

**PS-937**

## **SEMANTIC WEB IN HIGHER EDUCATION**

Lic. Marcelo Claudio Périssé (Universidad Nacional de La Matanza, Departamento de Ingeniería e Investigaciones Tecnológicas, Buenos Aires – Argentina) -

[mperisse@unlam.edu.ar](mailto:mperisse@unlam.edu.ar)

Dr. Jorge Luis Narváez (Universidad Nacional de La Matanza, Secretaría de Planeamiento y Control de Gestión, Buenos Aires – Argentina) - [jnarvaez@unlam.edu.ar](mailto:jnarvaez@unlam.edu.ar)

The ideas expressed in this project largely reflect the need for the satisfaction of the requirements of scientific and technical information, and generate new perspectives on a model that allows linking Directory of Scientific Publications, the Digital Libraries University and Scientific Communities with key players-users of these requirements, such as researchers, teachers and students participate in the model learning, in conjunction with the community producer of goods and services.

Keywords: Semantic Web, knowledge management, higher education, sistemas de información, infraestructura para la descripción de recursos.

### **LA WEB SEMÁNTICA EN LA EDUCACIÓN SUPERIOR**

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 los Directorios de Publicaciones Científicas, las Bibliotecas Digitales Universitarias y las Comunidades Científicas con los principales actores-usuarios de dichos requerimientos, como son los investigadores, los docentes y los alumnos partícipes en los modelos de enseñanza-aprendizaje, conjuntamente con la comunidad productora de bienes y servicios.

Palabras llaves: Web Semántica, gestión del conocimiento, educación, information systems, superior, Resource Description Framework.

## 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 UNESCO, la integración de los requerimientos mínimos para el desarrollo de aplicaciones informáticas intervinientes en el proceso Enseñanza-Aprendizaje, mediante la utilización de metodologías simplificadas que se apoyan en la modelización del sistema a través de las técnicas de modelado Unified Modeling Language (UML) y en el Rational Unified Process (RUP).

Concretamente el trabajo presenta la construcción de un catálogo de publicaciones científicas accesible para las Bibliotecas Digitales Universitarias, haciendo hincapié en la capacidad de generar actividades de almacenamiento, recuperación y gestión de información apropiada, de forma: ordenada, distribuida y compartida.

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 mezclar datos estructurados y semiestructurados que son exportados y compartidos a través de diferentes aplicaciones.

Además de la interoperabilidad de datos, el RDF provee una semántica para metadatos entendible por aplicaciones informáticas desarrolladas por la 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". Además 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. 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 XML (eXtensible Markup Language) 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.

El proyecto se encuentra enmarcado en las políticas definidas por la Rede ScienTI, liderada por Brasil, quien promueve un espacio público y cooperativo de interacción entre los actores de los sistemas y comunidades de ciencia, tecnología e innovación de sus países miembros y en la que se integran: la plataforma Lattes, como sistema que permite el acceso a la información científica a través de los currículum de los docentes-investigadores; el Modelo Scielo, como infraestructura para la descripción de publicaciones científicas y la Metodología LILACS, para la generación de bases de datos bibliográficas

A su vez se encuentra enmarcado en los lineamientos promovidos por el Sistema de Información Universitaria (SIU): SIU-Toba y SIU-Biblioteca, perteneciente a la Secretaría de Políticas Universitarias del Ministerio de Educación Ciencia y Tecnología de la Argentina; el cual ha formado parte, junto a UNESCO y la Library of Congress, en el desarrollo de IsisMarc, un software para bibliotecas que

