="es-AR"> Palabra clave. Derivada de un
    tesauo públicamente reconocido
  </xsd:documentation>
  </xsd:annotation>
</xsd:element>
```

Relaciones semánticas

Proporcionar relaciones semánticas con otras schemata, cuando sea apropiado y posible. Esto permite a los autores que usan el lenguaje, reutilizar sus conocimientos previos y las herramientas que ya tenían.

Técnica: esto puede hacerse implícitamente a través de la derivación/subclasificación de tipos existentes, por medio de la aserción de equivalencia de los tipos (ejemplo: el título en SVG) o trazando correspondencias con semánticas bien conocidas.

Ejemplo: creando una correspondencia entre el ejemplo de dcrelation proporcionado en la introducción a XHTML, utilizando XSLT:

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/TR/WD-xsl">
<xsl:template match="/">
<html>
<body>
<table border="1">
<tr>
<td>Título</td>
<td>Relación</td>
</tr>
<xsl:for-each select="*/dc" order-by="dctitle">
<tr>
<td>
<xsl:value-of select="dctitle" />
</td>
<td>
<xsl:if test="dcrelation[.='ADMG']">
<xsl:value-of select="dcrelation" />
</xsl:if>
<xsl:if test="dcrelation[.='MCI']">
<xsl:value-of select="dcrelation" />
</xsl:if>
</td>
</tr>
</xsl:for-each>
</table>
</body>
</html>
</xsl:template>
</xsl:stylesheet>

```

Información semántica del elemento

No asumir que los nombres de elemento o atributo proporcionan alguna información sobre la semántica del elemento. Un elemento nombrado puede tener un significado completamente contextualizado para el autor del esquema, pero es poco probable que signifique lo mismo para una persona que no habla la misma lengua o bien es parte de otra comunidad de conocimiento. Sacado de contexto, sin explicación semántica, los nombres de elemento tienden a perder su significado. Nombrar un elemento simplemente, no es suficiente para garantizar que los autores de documentos utilizarán ese elemento en conformidad semántica con la intención del autor del esquema.

Ejemplo: equivocado

```
<xsd:element name="dc:title" type="xsd:string" />
  <xsd:annotation>
    Título
  </xsd:annotation>
</element>
```

Aquí el nombre de elemento ha sido descrito utilizando sólo el nombre del elemento, lo que no añade valor semántico.

Ejemplo: Correcto

```
<xsd:element name="dc:title" type="xsd:string" />
  <xsd:annotation>
    <xsd:documentation>
      El nombre que se ha dado al recurso
    </xsd:documentation>
  </xsd:annotation>
</element>
```

Aquí el nombre de elemento ha sido descrito de manera alternativa para aclarar el significado semántico en vez de solo reforzar el nombre, repitiéndolo.

Estructuras navegables

Documentar las estructuras navegables. Describir cómo deben trabajar los dispositivos, secuenciales, estructurados, y los de navegación de búsqueda

Para navegar alrededor de un documento significativo, es útil para el lector saber qué elementos están disponibles para dicha navegación.

Técnica: acceso aleatorio a cualquier parte del documento vía una tabla de contenidos detallada, títulos numerados que pueden ser buscados, una vista jerárquica (basado en algún modelo taxonómico) habilitando el acceso rápido a las partes buscadas, y ayudas para la búsqueda.

CAPÍTULO IV
DESARROLLO DEL SISTEMA DE
GESTIÓN DEL CONOCIMIENTO

Capítulo IV – Desarrollo del sistema de gestión del conocimiento

Modelo Conceptual

En el modelo conceptual se describió todo el universo de la información que se consideró relevante, independientemente si se trataba de información dinámica o estática, con el fin de poder responder a todas las funcionalidades a las que se pueden presentar.

El Hipertexto como recurso de divulgación

En esta actividad de modelado, se buscó contar con una definición clara del objetivo que permita elaborar el programa de actividades del proyecto.

Para ello, se ha utilizado la herramienta del mapa conceptual (ver figura 16 y 17) como técnica para representar y organizar el conocimiento, empleando conceptos y frases de enlace entre estos conceptos, que atienda a la norma ISO/IEC 13250.

En la figura 16 pueden observarse los requerimientos de un recurso digital técnico o científico, desde la perspectiva de:

- Los contenidos.
- La estructura.
- La técnica de edición para la Web Semántica.

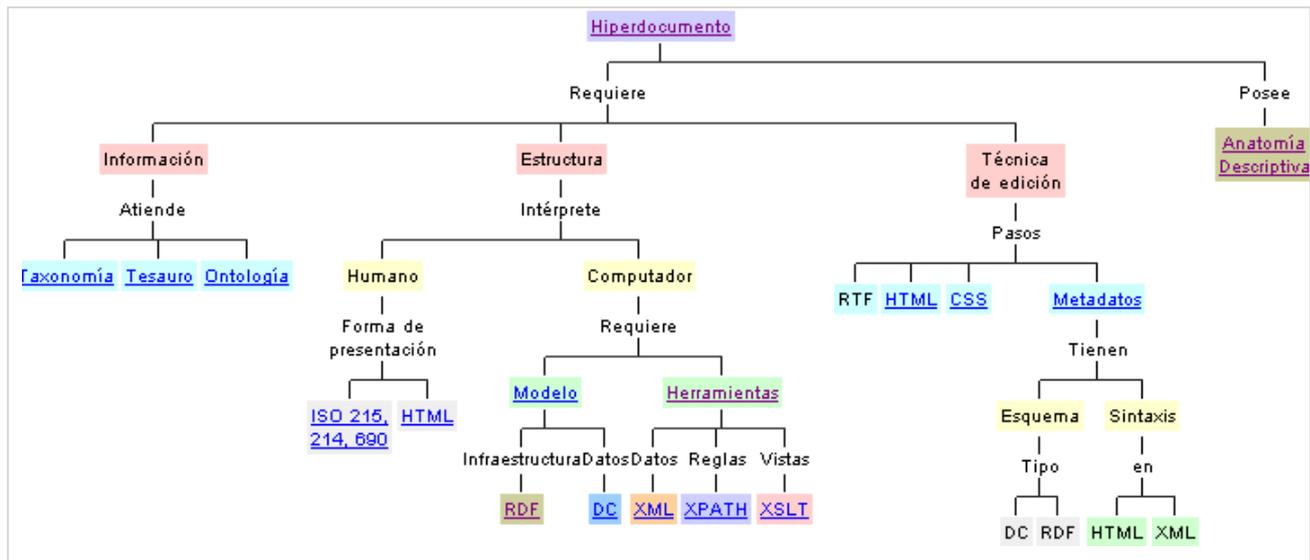


Figura 16. Mapa conceptual de los requerimientos del hipertexto

De estos requerimientos representados en la figura 16, en su mayor parte, ya nos hemos ocupado en los capítulos precedentes.

En cuanto a los requerimientos del recurso tanto de su concepción, composición y propiedades, son representados en la figura 17; y de los cuales describiremos a continuación (Rovira, 2007).



Figura 17. Mapa conceptual de la anatomía del hipertexto

Concepción, composición y propiedades del hipertexto

Los catálogos y directorios de publicaciones científicas, atentos a las transformaciones que están operándose en la comunicación científica y

sus paradigmas, y respondiendo al número creciente de recursos científicos y técnicos que han optado por la edición electrónica, ha definido criterios de calidad editorial.

A modo de ejemplo podemos listar algunos de estos criterios:

1. Mención del cuerpo Editorial

- a. Contenido
- b. Antigüedad mínima 1 año - 4 años
- c. Identificación de los autores
- d. Entidad editora
- e. Mención del Director
- f. URL del recurso

2. Características de presentación

- a. Navegación y funcionalidad (Navegación estructural que permita con un número prudente de clics acceder a sumarios y contenidos, así como descargar e imprimir sumarios y textos)
- b. Mención de periodicidad (o en su caso declaración de periodicidad continuada y cadencia con la que se cerrarán los volúmenes o números, o se incorporaran trabajos. Número mínimo de trabajos/año)
- c. Tabla de contenidos (presencia del sumario de la última entrega así como de los sumarios anteriores)
- d. Números publicados (Enlace/acceso a los artículos de otros números anteriores)
- e. Membrete bibliográfico al inicio del recurso (Al inicio de cada recurso deberá hacerse constar el título, año, vol. y número único de identificador)
- f. Filiación de los autores (Deberá hacerse constar siempre tanto la entidad a la que está adscrito el autor como su dirección electrónica o url)

- g. Recepción y aceptación de originales (estas fechas revisten en las revistas electrónicas aun más importancia)

3. Características de funcionales de los recursos

- a. Identificador único (las versiones electrónicas de las revistas convencionales deben de contar con un ISSN)
- b. Definición del recurso (breve descripción de los objetivos científicos de la publicación y mención del público al que se dirige)
- c. Sistema de arbitraje (Se deberá hacer constar el procedimiento empleado para la selección de originales para publicación)
- d. Evaluadores externos (Se considerarán externos aquellos evaluadores ajenos a la entidad editora)
- e. Autores externos (Al menos el 50% de los trabajos deberán ser de autores ajenos a la institución que edita la revista y a la directiva de la Asociación profesional o científica cuando esta es la entidad editora)
- f. Apertura editorial (Al menos 2/3 del comité editorial deberá ser ajeno a la entidad que edita la revista)
- g. Servicios de información (Califica positivamente si la revista esta indizada sistemáticamente en alguna base de datos bibliográfica)
- h. Cumplimiento de la periodicidad (Califica positivamente si la revista cumple la declaración de periodicidad que se contempla en el criterio 9)

4. Características de los contenidos

- a. Contenido original (Al menos el 40% de los artículos son trabajos de investigación, comunicación científica o creación originales)
- b. Instrucciones a los autores (Califica si aparecen claramente las instrucciones a los autores sobre el envío de originales y la elaboración de resúmenes)

- c. Elaboración de las referencias bibliográficas (En las normas a los autores deberán indicarse las instrucciones para redactar las referencias bibliográficas)
- d. Exigencia de originalidad (Califica positivamente si en la presentación de la revista o en las instrucciones a los autores se menciona esta exigencia de manera explícita)
- e. Resumen (Todos los artículos deberán ir acompañados de un resumen en el idioma original del trabajo)
- f. Resumen en dos idiomas (Califica si además de un resumen de cada artículo en el idioma original del trabajo acompaña otro en una lengua de amplia difusión)
- g. Palabras clave (Califica positivamente si se incluyen palabras clave en el idioma original del trabajo)
- h. Palabras clave en dos idiomas (Califica positivamente si se incluyen palabras clave en el idioma original del trabajo y en un segundo idioma de amplia difusión)
- i. Metaetiquetas (Dublín Core) (Califica positivamente si aparecen metaetiquetas siguiendo la norma Dublín Core en la página de presentación de la revista -código fuente-)
- j. Buscadores (Califica positivamente la presencia de algún motor de búsqueda que permita realizar búsquedas por palabras, por índices, utilizar operadores booleanos)
- k. Servicios de valor añadido (Califica positivamente si la revista ofrece alertas, enlaces hipertextuales, foros, guías de enlaces)

