

# Barreras y estrategias de solución para la implementación de un portal semántico en investigación en la Facultad de Ingeniería y Arquitectura de la Universidad de San Martín de Porres

Barriers and solution strategies for the implementation of a semantic portal research at the Faculty of Engineering and Architecture at the University of San Martin de Porres

Recibido: julio 09 de 2015 | Revisado: setiembre 18 de 2015 | Aceptado: octubre 15 de 2015

NORMA LEÓN LESCANO<sup>1</sup>  
MILAGROS QUISPE RODRÍGUEZ<sup>2</sup>  
CÉSAR PORRAS QUINTO<sup>3</sup>  
SANDRA REGGIARDO BARRETO<sup>4</sup>

## ABSTRACT

This paper describes the barriers identified for the collection and dissemination of knowledge in research centers of the FIA USMP and proposes strategies for the implementation of a semantics portal as a tool for research knowledge management.

**Keywords:** barriers, knowledge, dissemination, identified portal, management

## RESUMEN

En este trabajo, se describen las barreras identificadas para la obtención y difusión del conocimiento en los centros de investigación de la FIA USMP y en él se plantean las estrategias para la implementación de un portal semántico como herramienta para la gestión del conocimiento en investigación.

**Palabras clave:** barreras, conocimiento, difusión, identificación, portal, semántico, gestión

---

1 Facultad de Ingeniería y Arquitectura de la Universidad de San Martín de Porres  
nleonl@usmp.pe  
2 aquisper@usmp.pe  
3 cporras@usmp.pe  
4 sreggiardob@usmp.pe

## Introducción

El propósito de este trabajo es describir las barreras existentes en las unidades de investigación de la Facultad de Ingeniería y Arquitectura (FIA) de la Universidad de San Martín de Porres (USMP) y las estrategias que se proponen para la gestión del conocimiento mediante la implementación de un portal semántico.

Para el desarrollo de esta investigación, se ha analizado el estado del arte de los portales semánticos como medio de captura y difusión de información compleja como es la que se manipula en investigación.

Luego se ha realizado un levantamiento de información entre los distintos grupos de investigación de la FIA y se obtuvo como resultado la identificación de un proceso de producción académica común a todas las unidades de investigación. Se identifica este proceso alineado a la necesidad urgente de mostrar la producción escrita dentro de las unidades de investigación y ser visibles para las organizaciones que elaboran los rankings de producción científica formal (QS, 2012) e informal. Nos referimos a la publicación informal, en que nacen y se difunden desde los centros de investigación sin necesidad de ser indexadas o pasar por procesos formales de publicación, pues se entiende que en unidades de investigación jóvenes, la formalidad tiene su tiempo de madurez y no es un buen aliado para la visibilidad de los resultados de la investigación.

Este proceso ha permitido identificar barreras internas que han colaborado con la característica de invisibilidad de la producción académica por lo que se plantean estrategias para salvar estas barreras ayudando a la gestión de la información y hacer uso de la implementación de una tecnología emergente capaz de gestionar el conocimiento integrado y realizar inferencia sobre la misma.

## Problema objetivo

En los centros de investigación, existe una creciente necesidad de acceso y distribución de la información de manera constante y fluida para el desarrollo de sus actividades. La información, por lo general, se encuentra distribuida en diferentes tipos de repositorios, propios, privados y públicos siendo difícil la obtención de todos estos repositorios a la vez, y casi imposible que los buscadores informáticos puedan realizar inferencia directa sobre la información obtenida. Además, es característica de las organizaciones de investigación tener autonomía sobre la información requerida y generada, lo que ha permitido que la producción de información haya ido en aumento contribuyendo esta característica a la difusión de la ella en diferentes medios lo que se ha convertido en un parámetro de medición de la performance de las instituciones educativas.

Las unidades de investigación de la FIA no escapan a esta realidad. Cuentan con sistemas de gestión tradicionales: archivos físicos en diferentes formatos, un cúmulo de links de páginas web, bibliotecas digitales existentes, repositorios en nube, entre otros. Ellas permiten al investigador realizar el proceso de recolección de información científica y técnica, aunque no con la velocidad y capacidad necesaria a los intereses de cada investigador. La producción escrita, sin embargo, no ha registrado avances de acuerdo con el crecimiento de las actividades de investigación que se vienen realizando.

Este trabajo identifica las barreras que han colaborado con esta característica y plantea como estrategia la gestión de la información usando la tecnología semántica.

## Problema

¿En qué medida un portal semántico de investigación ayudará a las estrategias planteadas ante las barreras de difusión del conocimiento de la investigación FIA?

## Objetivo

Alinear la implementación de un portal web semántico a las estrategias de solución mediante las barreras de difusión de la producción escrita de la investigación en la FIA.

## Marco de trabajo

En esta sección, se describe el proceso de trabajo realizado, comenzando por el estado del arte de una de las herramientas de gestión del conocimiento (KM), la descripción del proceso de creación y difusión de los resultados de la investigación, las barreras existentes, y las propuestas de estrategias de difusión de producción de investigación mediante la implementación del portal que ayude en la gestión de conocimientos de investigación.

## Estado del arte

La gestión del conocimiento en unidades de investigación que existen al interior de instituciones educativas ha permitido que las instituciones educativas como las universidades que mejor manejan la gestión de su producción científica y académica estén posicionadas en el ranking de las mejores universidades a nivel mundial (Scientometrics, 2012). Son estas instituciones las que se han encargado de proponer tecnología que pueda soportar sus necesidades de obtención y difusión de información.