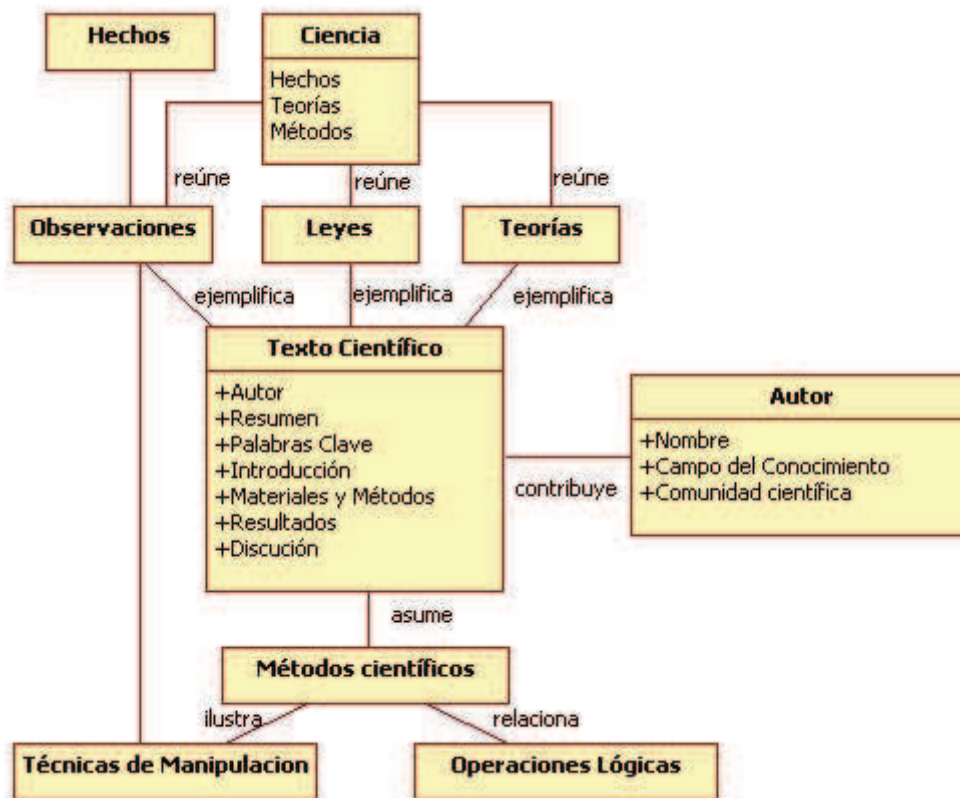
permite el ingreso de información para el formato bibliográfico MARC21 sobre base de datos de tecnología Isis. Marc21 es un estándar internacional de catalogación que permite el intercambio de registros catalográficos entre bibliotecas.

Dentro de este marco de 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 (OAI) quien desarrolla y promueve estándares de interoperabilidad cuyo fin es facilitar la eficiente en la diseminación de contenidos. Siendo que 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.

Es de destacar que la Investigación y Desarrollo del presente proyecto incluye el estudio los vocabularios controlados (tesauro), su modelo estructural (taxonomía) y su representación explícita y formal como estructura conceptual de al base de conocimiento (ontología).