Respecto a las propiedades podemos enumerar las siguientes:

1. Grado: mide las posibilidades de navegación e interactividad global que presenta un hiperdocumento.
 - a. Hipertexto Grado 1: simples conjuntos de secciones unidas en forma de red (secciones, enlaces anclajes), carente de

- herramientas de navegación y de una estructura lógica; provocando un desbordamiento cognitivo.
- b. Hipertexto Grado 2: G1 más una herramienta de navegación como un sumario o un índice, o ambos que permitan el acceso a la información.
 - c. Hipertexto Grado 3: G2 más un sistema de recuperación (Filtro, buscador).
 - d. Hipertexto Grado 4: G3 más alguna forma de interactividad, como mapas 3d, realidad virtual o cuadros estadísticos.
2. Estructura: los hiperdocumentos poseen una doble estructura que son:
- a. Física: representación lineal y secuencial de la información
 - b. Lógica: es la que se manifiesta ante el lector dada por la navegabilidad y representación de la información.
3. Conectividad, la conectividad posee dos dimensiones:
- a. La micronavegación: desplazamientos dentro del mismo hiperdocumento, pudiendo ser:
 - i. Mínima: Sumario Global
 - ii. Media: Mínima + Sumarios Locales
 - iii. Máxima: Media + Índices
 - b. La macronavegación: Relación de un Hiperdocumento con otros, sus dimensiones son:.
 - i. Luminosidad: citas, o enlaces que parten del hiperdocumento hacia otros.
 - ii. Visibilidad: citas o enlaces de otros documentos hacia el hiperdocumento considerado.
4. Autoría / Fuente -Clases de autoría / fuente en hiperdocumentos según la declaración incluida en ellos:

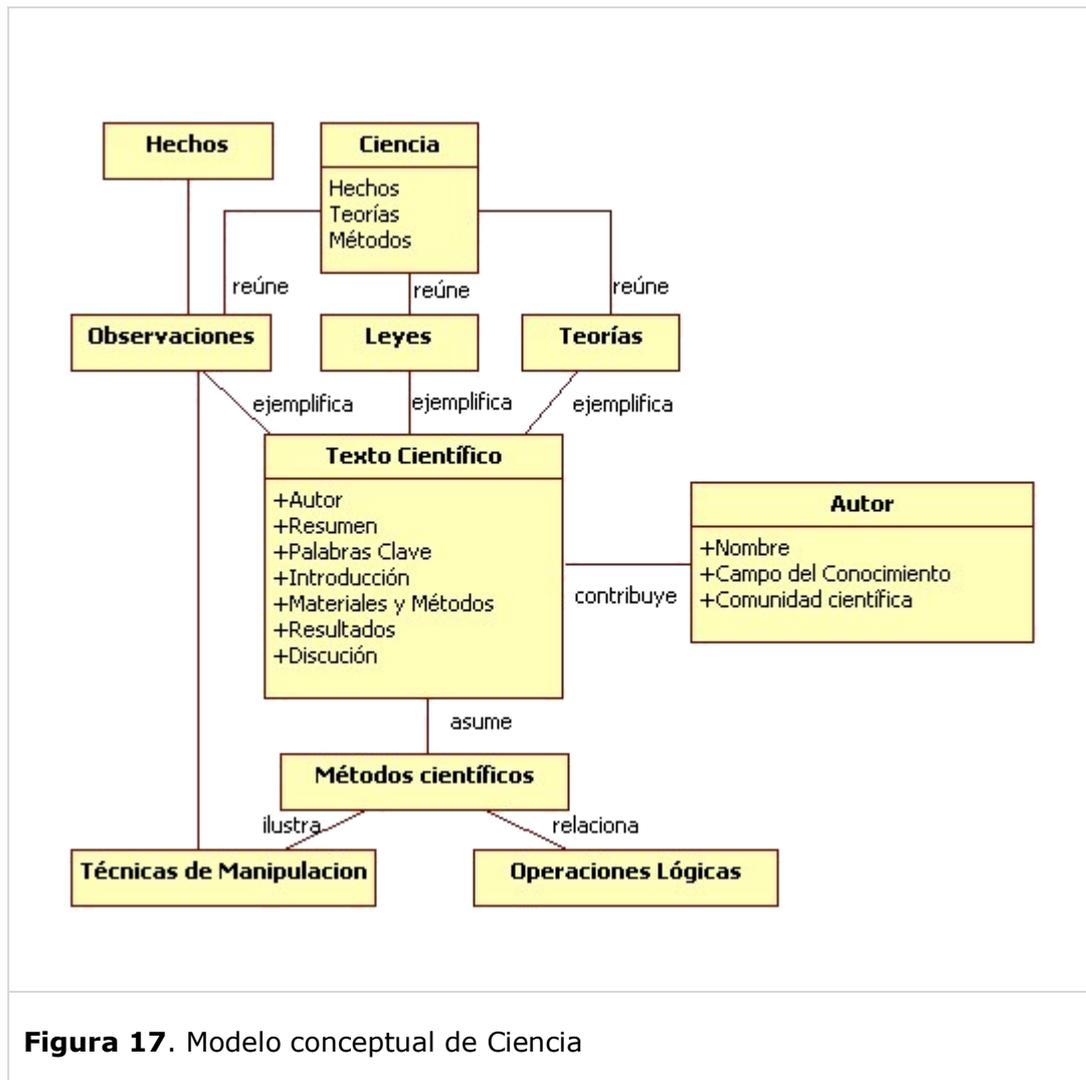
- a. Anónima: sin indicación de autoría o fuente, por descuido o por estrategia de ocultación de la información.
 - b. Implícita: simples indicaciones de nombres propios sin información adicional, curricular o de contexto.
 - c. Explícita: con indicaciones de autoría y editoriales.
5. Metainformación - Información autorreferencial; información sobre un documento incluida en el propio documento. Tipos habituales de metainformación en sedes Web:
- a. Metadatos con etiquetas estándares en HTML: elemento "meta" y los atributos name y content; como keyword, Description o Creator.
 - b. Metadatos con etiquetas Dublin Core

El hipertexto en la gestión de la información académica

Basados en la conceptualización de ciencia como: la constelación de hechos, teorías y métodos reunidos en los libros de texto actuales y en donde los científicos son hombres que, obteniendo o no buenos resultados, se han esforzado en contribuir con alguno que otro elemento a esa constelación particular.

Asumiendo dicha conceptualización, se procedió a elaborar un primer modelo conceptual, el cual se encuentra representado en la figura 17.

En él se describen a los recursos científicos como contenedores de las observaciones, leyes, teorías, métodos científicos, técnicas de manipulación utilizadas en la reunión de datos y las operaciones lógicas empleadas para relacionar esos datos con las generalizaciones teóricas del recurso en cuestión (PERISSÉ, 2008).



Otro modelo conceptual que se desarrolló y con el cual se trabajó para determinar la infraestructura necesaria que permita realizar una apropiada gestión del conocimiento producido por la Universidad, representado en la figura 18, cuenta con los siguientes sistemas:

- Una Base de Datos, conteniendo el currículum de los docentes universitarios que permite evaluar los progresos de los mismos y su integración al modelo de gestión del conocimiento a través de la producción bibliográfica y los reportes de su producción técnica.

- Una Biblioteca Digital, que contenga prioritariamente:
 - un catálogo de publicaciones científicas,
 - un acervo de disertaciones y tesis aprobadas por la universidad,
 - un contenido de publicaciones monográficas relevantes para la comunidad
- Un Directorio de publicaciones seriadas, editadas por las distintas unidades académicas.
- Un órgano para desarrollar ontologías, que permita el intercambio de información entre las distintas unidades ligadas a la ciencia, tecnología e información universitaria.

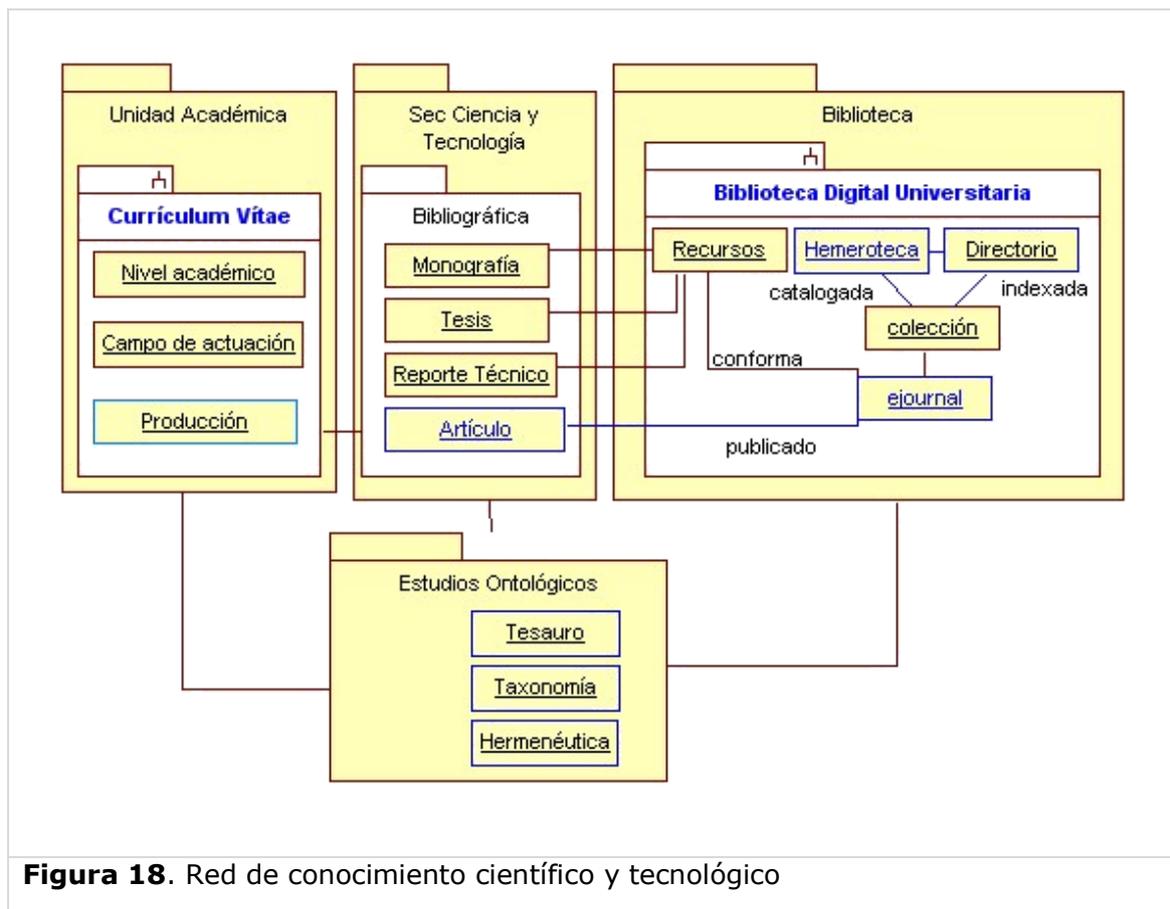


Figura 18. Red de conocimiento científico y tecnológico

A partir de dicho modelo conceptual se procedió a la construcción de los diagramas de clases para la construcción del:

- Catálogo de publicaciones seriadas científicas, que se incorporará a la Biblioteca Digital Universitaria, como se puede observar en la figura 19.

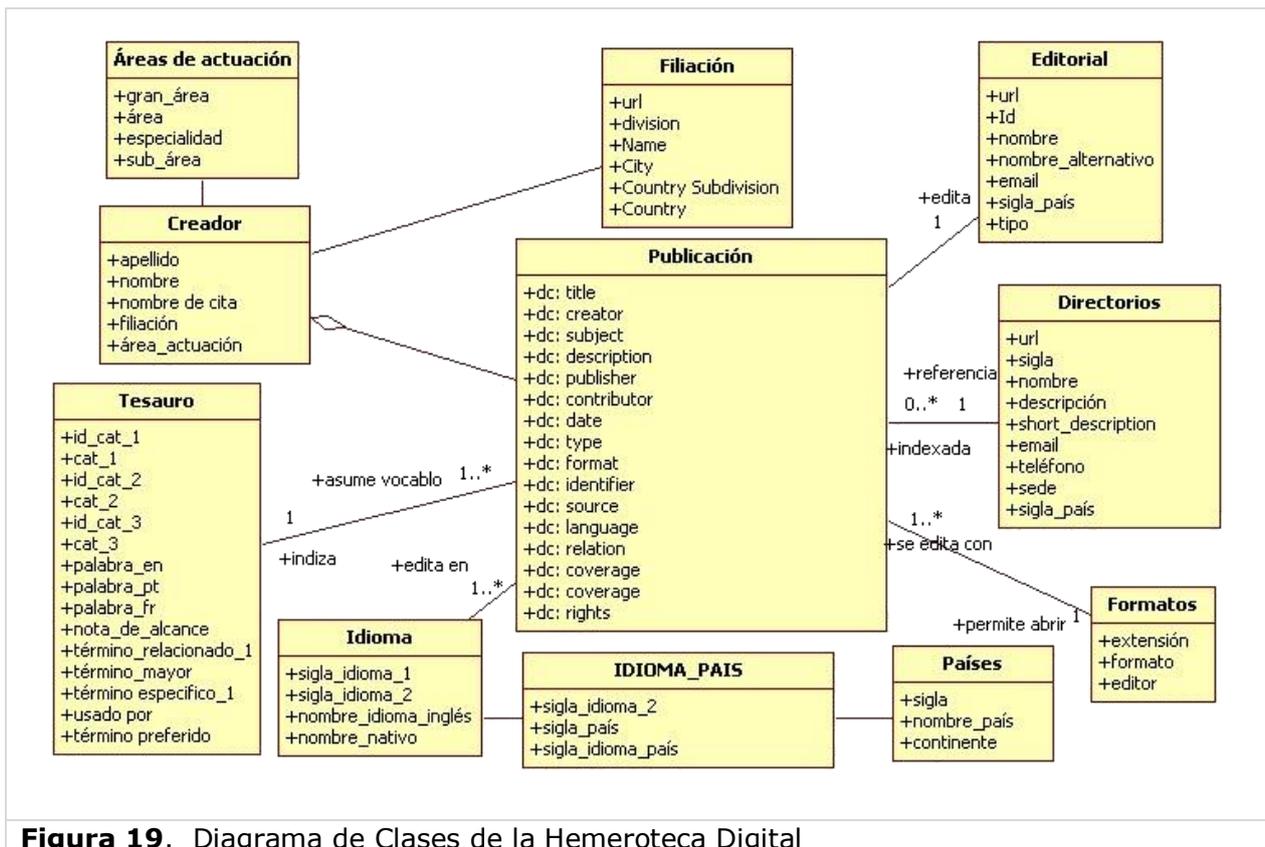


Figura 19. Diagrama de Clases de la Hemeroteca Digital

- Modelo de tesis, que permite la creación, el uso, la diseminación y preservación de las tesis digitales tal se representa en la figura 20.

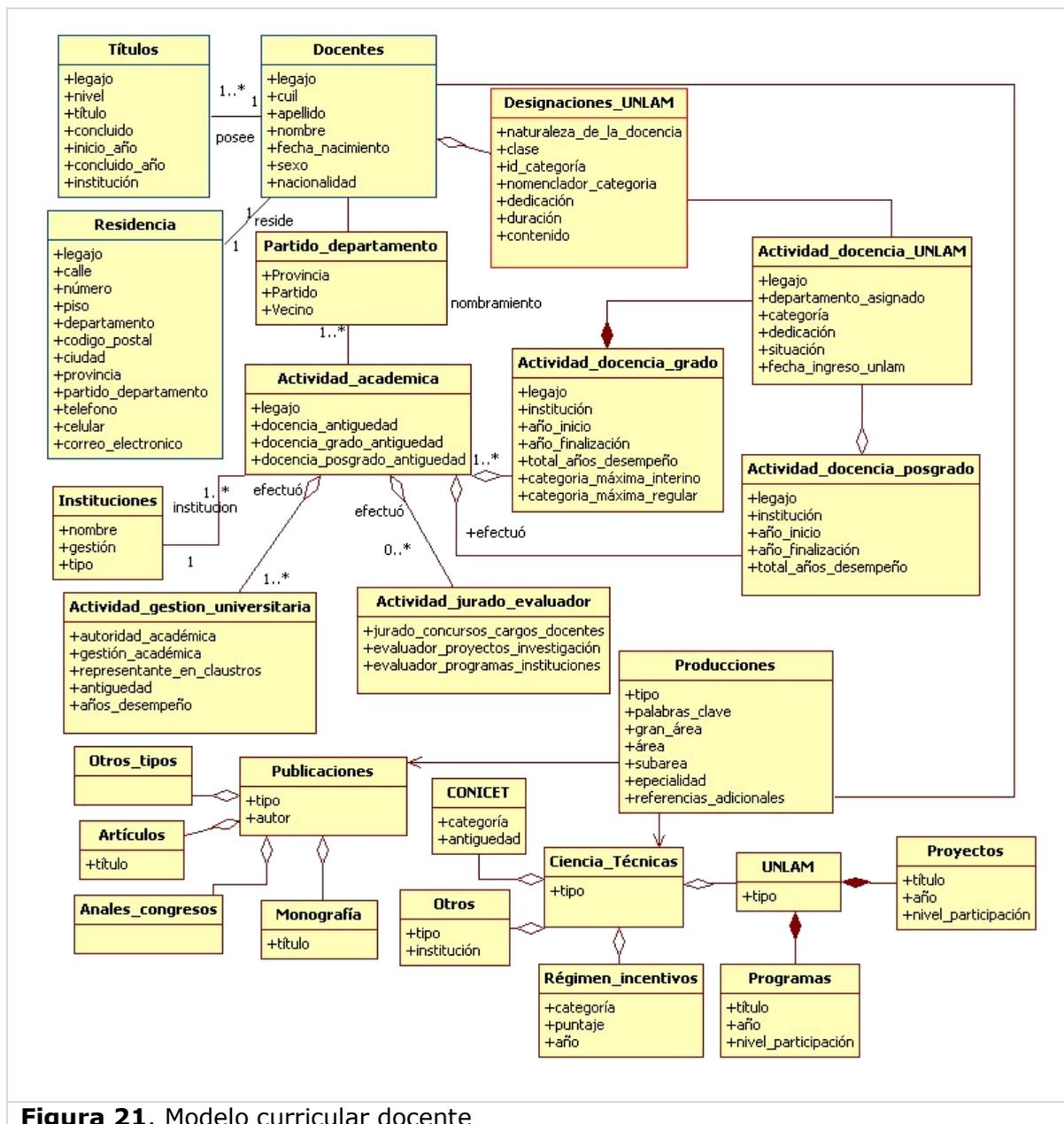


Figura 21. Modelo curricular docente

Los modelos se diseñaron bajo los siguientes requisitos del sistema, en el que el programa permitirá:

- La integración de los sistemas de archivos abiertos reconocidos por la comunidad científica.
- El mantenimiento de un archivo distribuido de contenidos científicos.

- La implementación del protocolo OAI para el intercambio de metadatos.
- La implementación de un recolector (harvester) para información tanto en español como en otras lenguas.
- La creación de un proveedor de metadatos destinado a mantener uno o más repositorios que soporten OAI, para que cada revista pueda poner a disposición de cualquier usuario y aplicación, los metadatos referentes a sus contenidos.
- la creación de un proveedor de servicios (service provider) que tome datos de los proveedores y genere un valor agregado.

La arquitectura lógica del programa está basada en el modelo Universal Preprint Service (UPS) cuyos niveles para la provisión de los datos, gestión de los datos y la prestación de servicio son:

- DP (data provider),
- HARVESTING (OAI-PMH),
- SP (service provider).

Modelo navegacional

Basados en la semántica de los nodos se han construido los enlaces de las aplicaciones hipermedia y las estructuras de acceso, tales como índices o recorridos guiados, como representación de los posibles caminos de acceso a los nodos.

Sobre este modelo representado en la figura 21, correspondiente al diagrama de clase del sistema de información de posgrado para las Tesis Digitales, se realizó el bosquejo del modelo de navegabilidad, en el cual se especifican los distintos tipos de objetos y principalmente la tesis digital y cómo serán presentados a los potenciales usuarios (actores), como los: docentes, alumnos, coordinadores y bibliotecarios.

Este modelo derivado del esquema conceptual, tiene por objetivo construir las diferentes visiones, considerando los distintos perfiles de los usuarios.

Estos usuarios con distintos perfiles podrán, ver la misma información con distintas perspectivas.

En esta actividad de proyectar la estructura de navegación, se consideraron los siguientes aspectos (ver figura 22):

- Cómo se relaciona la Tesis Digital con los distintos tipos de usuarios y con el resto de los objetos del sistema.
- Cuál es la estructura subyacente a la navegación y en qué contexto se irá a presentar ante los usuarios.

Así se espera contar con una representación de la aplicación a nivel navegacional, a través de un modelo de clases que especifique los objetos de información y del modelo de contexto, donde podrán ser especificados las distintas alternativas que los usuarios podrán seleccionar para explorar los distintos objetos de información.