### **Materiales y Métodos**

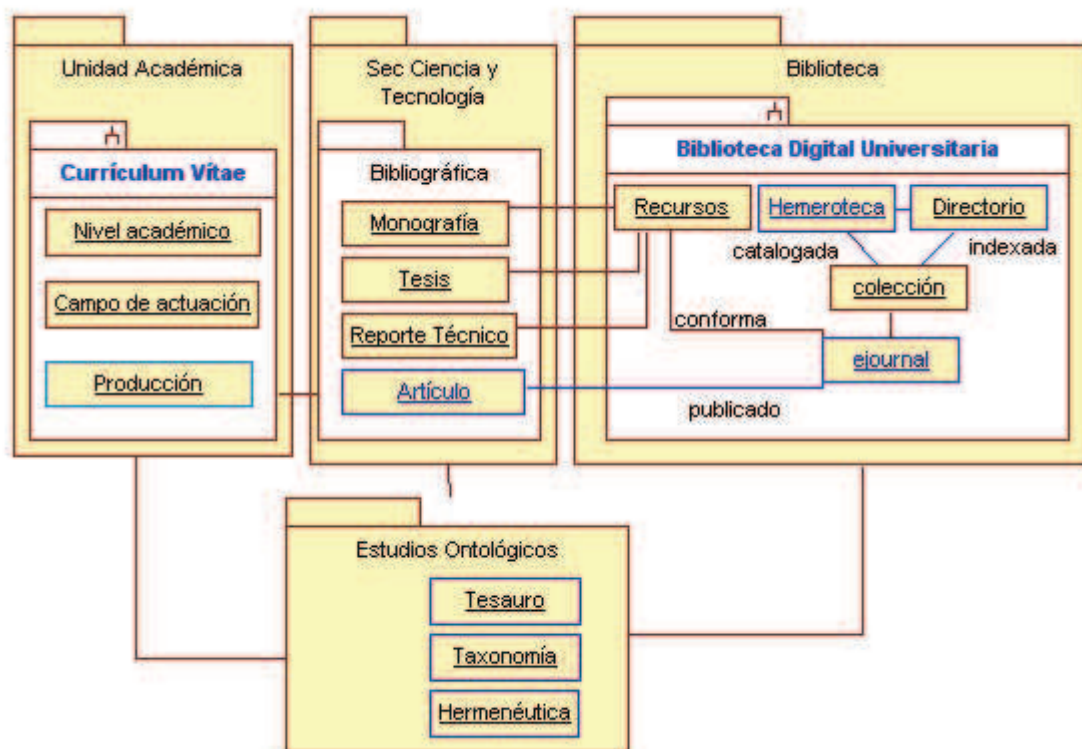
Basados en la conceptualización de ciencia como: la constelación de hechos, teorías y métodos reunidos en los libros de texto actuales y en donde los científicos son hombres que, obteniendo o no buenos resultados, se han esforzado en contribuir con alguno que otro elemento a esa constelación particular; (KHUN:1962) se procedió a un primer modelo conceptual tal lo representado en la figura 1 y en el que se describe 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.



**Figura 1. Modelo conceptual de Ciencia basado en Thomas Khun (1962)**

Otro modelo conceptual con el cual se trabajó para determinar la infraestructura necesaria que permita realizar una apropiada gestión del conocimiento producido por la Universidad, cuenta con los siguientes sistemas y es representada en la figura 2:

- Una base de datos que contiene el currículum de los docentes universitarios que permite principalmente evaluar los progresos de los docentes 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:
  - a. un catálogo de publicaciones científicas,
  - b. un acervo de disertaciones y tesis aprobadas por la universidad
  - c. 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.



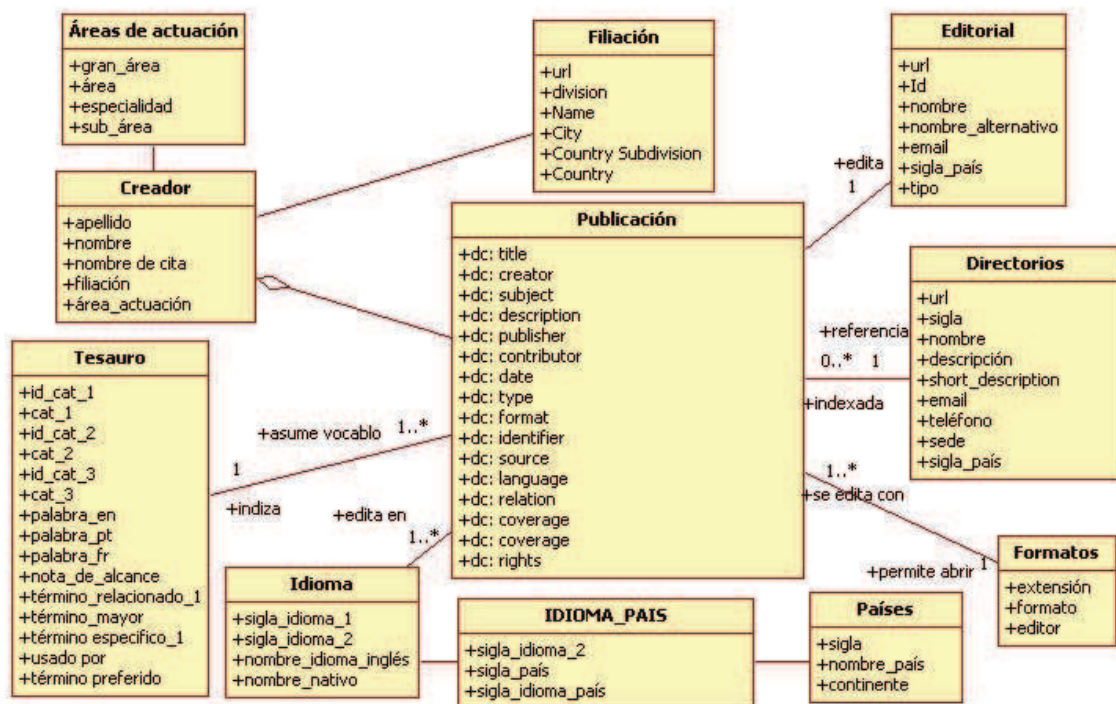
**Figura 2. Red de conocimiento científico y tecnológico**