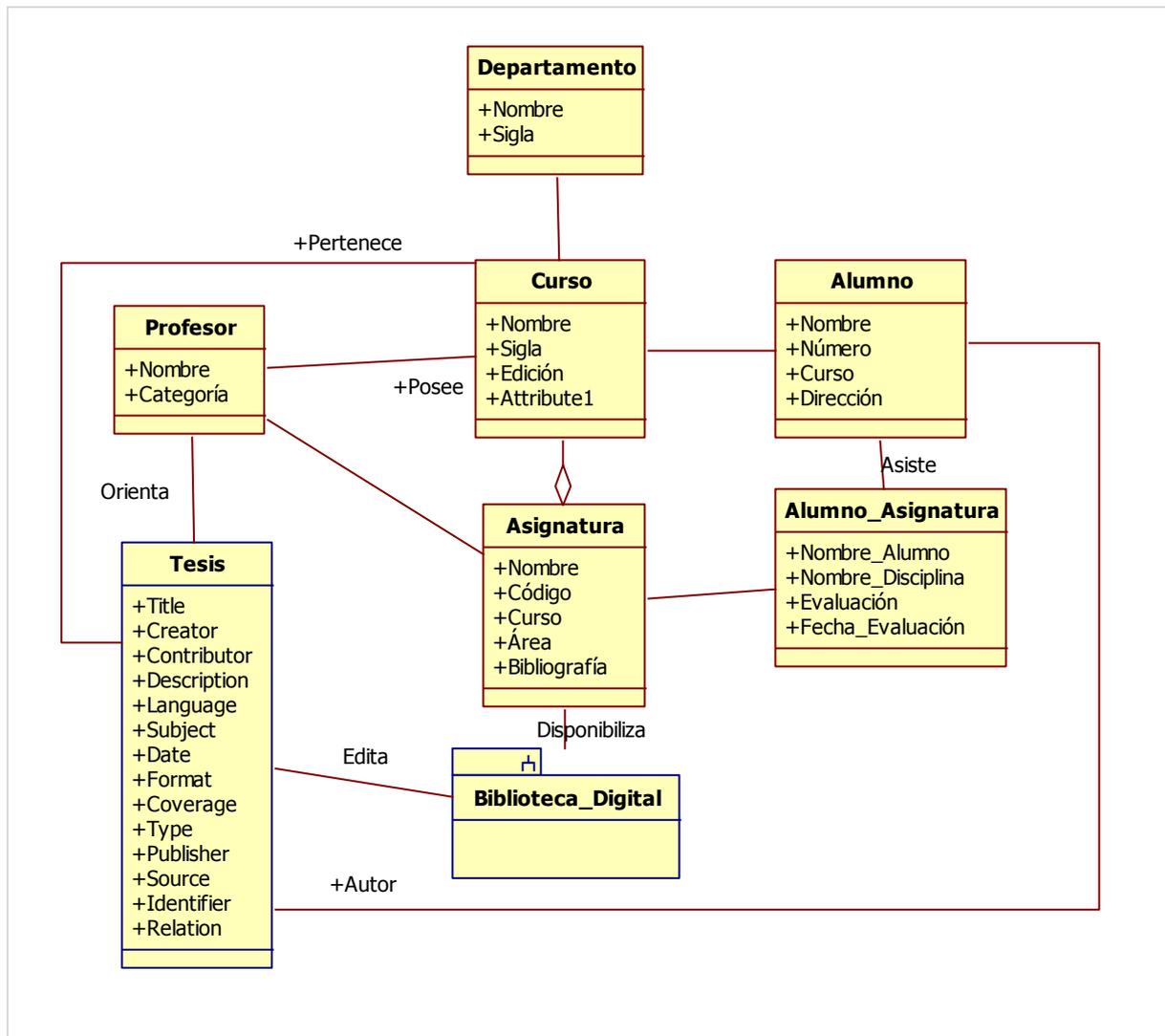


Figura 22. Esquema de clases de la Tesis Digital, con la visión de los actores profesor y coordinador

En la figura 23 se representa, a modo de ejemplo, el modelo navegacional correspondiente a la aplicación de tesis digitales.

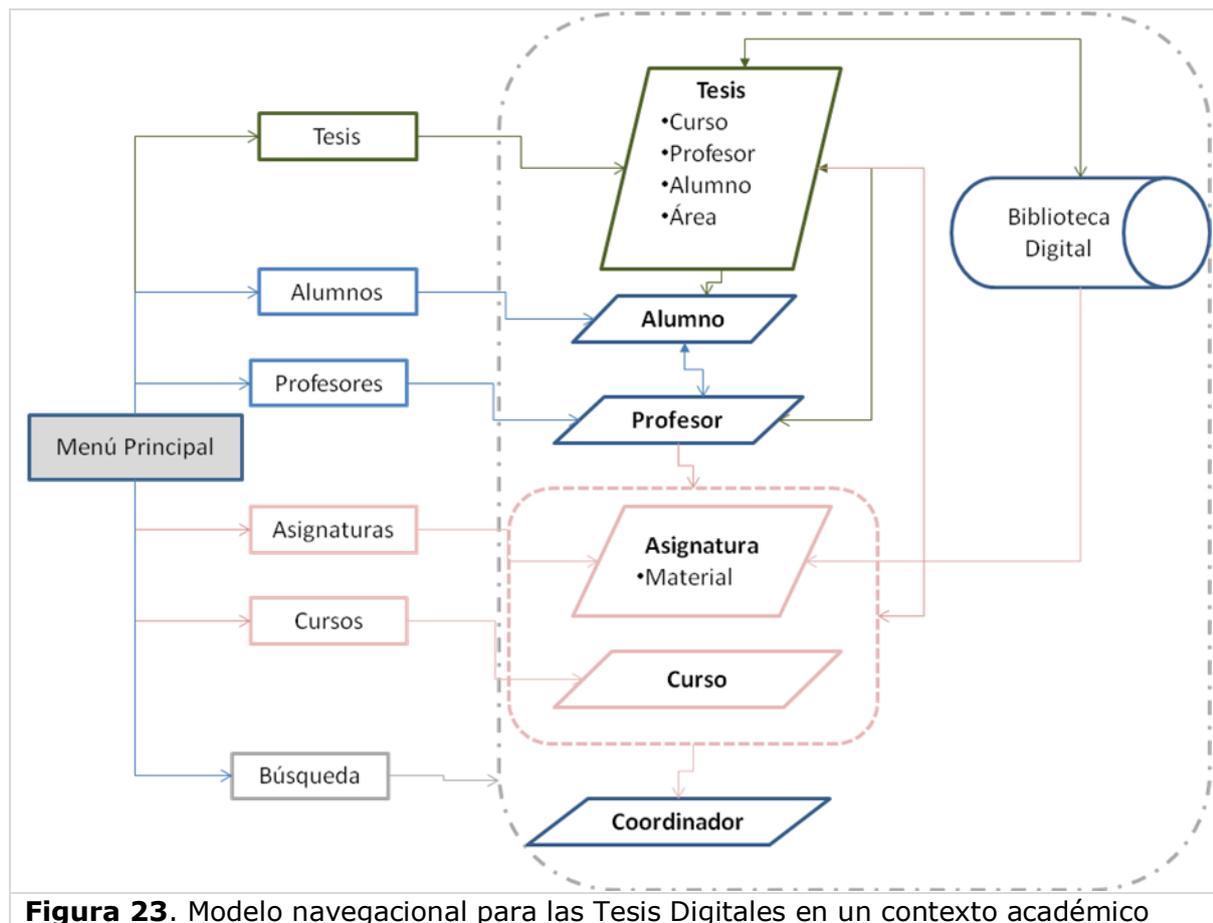


Figura 23. Modelo navegacional para las Tesis Digitales en un contexto académico

Este modelo derivado del esquema conceptual, tiene por objetivo construir las diferentes visiones, considerando los diferentes perfiles de los usuarios. Estos usuarios con distintos perfiles podrán ver la misma información con distintas perspectivas.

En esta actividad de proyectar la estructura de navegación se consideraron los siguientes aspectos:

- Cómo se relacionan los recursos digitales con los distintos tipos de usuarios y con el resto de los objetos del sistema.
- Cuál es la estructura subyacente a la navegación y en qué contexto se irá a presentar ante los usuarios.

Así se espera contar con una representación de la aplicación a nivel navegacional a través de un modelo de clases que especifique los objetos de información y del modelo de contexto donde podrán ser especificados las distintas alternativas que los usuarios podrán seleccionar para explorar los distintos objetos de información.

Modelo de interfaces abstractas

En este modelo se especifica la estructura de los objetos de la interface, especificando cómo se activa la navegación entre los objetos, las barras de herramientas y los botones de control; entre otras propiedades correspondientes a la interface.

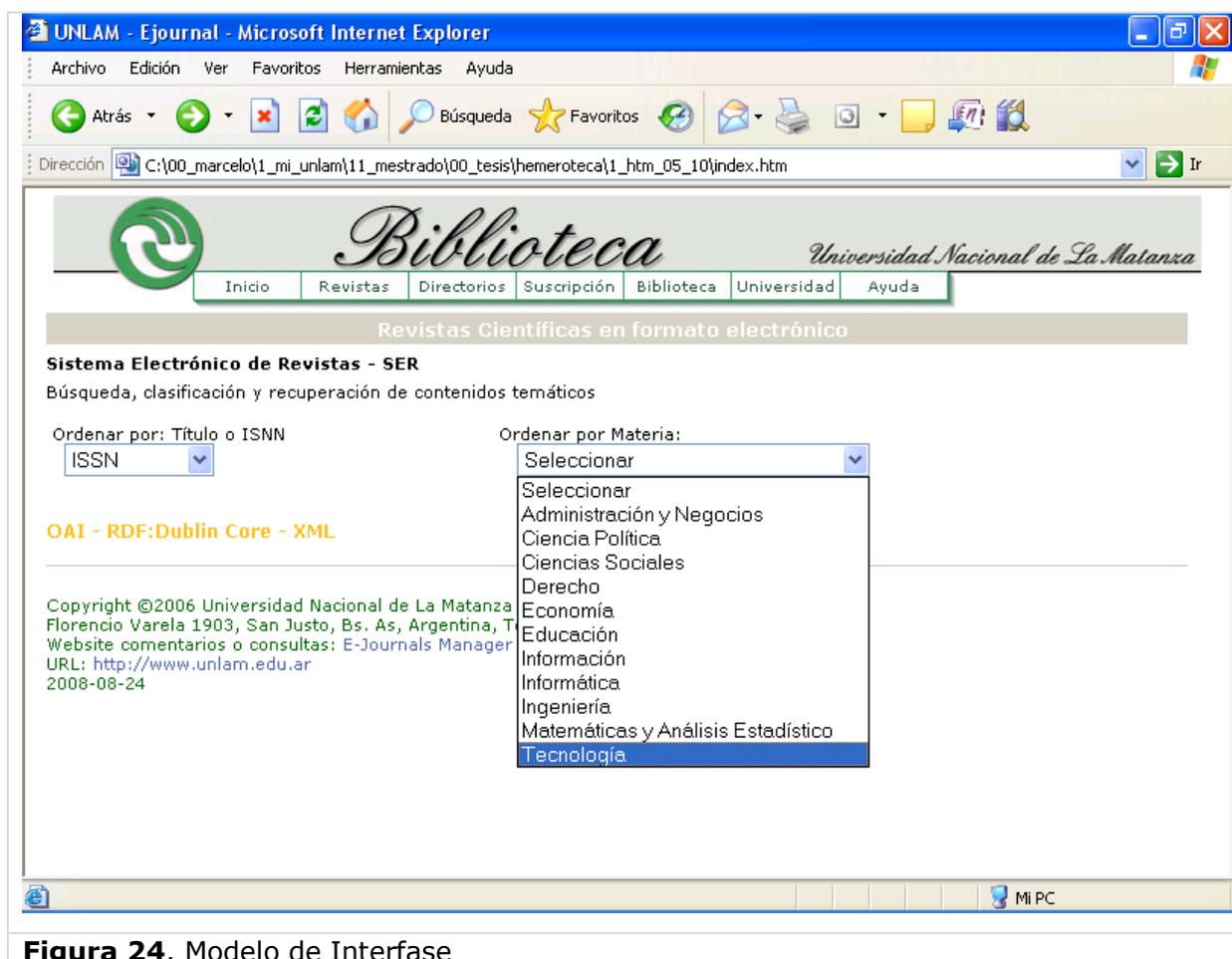


Figura 24. Modelo de Interfase

La metodología OOADM, seguida en el proyecto de la Web Semántica, utiliza Abstract Data View para especificar este modelo; en nuestro caso aún es materia de estudio.

Construcción de la Base de Datos

El conjunto unificado de información, resultante de nuestro proyecto informático, que será compartido por los diferentes usuarios, va a conformar la denominada Base de Datos.

Independientemente de la Base de Datos que será implementada, ésta necesita de un Sistema de Gestión de Base de Datos (SGBD). Los sistemas de Gestión de Base de datos, son programas de software para la administración de las Bases de Datos; y en particular, para: almacenar, manipular y recuperar datos en una computadora. El SGBD también se encargará de la comunicación entre el usuario y la base de datos, proporcionándole al usuario, los medios necesarios para poder obtener información, introducir nuevos datos y actualizar los ya existentes.

El modelo propuesto parte de la arquitectura clásica de bases de datos ANSI/SPARC, y considera tres capas: documentos, esquemas y ontología. Más adelante se muestra cómo este modelo se adapta a los paradigmas existentes para la integración de la información, y a la vez proporciona un marco conceptual para guiar nuevos desarrollos en ese campo. (Ismael Sanz, 2002)

Arquitectura de la Base de Datos

Para definir una arquitectura modelo, se han estudiado las soluciones que se están aplicando en los proyectos existentes de integración semántica de la información. En general, se han encontrado sistemas basados en

recubridores y mediadores, sistemas basados en esquemas federados, y soluciones ad hoc. Por otro lado, hay muchos trabajos enfocados a resolver problemas más puntuales, por ejemplo las investigaciones sobre técnicas para la alineación de ontologías.

Un requisito importante es que la arquitectura debe integrar de modo natural este trabajo ya realizado.

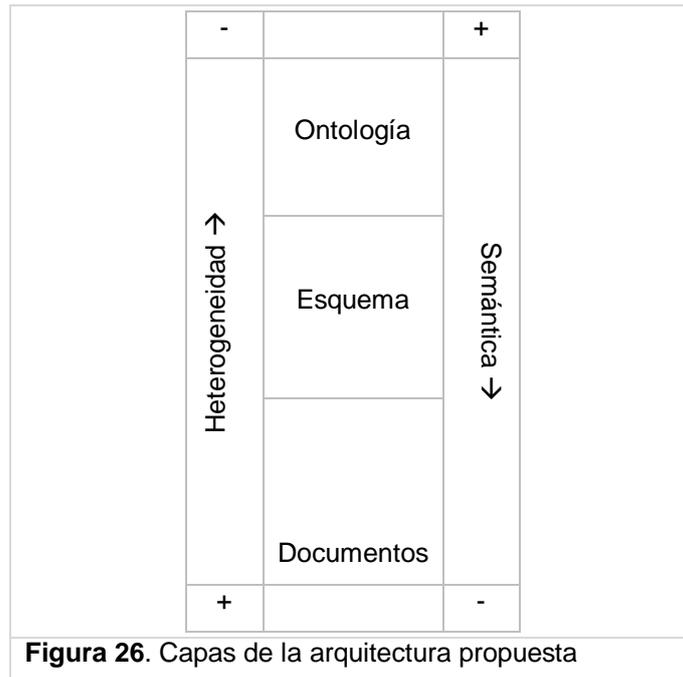
El otro requisito fundamental es que la arquitectura ayude a situar los problemas de investigación abiertos en el campo, y a calibrar su importancia.

Esto incluye desde la falta de estandarización de los formatos, a cuestiones más fundamentales que se presentan a la hora de realizar consultas sobre modelos muy expresivos.

También se pueden incluir problemas más prácticos, como la falta de datos más allá de simples documentos XML sin un DTD que los describa, y desde luego la falta de ontologías ya definidas.

A la vista de estos requerimientos, hemos empezado a utilizar un modelo basado en la arquitectura de referencia ANSI/SPARC de tres capas (esquema físico, esquema conceptual y vistas). Esta división proporciona un marco de referencia que muestra las necesidades de una sistema de integración, proporcionando un marco claro para estudiar soluciones al problema (PÉRISSÉ, 2001).

En nuestra versión adaptada, contamos con tres capas análogas a las de la arquitectura ANSI/SPARC (ver figura 26): documentos, esquema y ontología.



En donde la capa de documentos contiene los datos que se pretenden integrar, en su formato nativo. No es una simplificación excesiva considerar que se trata de documentos XML, ya que siempre se puede construir un recubridor que efectúe la conversión necesaria.

Mientras que la capa de esquema describe la estructura global de los documentos que componen la capa inferior. Se tratará de descripciones mediante DTD, o si es posible XML Schema o RDF.

Y por último la capa de ontología proporciona una visión semánticamente coherente de la información, mediante el uso de ontologías que describan el dominio del sistema. El usuario siempre interactúa con esta capa, que proporciona una visión simplificada y de alto nivel, ocultando la heterogeneidad del sistema subyacente.

En el marco de este trabajo, es importante resaltar la enorme semejanza de un conjunto dado de metadatos RDF con la tabla de una base de datos relacional. En primer lugar, podemos observar que la tabla 1 se puede transformar en una tabla de una base de datos relacional simplemente cambiando el título de las columnas.

Tabla 1. Parte de los metadatos de los recursos de la hemeroteca expresados como una tabla relacional

<i>Identifier</i>	<i>Title</i>	<i>Creator</i>	<i>Description</i>	<i>Subject</i>
http://cyta.com.ar	El trabajo de tesis: estilo	Wymarczyk, Hilario	Consejos de estilo y el manejo...	Tesis

En efecto, en el modelo relacional, cada entidad es una fila de una tabla, y cada columna es un atributo de la entidad. En la siguiente tabla podemos ver las equivalencias entre ambos modelos.

Tabla 2. Equivalencias entre modelo relacional y RDF

<i>Modelo relacional</i>	<i>Modelo RDF</i>	<i>Significado común</i>
Entidad	Recurso	Aquella parte del mundo real o conceptual sobre la cual se aporta alguna (meta) información
Atributo	Propiedad	Características relevantes de la entidad o del recurso
Valor	Valor	Dato concreto que asume el valor de una entidad o de un recurso

Posteriormente fue aplicado este modelo en una base de datos relacional, en nuestro caso una base de datos Microsoft Access, dando como resultado el diagrama de Entidad Relación que se observa en la Figura 26

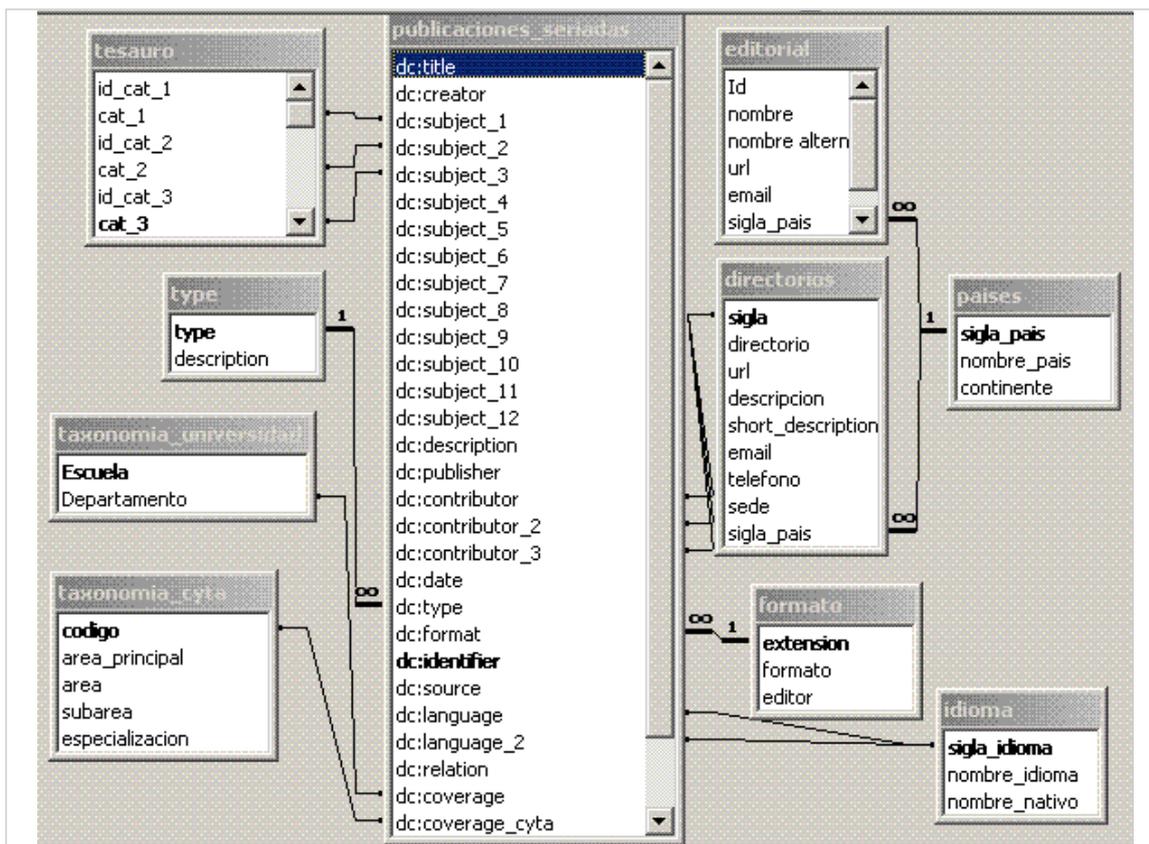


Figura 26. Hemeroteca digital - diagrama de relacional de la Base de Datos Access

Estructurando la información bajo el modelo Resource Description Framework (RDF) y Dublin Core en sintaxis XML, se obtiene un recurso del siguiente tipo.

```

<?xml version="1.0" encoding="UTF-8" ?>
-<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/metadata/dublin_core#"
  xmlns:dcq="http://purl.org/metadata/dublin_core_qualifiers#"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.cyta.com.ar/xml/rdf_dc.xsd">
-<rdf:description about="http://www.scielo.org.co/scielo.php?... ">
<dc:title>Cuadernos de Administración</dc:title>
<rdf:Bag>
<dc:subject_1>ADMINISTRACION</dc:subject_1>
...
<dc:subject_4>ADMINISTRACION</dc:subject_4>
</rdf:Bag>
<dc:description>Cadernos de Administração é uma publicação ...</dc:description>
<dc:contributor>DOAJ</dc:contributor>
<dc:type>científica</dc:type>
<dc:identifier>01203592</dc:identifier>
- <dc:source>http://www.scielo.org.co/scielo.php?...></dc:source>
<dc:coverage>ciencias economicas</dc:coverage>
</dc>
...