A partir de dicho modelo conceptual se procedió a la construcción del diagrama de clases para la construcción del catálogo de publicaciones seriadas (hemeroteca) científicas que se incorpora a la Biblioteca Digital Universitaria, tal se puede observar en la figura 3; 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 en 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:

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



**Figura 3 Diagrama de Clases de la Hemeroteca Digital**

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 4



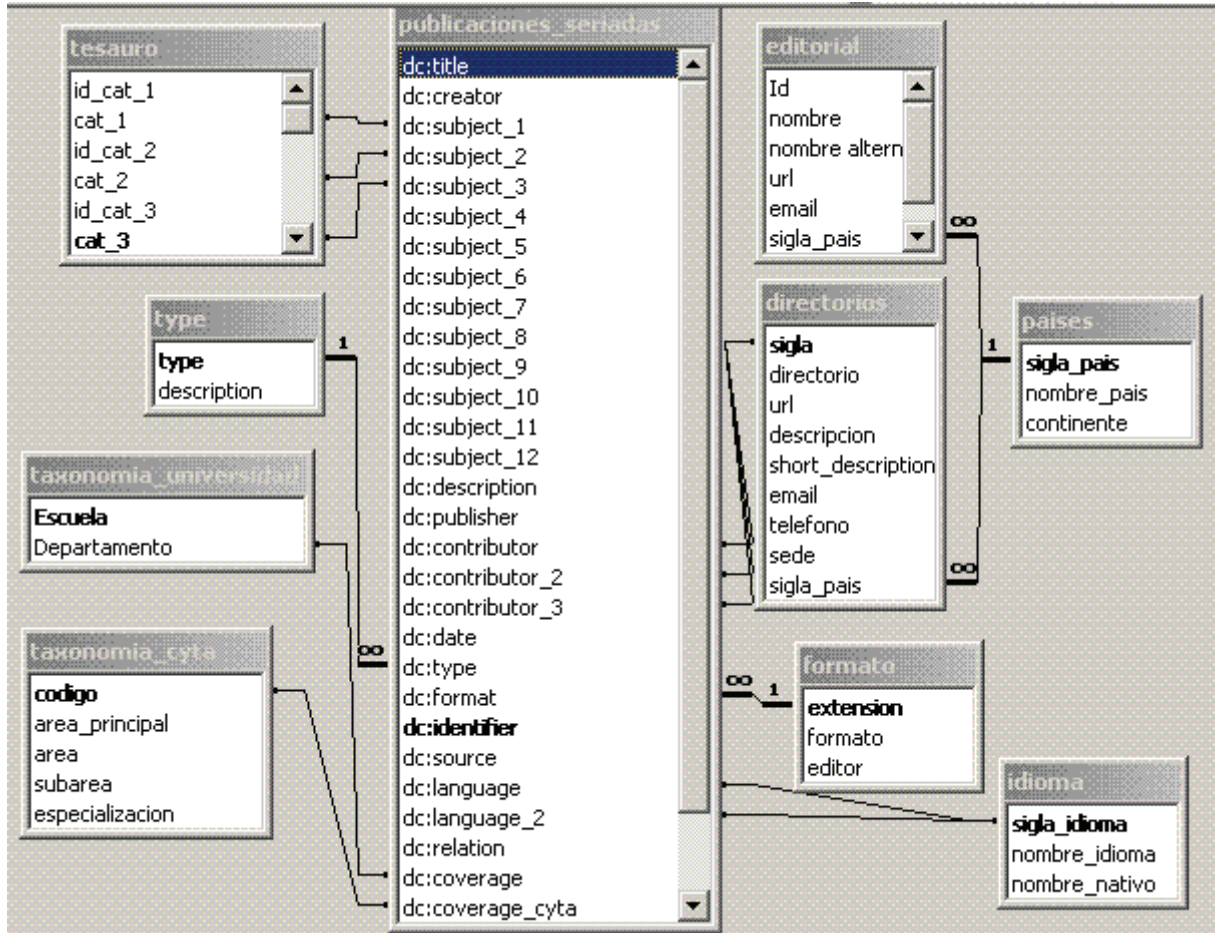


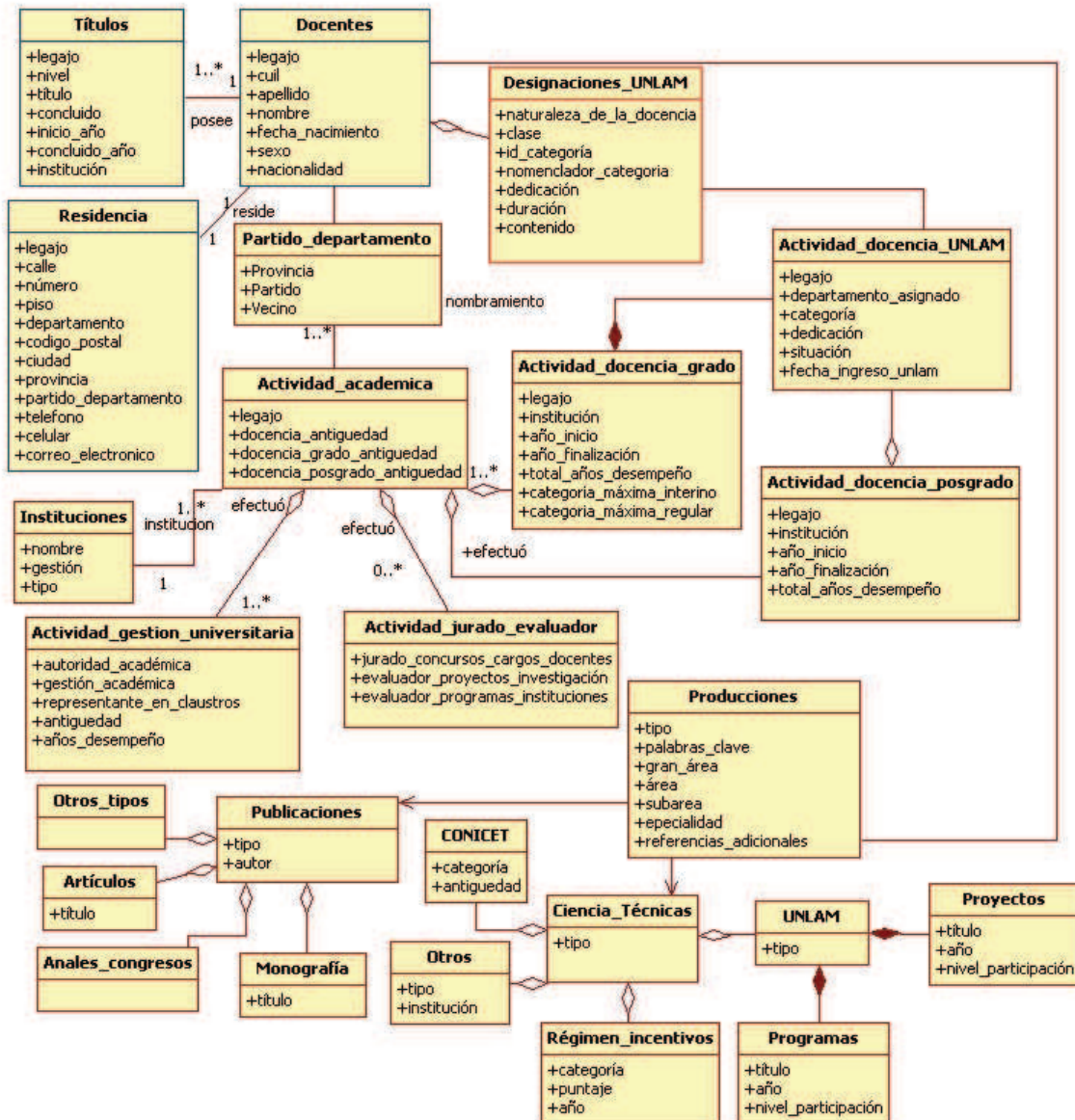
Figura 4. Hemeroteca digital - diagrama de relacional de la Base de Datos Access