```

Figura 27. Modelo RDF:DC en sintaxis XML

Este primer modelo de gestión de recursos bibliográficos se ve complementado por el modelo curricular docente, en el que se vinculan principalmente por medio de la producción científica de los mismos (ver figura 28). Este proceso de integración permite realizar la relación entre los campos del conocimiento y las comunidades científicas.

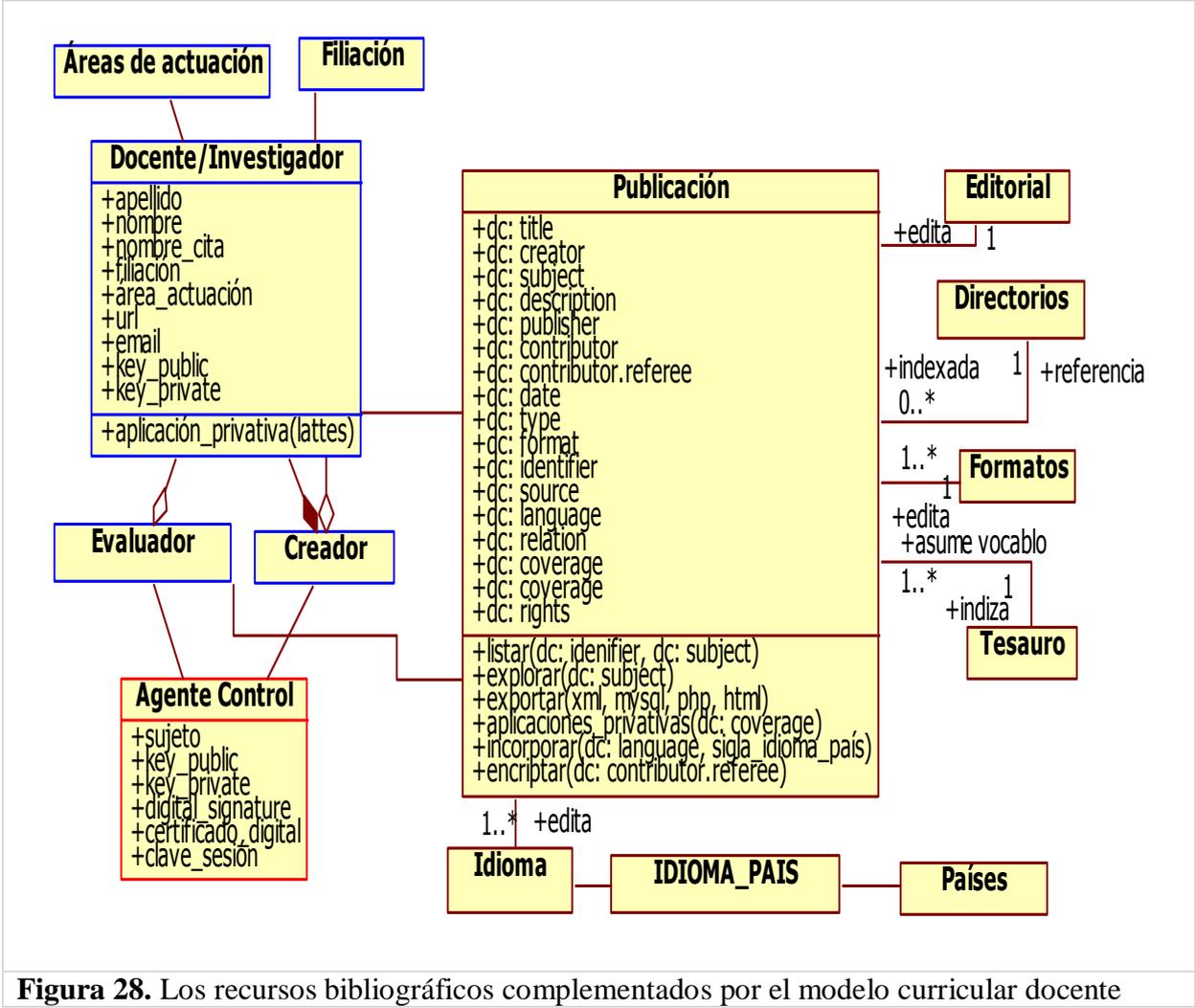


Figura 28. Los recursos bibliográficos complementados por el modelo curricular docente

En tanto a través de la figura 29 puede visualizarse parte de la estructura de la entidad docente, expresada en sintaxis XML.

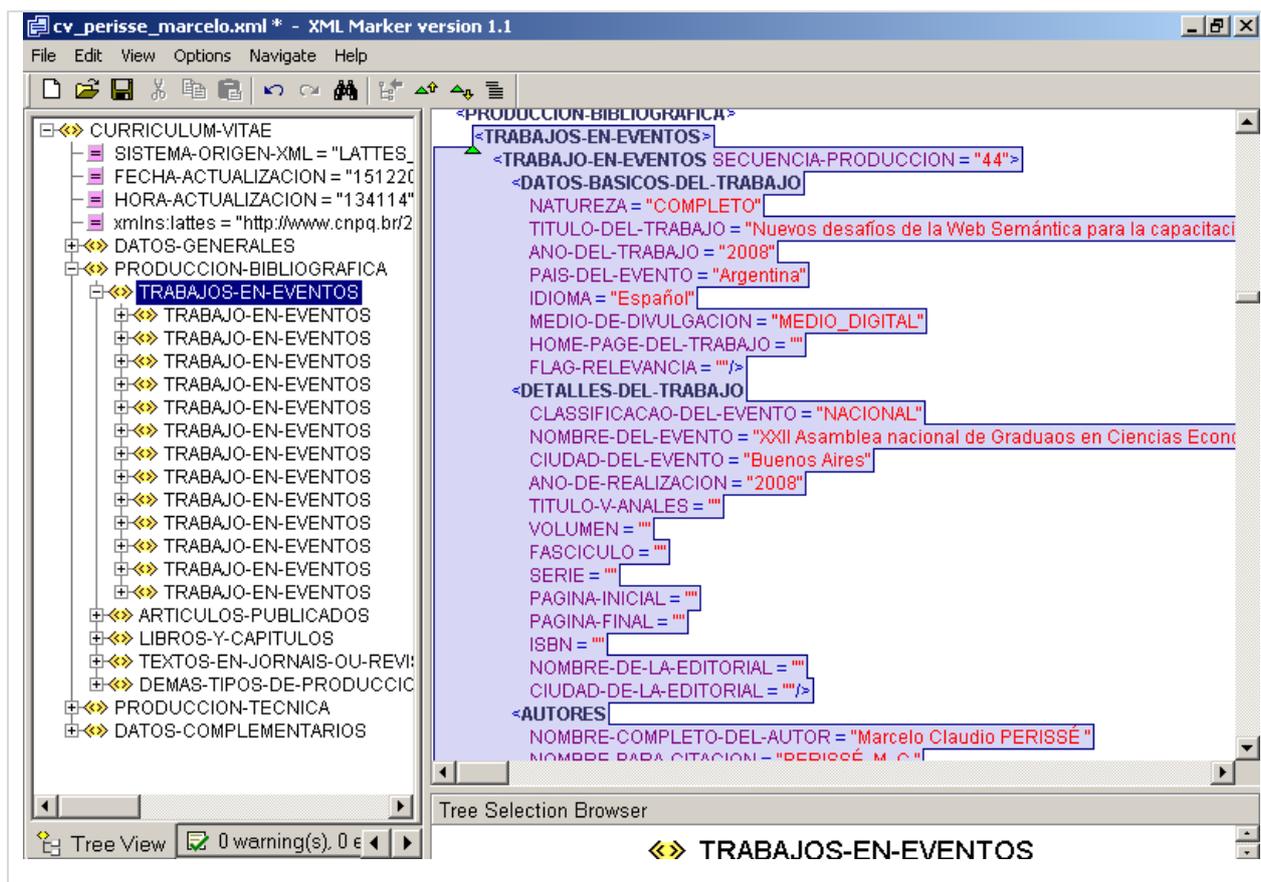


Figura 29. Vista parcial del archivo XML correspondiente a la entidad docente

CAPÍTULO V
CONCLUSIONES Y FUTUROS TRABAJOS

Capítulo V - Conclusiones y Futuros trabajos

El primer resultado obtenido es el de una plataforma que permite alojar revistas, reportes técnicos, artículos, tesis y monografías con cumplimiento de requisitos de calidad.

Al desarrollar un portal de acceso abierto (OAI), se crea un nuevo canal que permite la cooperación entre los diversos agentes de ciencia y tecnología.

Este portal, se convierte en un medio de comunicación apropiado para que los resultados de los proyectos de investigación y desarrollo lleguen al sector empresarial.

Además esta fuente de información y gestión del conocimiento se presenta, para dicho sector empresarial, como en un mecanismo más de contribución a la mejora de los procesos productivos; conformándose así en un servicio de utilidad pública, con ámbito de actuación multisectorial y nacional.

Adicionalmente, el sistema puede ser considerado como una herramienta útil para capacitar a los Docentes-Investigadores sobre la correcta estructuración de sus resultados de investigación como son: los reportes técnicos, los artículos, las tesis y las monografías, a ser alojados en la Web Semántica.

Con referencia a las técnicas simplificadas aplicadas, cabe aclarar que el término simplicidad de ninguna manera puede confundirse o asociarse al de sencillo, tosco o burdo; una técnica simplificada significa que cuando un proyecto informático se cumple satisfactoriamente y para ello se han empleado en él: menor cantidad de personal, menor cantidad de recursos técnicos y además se ha realizado en un tiempo menor, resulta pues un perfeccionamiento evidente del acto informático.

Puede decirse entonces que en informática aplicada, simplificación y perfeccionamiento son la misma cosa, con la inteligencia necesaria para cumplimentar el requisito fundamental de todo proyecto que es la solución de un problema, o al menos el mejoramiento de una situación dada.

En este trabajo se presentó una metodología *simplificada* para la construcción de un *Sistema de Gestión del Conocimiento* basado en la Web Semántica, haciendo hincapié en la capacidad de generar actividades de almacenamiento y recuperación de información apropiada.

En otras palabras, llegar al final por el camino más corto y sencillo significa, indudablemente, un progreso; y cuando ante cualquier procedimiento se deja en el ánimo del observador la impresión de que el mismo ha sido fácil, es sin duda alguna, porque se ha efectuado con: pulcritud, sencillez, tiempos bien reglados y con prescindencia de procedimientos redundantes y recursos inútiles; vale decir, se ha efectuado con una buena técnica.

Partiendo de estos razonamientos y hechos de observación, los conceptos de simplificación y orden riguroso, son inversos a los equipos de trabajo numerosos y al instrumentismo exagerado. El secreto estará en saber disponer de las herramientas necesarias, seleccionarlas según sea el problema a ser abordado, y desarrollar un buen plan para su uso.

Destaco tanto el aspecto de las técnicas simplificadas, pues pienso en las distintas alternativas que vivenciará un profesional; pues si al profesional le tocara desarrollar su actividad en un medio mal provisto, sentirá un gran alivio de poder hacerlo sin mayores recursos de personal y herramientas. Si, por el contrario, tuviera que desempeñar su profesión en un medio mejor dotado, tendrá tiempo para rodearse de comodidades; que no siempre redundarán en beneficio del usuario.

El trabajo también ha buscado dar respuesta a cierto desconocimiento general sobre la *Tecnología de la Información*, que lleva a las personas a tener cierto *pensamiento mágico* sobre la utilización de la computadora en el proceso de gestión de la información, por la que pasan a considerar a esta herramienta como un elemento que prácticamente puede hacer absolutamente todo por nosotros.

Dentro de este *pensamiento mágico* se cree que, por medio de la comunicación en Internet, se pueden realizar sin método alguno, innumerables o vastas búsquedas que permiten encontrar y recopilar infinitos datos de numerosas fuentes y toda la información técnica y científica existente.

Como respuesta a ello el trabajo presenta a la computadora como una herramienta que puede manejar datos, procesarlos a fin de obtener información, almacenar datos e informes que facilitan la adquisición del conocimiento y por último automatizar reglas que nos llevan a una mejor comprensión de los hechos; basando la evaluación de la eficiencia en una lógica que no presenta inconvenientes de ser programada mediante diversos sistemas de inteligencia basados en reglas del tipo tripleta: sujeto, predicado, objeto, el que se adecua al esquema del modelo Resource Description Framework (RDF) .

En la categorización de sitios web, se utiliza la taxonomía basada en metadatos aplicando el modelo Dublin Core para saber qué elementos son representados, y la sintaxis del XML para saber cómo se representan dichos elementos. Como lazo de unión hacia el aspecto pedagógico, las ontologías permiten un desarrollo semántico de gran profundidad que proporcionan una descripción lógica y formal de la información que almacenan, y por lo tanto puede ser interpretada tanto por usuarios humanos como por programas informáticos. Para ello se utiliza el Resource Description Framework como herramienta en el armado de este

esqueleto semántico que sustenta a los tesauros en la representación y recuperación de información, con énfasis en la coherencia y riqueza relacional de la estructuración conceptual.

En lo que respecta al proceso Enseñanza-Aprendizaje se considera que la correcta utilización de la Tecnología de la Información puede ayudar al proceso educativo a reunir y relacionar los datos, transformarlos en información y almacenar el conocimiento necesario, que le permita al educando descubrir y generar la comprensión y el desarrollo de la sabiduría.

Es decir que constituyen un aporte para el mejoramiento del desempeño del sistema educativo. Para ello la retroalimentación entre aulas y bibliotecas digitales son la base en la que se sustenta todo este camino de aprendizaje.

Es aquí donde se precisa que todo proceso de Enseñanza-Aprendizaje en un entorno digital cuente con los siguientes elementos:

- a) Un tipo de lenguaje documental que represente la estructuración conceptual de un determinado campo del conocimiento (Tesauro).
- b) Un tipo de vocabulario controlado en que todos los términos están conectados mediante algún modelo estructural (jerárquico, arbóreo, facetado) y especialmente orientado a los sistemas de navegación, organización y búsqueda de contenidos de los sitios web (Taxonomía).
- c) Una representación explícita y formal de una conceptualización compartida, que implica una perspectiva sobre cierta realidad y que se constituye en la estructura conceptual de una base de conocimiento (Ontología).

- d) Un sistema de metadatos para la recuperación de información, que adopte al XML como formato de codificación y al Dublin Core como modelo para la descripción de los recursos de información.
- e) Una base para procesar metadatos que proporcione interoperabilidad entre aplicaciones que intercambian información en la Web, como es el modelo Resource Description Framework (RDF).

Es de esperar, que la instrumentación del presente modelo permita que el lenguaje disciplinar deje de ser una barrera difícil de transponer; favoreciendo así las actividades de investigación pluridisciplinarias, interdisciplinarias en sus tres grados: de aplicación, epistemológico y de generación de nuevas disciplinas; y fundamentalmente las actividades de investigación *transdisciplinarias*.

En este proyecto de Investigación y Desarrollo sobre la gestión del conocimiento quedan algunos sistemas relevantes a ser integrados como son:

- Aumentar la fidelización de los usuarios y optimizar la comunicación y difusión científica y tecnológica a todos los sectores de la sociedad.
- Establecer o implantar un sistema de calidad normalizado que permita crear un marchio de calidad consensuado por la comunidad científica.
- Proveer sistemas de valor añadido: Sistema de Alerta (RSS), acceso personalizado, estadística de uso, estudio de medidas de impacto (bibliometría).
- Proporcionar servicios de hospedaje para la producción científica de las distintas unidades académicas.
- Implementar un sistema de análisis bibliométrico para la Información Técnica y Científica, conceptualizando a la bibliometría como un método que constituye la aproximación cuantitativa que

permite el desarrollo el estudio descriptivo e inferencial o inductivo de todas las formas de comunicación adoptadas por la literatura científica.

Es de esperar, que la instrumentación del presente modelo permita que el lenguaje disciplinar deje de ser una barrera difícil de transponer; favoreciendo así las actividades de investigación pluridisciplinarias, interdisciplinarias en sus tres grados: de aplicación, epistemológico y de generación de nuevas disciplinas; y fundamentalmente las actividades de investigación transdisciplinarias.

Los proyectos que integran el presente programa han demostrado haber avanzado tanto en la acumulación de conocimiento como en la valoración de los objetivos inicialmente propuestos, viéndose plasmados en las aplicaciones desarrolladas, logrando mantener una hegemonía en el manejo de la sintaxis, la semántica aplicadas y en el uso de herramientas como: el XML, XSLT, XML Schema, el Lenguaje de Bases de Datos Relacional Structured Query Language (SQL), el sistema de gestión de bases de datos SQL Open Source MySQL, el lenguaje de código abierto PHP, el modelo semántico RDF con su respectivo esquema RDF Schema, el DC y los modelos de Firma Digital aplicados a documentos XML.

CAPÍTULO VI
BIBLIOGRAFÍA

Capítulo VI – Bibliografía

ACKOFF, Rusell. *Fábulas antiburocráticas de Ackoff*. Barcelona: Garneca, 1993

ARANO, Silvia. "Los Tesoros y las Ontologías en la Biblioteconomía y la Documentación". *Hipertext.net* [en línea]. 2005, núm. 3. Disponible en Internet: <<http://www.hipertext.net>>. [Consulta: 05/09/2005]. ISSN 1695-5498

BARCELÓ VALENZUELA, Mario, SÁNCHEZ SCHMITZ, Guzmán Gerardo Alfonso y PÉREZ SOLTERO, Alonso. "La Web Semántica como apoyo a la Gestión del Conocimiento y al Modelado Organizacional". *Revista Ingeniería Informática*. abril 2006, Edición 12.

BARRUECO, José Manuel y COLL IMMA, Subirats. OAI-PMH: Protocolo para la transmisión de contenidos en Internet-

BERNERS-LEE, Tim; HENDLER, James; LASILA, Ora. "La red semántica", *Investigación y Ciencia*. 2001, Jul.

BROWN, Richard D. Digital Signatures for XML. Disponible en Internet: <http://www.w3.org/Signature/Drafts/xmlldsig-signature-990618.html>-

BRUN E., Ricardo. "XML y la gestión de contenidos". *Hipertext.net* [en línea]. , 2005, núm 3. Disponible en Internet: <<http://www.hipertext.net>>. [Consulta: 05/09/2005]. ISSN 1695-5498

CENTELLES, Miquel. "Taxonomías para la categorización y la organización de la información en sitios web". *Hipertext.net* [en línea]. 2005, núm 3. Disponible en Internet: <<http://www.hipertext.net>> [Consulta: 05/09/2005]. ISSN 1695-5498

CODINA, L. *Información documental e información digital*. In YAPES L., José. *Manual de Ciencias de la Documentación*. Madrid: Pirámide, 2002. p.301-315

CODINA, L. y ROVIRA, C. "Recursos sobre la Web Semántica". *Revista Española de Documentación Científica*, Abril-Junio 2006, v. 29, n. 2, p. 297-305

CONSORCIO WEB (W3C). Resource Description Framework (RDF) marzo 2000. Disponible en Internet:

<http://www.sidar.org/recur/desdi/traduc/es/rdf/rdfesp.htm>-

CONSORCIO WEB (W3C). Resource Description Framework (RDF) Especificación del Esquema 1.0. marzo de 2000. Disponible en <Internet:<http://www.w3.org/TR/2000/CR-rdf-schema-20000327>>-

DAVID, Berlo, *The Process of Communication*. New York: Holt, Reinhart and Winston, 1960. pag, 72

DAVID, Harel. "On Visual Formalisms". *Communications of the ACM*, May 1988, Volume 31 Number 5, Pag. 514.