Estructurando la información bajo el modelo 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#">
- <dc>
  <dc:title>Journal of Research for Consumers</dc:title>
  <dc:subject_1>ADMINISTRACION</dc:subject_1>
  <dc:subject_3>NEGOCIOS</dc:subject_3>
  <dc:description>The primary objective of the Journal of Research for Consumers is to provide
  information to individuals to enable them to better understand their roles as consumers.
  Also, the Journal aims to encourage consumer researchers to use their skills for the direct
  benefit of consumers, without the need to produce research outcomes of managerial
  relevance. It is hoped that the dissemination of research information that has been
  generated for the specific purpose of educating consumers will provide some
  counterbalance to the quantity of consumer research that is currently published in the
  interests of marketers</dc:description>
  <dc:contributor>DOAJ</dc:contributor>
  <dc:type>científica</dc:type>
  <dc:identifier>14446359</dc:identifier>
  <dc:source>http://www.jrconsumers.com/</dc:source>
  <dc:coverage>ciencias economicas</dc:coverage>
</dc>
```

Figura 5. 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. Este proceso de integración permite realizar la relación entre los campos del conocimiento y las comunidades científicas.



## Resultados

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, ayudando así a que los resultados de los proyectos de investigación y desarrollo lleguen al sector empresarial y contribuyan a la mejora de sus procesos



productivos, conformándose un servicio de utilidad pública, con ámbito de actuación nacional y multisectorial.



Adicionalmente el programa permite capacitar a los investigadores sobre la correcta estructuración de sus resultados (reportes técnicos, artículos, tesis, monografías) para ser alojada en la Web Semántica y la utilización de herramientas, técnicas y métodos para la recopilación de información científica en el mismo espacio.

La computadora es 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; además la evaluación de la eficiencia se basa en una lógica y esta evaluación no tiene ningún inconveniente de ser programada en una computadora para ser automatizada mediante sistemas de inteligencia artificial basados en reglas del tipo tripleta (sujeto, predicado, objeto) como el prolog el cual se adecua al esquema del modelo RDF .

## Discusión

En este proyecto de Investigación y Desarrollo sobre la gestión del conocimiento quedan algunos de 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 marchamo 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 científica y técnica, conceptualizando a la bibliometría como un método científico de la ciencia de la información o ciencia de la documentación que constituye la aproximación cuantitativa que permite el desarrollo de la teoría general de esta

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

Respecto al proceso Enseñanza-Aprendizaje, es de esperar que en la instrumentación del presente modelo permita que el lenguaje disciplinar deje de ser una barrera aparentemente imposible 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.

## Conclusión

En este trabajo se presenta 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 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 consideramos 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, etc.) 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).

## Bibliografía

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