DUVAL, E., HODGINS, W., SUTTON, S.y WEIBEL, S. "Metadata principles and practicalities". *D-Lib magazine* [en línea]. 2002, april. Disponible en Internet: <<http://www.dlib.org/dlib/april02/weibel/04weibel.html>>

DANGER MERCADERES, Roxana María. "Extracción y análisis de información desde la perspectiva de la Web Semántica". Universidad de Jaume, Castellón – España, febrero, 2007. Pag, 38 y 39-

DÍAZ ORTUÑO, Pedro Manuel. Problemática y tendencias en la arquitectura de metadatos web. *Anales de documentación*, N° 6, 2003, PÁGS. 35-58-

ETSI, Security Plugtests Event – PKI/XAdES, 29 de junio de 2004.

Disponibile en Internet: <http://www.etsi.org>-

_____, XML Advanced Electronic Signatures (XAdES), abril de 2004.

Disponibile en Internet: <<http://www.etsi.org>>. Acesso em: 12 dez. 2004-

_____, ETSI XAdES PlugTests – Final Report, novembro de 2003.

Disponibile en Internet: <<http://www.etsi.org>>. Acesso em: 12 dez. 2004-

_____, Signature policy for extended business model, março de 2003.

Disponibile en Internet: <http://www.etsi.org>-

_____, XML format for signature policy, abril de 2002. Disponibile en

Internet: <http://www.etsi.org>-

FERREIRA, Carlos; TEIXEIRA, Leonor y RUI, Santiago. Modelização de Aplicações Web: o contributo da UML e do OOHDM na modelização de um sistema de informação académico-

Internet Engineering Task Force IETF. *Internet X.509 Public Key Infrastructure Time-Stamp Protocol (TSP)*. RFC-3161: agosto de 2001.

Disponibile en Internet: <<http://www.faqs.org/rfcs/rfc3161.html>>

Internet Engineering Task Force IETF. *Cryptographic Message Syntax*.

RFC-3852: julio de 2004. Disponibile en Internet:

<<http://www.faqs.org/rfcs/rfc3852.html>>

Internet Engineering Task Force IETF. *Cryptographic Message Syntax*.

RFC-3369: agosto de 2002. Disponibile en Internet:

<<http://www.faqs.org/rfcs/rfc3369.html>>

Internet Engineering Task Force IETF. *An Internet Attribute Certificate Profile for Authorization*. RFC-3281: abril de 2002. Disponibile en Internet:

<<http://www.faqs.org/rfcs/rfc3281.html>>

Internet Engineering Task Force IETF. *XML-Signature Syntax and Processing* RFC-3275: marzo de 2002. Disponible en Internet:
<<http://www.faqs.org/rfcs/rfc3275.html>>

Internet Engineering Task Force IETF. *Cryptographic Message Syntax* RFC-2360: junio de 1999. Disponible en Internet:
<http://www.faqs.org/rfcs/rfc2630.html>

Ismael Sanz, Juan Manuel Pérez, Rafael Berlanga. *Hacia una arquitectura de referencia para la integración semántica de información*. En REDBD'2002: Taller sobre Integración Semántica de Fuentes de Datos Distribuidas y Heterogéneas. El Escorial (Madrid), 2002

ISO 17799:2000. *Information technology. Code of practice for information security management*.

KUHN, Thomas. *La estructura de las Revoluciones científicas*. Buenos Aires: FCE, 1990.

LAMARCA LAPUENTE, María Jesús. "Hipertexto, el nuevo concepto de documento en la cultura de la imagen". Tesis doctoral. Universidad Complutense de Madrid. Disponible en Internet:
<<http://www.hipertexto.info>>. Fecha de Actualización: 30/05/2006-

LÓPEZ GUZMÁN, Clara. Modelo para el Desarrollo de Bibliotecas Digitales Especializadas. Disponible en Internet:
<http://www.bibliodgsca.unam.mx/tesis/tes7cllg/tes7cllg.htm>-

LÓPEZ LÓPEZ, Pedro. *Introducción a la bibliometría*. Valencia: Promolibro, 1996.

LORSCH, Jay W. y SHELDON, Alan. "The individual in the organization: A system View". *Managing Group and Intergroup Relations*. Ed. Jay W.

Lorsch y Poul R. Lawrence, Homewood, Ill: Irwin-Dorsey, 1972, pp. 161-82.

MARTÍNEZ, Rubén. Servicio piloto de certificación en RedIRIS. Boletín de RedIRIS, número 39, autenticación, confidencialidad, piloto de certificación, criptografía, criptosistemas, IRIS-PCA, X.509, protocolos, claves PEM. RedIRIS. España. 2007-11-22-

Ministerio De Educación, Ciencia y Tecnología Secretaría De Ciencia, Tecnología e Innovación Productiva. Bases para un Plan Estratégico de Mediano Plazo en Ciencia, Tecnología e Innovación. Buenos Aires, julio de 2005-

Ministerio de Educación Ciencia y Tecnología. Resolución 193 SPU y 854 SCTIP, 9 de junio de 2005-

Ministerio de Educación Ciencia y Tecnología. Resolución 1054 SPU y 854 SCTIP, 12 de julio de 2005-

MySQL 5.0 Reference, 1997-2007 MySQL AB Manual <http://dev.mysql.com> Documento generado el: 2007-05-15 (revisión: 436)-

NICOLESCU, Basarab. *O Manifesto da transdisciplinaridade*. Lúcia Pereira de Souza (trad.) São Paulo: Triom, 1999.

NAKOV, Svetlin. How Digital Signatures Work: Digitally Signing Messages. Disponible en Internet: <http://www.developer.com/java/ent/article.php/3092771->

NOVAK, J y GOWIN, D. *Learning How to Learn*. Cambridge: Cambridge University Press, 1984.

PEMAU ALONSO, Julio y BARROSO CORROTO, Jose Angel. Introducción a OAI-PMH y su implantación en el portal e-Revistas. Disponible en

Internet:http://www.emwis.org/documents/pdf/20050404_JA_cursos_erevistas_completo.pdf

PÉRISSÉ, Marcelo Claudio. *Proyecto informático: una metodología simplificada*. Buenos Aires: Ciencia y Técnica Administrativa, 2001. ISBN 987-43-2947-5

PÉRISSÉ, Marcelo Claudio. "Semantic Web in higher education". *Journal of Information Systems and Technology Management*. 2008, Vol. 5, N.2/ Mai.Ago./08.

PÉRISSÉ, Marcelo Claudio. "Nuevos desafíos de la Web Semántica para la capacitación superior". En: Actas de la X Conferencia de Ciencias Empresariales y Sociales (Cuba, Junio 2008). ISBN 978-959-16-0737-9

PÉRISSÉ, Marcelo Claudio. *Gestión del conocimiento en una Web Semántica segura para la educación superior*

PÉRISSÉ, Marcelo Claudio. "La computadora en la capacitación". En: Actas de la Primera Jornada De Bibliotecas Universitarias. (Buenos Aires, 2003)

PÉRISSÉ, Marcelo Claudio y PEPE, M. "Infraestructura para la descripción de recursos electrónicos en la capacitación". En: Actas de la 4ta. Jornada de Bibliotecas Digitales Universitarias, (Mendoza 2006)

PÉRISSÉ, Marcelo Claudio y PEPE, M. "Infraestructura para la descripción de recursos electrónicos en la capacitación: marco tecnológico para la integración de los actores del proceso de Enseñanza-Aprendizaje". En: Actas del 3rd International Conference on Information Systems and Technology Management, (São Paulo, May 31st, 2006). P 549-560

POZO, Juan R. *Espacio de Nombre en XML*. World Wide Web Consortium, Massachusetts Institute of Technology, Institut National de Recherche en

Informatique et en Automatique, Keio University. Disponible en Internet:
<<http://html.conclase.net/w3c/xml-names-es>>

RICH, Elaine. *Inteligencia Artificial*, 2ª ed. México: McGraw-Hill, 1994. P.4

ROVIRA, Cristòfol; Marcos, Mari-Carmen; Codina, Lluís. "Repositorios de publicaciones digitales de libre acceso en Europa: análisis y valoración de la accesibilidad, posicionamiento web y calidad del código". *El profesional de la información*. 2007, enero-febrero, v. 16, n. 1, pp. 24-38.

RUMBAUGH, James; JACOBSON, Ivar y BOOCH, Grady. *El lenguaje Unificado de Modelado: manual de referencia*. Madrid: Pearson Educación S.A., 2000.

SENSO RUIZ, José A. "Resource Description Framework". *Hipertext.net* [en línea]. 2003, núm 1. Disponible en Internet:
<<http://www.hipertext.net>> [Consulta: 07/06/2006]. ISSN 1695-5498

SOUSA, Artur Afonso. *Base de Datos, Web e XML*. Lisboa: FCA Editora de Informática, 2002.

TURNER, Arthur y LOMBARD, Geororge. *Interpersonal Behavior and Administration*. New York: The Free Press, 1939, p. 12.

WINSTON, Patrick. *Inteligencia Artificial*, 3ª ed. Addison-Wsley Iberoamericana, 1999. Pag.299.

W3C. *Extensible Markup Language (XML) 1.0*, 04 de febrero de 2004.
Disponible en Internet: <http://www.w3.org/TR/2004/REC-xml-20040204/>

W3C. *Markup Language (XML) 1.0*. Traducción al español: Fabio Arciniegas A. Disponible en Internet: <<http://www.w3.org/TR/1998/REC-xml-19980210>>

W3C. *Canonical XML*. Disponible en Internet:

<<http://www.w3.org/1999/07/WD-xml-c14n-19990729>>

W3C. *Directrices de Accesibilidad XML*. Disponible en Internet:

<http://www.w3.org/TR/2001/WD-xmlgl-20010829>

W3C. *Esquema XML Parte 0: Fundamentos*. 2001. Disponible en Internet:

<<http://www.w3c.es/Traducciones/es/TR/2001/REC-xmlschema-0-20010502/>>

W3C. *Extensible Stylesheet Language (XSL)*. Version 1.1. 2003.

Disponible en Internet: <<http://www.w3.org/TR/2003/WD-xsl11-20031217/>>

W3C. *Guía Breve de Privacidad y P3P*. Disponible en Internet:

<<http://www.w3c.es/divulgacion/guiasbreves/PrivacidadP3P>>

W3C. *Guía Breve de Seguridad*. Disponible en Internet:

<<http://www.w3c.es/divulgacion/guiasbreves/Seguridad>>

W3C. *HTML 4.01 Specification*. 1999. Disponible en Internet:

<<http://www.w3.org/TR/html401>>

W3C. *Lenguaje de Ontologías Web (OWL): Vista General*. 2004. Disponible en Internet:

<<http://www.w3.org/TR/2004/REC-owl-features-20040210/>>

W3C. *XML Signature Syntax and Processing*. 2002. Disponible en Internet:

<http://www.w3.org/TR/2002/REC-xmlsig-core-20020212/>

ZWICKER Ronaldo, *Aprendizagem e Uso de Sistemas*. Anais do XXIV Congresso Nacional de Informática, Sao Paulo, 1991