BERNERS-LEE, Tim; HENDLER, James; LASILA, Ora. La red semántica. *Investigación y Ciencia*\_\_jul. 2001

BRUN E., Ricardo. XML y la gestión de contenidos [on line]. *Hipertext.net*, núm. 3, 2005. <<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 [on line]. "*Hipertext.net*", núm. 3, 2005. <<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

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. 1988. On Visual Formalisms. In: *Communications of the ACM*, May 1988, Volume 31 Number 5, Pag. 514.

LÓPEZ GUZMÁN, Clara. Modelo para el Desarrollo de Bibliotecas Digitales Especializadas, <http://www.bibliodgsca.unam.mx/tesis/tes7c1lg/tes7c1lg.htm>

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

LORSCH, Jay W. y SHELDON, Alan, The individual in th organization: A system View, In: *Managing Group and Intergroup Relations* Ed. Jay W. Lorsch y Poul R. Lawrence, Homewood, Ill: Irwin-Dorsey, 1972, pp. 161-82.

NICOLESCU, Basarab *O Manifesto da transdisciplinaridade*. São Paulo, Triom: 1999. Tradução do Francês por Lúcia Pereira de Souza.

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

PEMAU ALONSO, Julio. Introducción a OAI-PMH y su implantación en el portal e-Revistas. Disponible en:  
[http://www.emwis.org/documents/pdf/20050404\\_JA\\_cursos\\_erevistas\\_completo.pdf](http://www.emwis.org/documents/pdf/20050404_JA_cursos_erevistas_completo.pdf)

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

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

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

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

Enlaces:

Tecnociencia: <http://www.tecnociencia.es>

E-Revistas: <http://www.teconociencia.es/e-revistas>

Dublin Core: <http://www.dublincores.org>

OAI: <http://www.openarchives.org>

DOAJ: <http://www.doaj.org/home>

DIALNET: <http://dialnet.unirioja.es>

Directorio y recolector de recursos digitales:  
[http://roai.mcu.es/es/consulta/exportar\\_registro.cmd?id=39406&exportar=Exportar](http://roai.mcu.es/es/consulta/exportar_registro.cmd?id=39406&exportar=Exportar)

SIU-Toba: <http://www.siu.edu.ar/soluciones/toba/acercade/>

SIU-Biblioteca: <http://www.siu.edu.ar/soluciones/bibliotecas/>

SciELO: [http://www.scielo.org/criterios\\_certificacao\\_sites\\_es.html](http://www.scielo.org/criterios_certificacao_sites_es.html)



## Currículo Vitae

### Marcelo Claudio Périssé

Licenciado en Administración de Empresas. Actualmente es Profesor Investigador de la Universidad Nacional de La Matanza y Profesor Titular de la Universidad de Ciencias Empresariales. Director de la Editorial Ciencia y Técnica Administrativa y Director y Editor Responsable de la revista científica Técnica Administrativa (issn 1666-1680). Tiene experiencia en el área de Administración con énfasis en Métodos Cuantitativos y Tecnología de la Información.

### Jorge Luis Narváez

Ph D in Business Administration, Master en Comercio Internacional (Univ. Complutense), Contador Público y Lic. en Administración (UBA). Actualmente es Secretario de Planeamiento y Control de Gestión en la Universidad Nacional de La Matanza y Profesor Titular Ordinario del Área Administración de su Departamento de Ingeniería e Investigaciones Tecnológicas. Se ha desempeñado como Secretario Académico de la UNLAM, Secretario de Investigación y Desarrollo en la Universidad Abierta Interamericana y Decano de la Facultad de Ciencias Empresariales de la misma universidad, Vice Decano del Departamento de Ciencias Económicas en la UNLAM y Responsable del Área Management de UADE Senior. Autor de once libros sobre su especialidad.