La web fue una herramienta muy buena para esto, pero pronto rebasó las expectativas, de ser una solución adecuada y controlada. Inicialmente, se transformó en una colección de telarañas independientes sin la capacidad de integración entre ellas. Entonces nace la necesidad de integrar la información, independientemente de los formatos y características de cada grupo. La información había crecido y diversificado en demasía, se empieza a trabajar en el desarrollo de herramientas que soporten esta característica. La web semántica ha contribuido a este esfuerzo dándoles a los investigadores información

altamente relevante por la capacidad de la tecnología semántica de realizar inferencia sobre la información solicitada, proporcionando al investigador información organizada y coherente.

La web semántica con sus esfuerzos para enfocarse en la omnipresencia de los protocolos web / estándares, el énfasis en los micro formatos, la evolución de los vocabularios como folksonomías y tagclouds, la definición de ontologías y la explotación de las máquinas de aprendizaje con la inclusión de los Web Service Modelling Ontology (Wang, Gibbins, Payne, Saleh, & Sun, 2007) proporciona a las computadoras la capacidad de almacenamiento, cálculo, gestión e indexación de enormes cantidades de datos existentes en la web que posibilitan, de manera automática, el descubrimiento, la agregación, la organización, correlación, análisis, interpretación e inferencia de la información existente.

Esto permite que el investigador tenga acceso a información relevante en menor tiempo. Por lo que el mundo académico de investigación se ha enfocado, en el gran ecosistema de servicios y herramientas, que se integran a esta gran malla de datos (Parastatidis, Viegas, & Hey, 2009). Nace entonces la necesidad de agrupar estas herramientas y servicios emergentes que permiten la incorporación de inferencias apoyadas con los motores de búsqueda que entienden el problema del investigador y que le permiten enfocarse, en su producción, sin invertir demasiado tiempo en la obtención de la información inicial. Los portales semánticos nacen de esta necesidad del mundo académico, en que la semántica y los agentes de software cobran gran importancia como su principal característica.

SEAL es una de las primeras propuestas para hacer accesible la información a los investigadores y a agentes de software usando una web semántica, que fue desarrollada por la Universidad de Karlsruhe. SEAL permite

que los sitios web puedan alimentarse directamente de la web semántica. Asimismo, se centra en la propuesta de una estructura basada en los medios que hacen que los sitios web semánticos sean accesibles desde el exterior, Figura 2 SEAL Communication between human and/or software agents, es decir, la semántica basada en la navegación, la consulta semántica, la consulta con la similitud semántica, y el acceso a las máquinas, a la información semántica (Maedche, Staab, Stojanovic, & Stojanovic, 2001).

Luego encontramos el portal JuNii Plus-que que propone la recuperación de la información basada en una combinación de ontologías y un modelo vectorial, (Figura 3), auspiciado por el Instituto Nacional de Informática y la empresa privada para las universidades e instituciones académicas japonesas. Posibilita al investigador ver las relaciones de búsqueda, las ontologías para poder inferir conocimiento sobre estas relaciones (Kando, Kanazawa, & Miyazawa, 2006).

También encontramos la propuesta My Experiment de la Universidad de Manchester que entre sus proposiciones lanza una estructura ontológica capaz de soportar el conocimiento de comunidades de investigación Figura 6. My Experiment se basa en cuatro pilares: a) Facilitar la gestión y el intercambio de objetos de investigación (OR); b) Apoyar un modelo social; c) Proveer un ambiente abierto extensible; d) Proporcionar una plataforma para la investigación-acción. (Newman, Bechhofer, & De Roure, 2009).

Los portales semánticos han ido en progreso, en sus implementaciones hacia el soporte de la investigación y los investigadores, sin preocuparse mucho por la usabilidad dando por entendido que quienes lo usan son personas altamente especializadas; no se tomó en cuenta que la velocidad de obtención de la información es parte de los requerimientos no funcionales de todo sistema. Para el año 2010, el problema de usabilidad ya era evidente, los portales habían

crecido en funcionalidad al investigador, pero eran difíciles de usarlos para mostrar la navegación y consulta de datos semánticos.

La universidad de Indiana propone una solución a esta deficiencia de las tecnologías de la web semántica. En la Figura 8, se presenta el Portal Web Semántico (SWP) como una plataforma ligera de ayuda a los usuarios para organizar, buscar y visualizar los datos pertinentes semánticos de una manera significativa. La propuesta del SWP se ha demostrado, probado y evaluado en varios casos de uso diferentes, tales como un estudio a un grupo mediano de usuarios del mismo portal, un conjunto de datos de catálogo de un portal del gobierno, un portal de centro de salud del paciente y un portal de Linked Open Data para datos de bio-química (Ding et al., 2010), los resultados fueron exitosos.

Con el enfoque SWP se mejora la usabilidad de un portal semántico. La Universidad de Kebangsaan propone una plataforma ligera orientada hacia la investigación dentro de un campus universitario. Este portal, orientado netamente a fortalecer la distribución y obtención de información de investigación relacionada con publicación de los departamentos, organizaciones, eventos y grupos de investigadores entre otros, tiene entre sus características la agilización del proceso de emisión de resultados, permitiendo al usuario académico, personal de apoyo, y otros miembros de la institución publicar la información sobre los estudios de investigación. Además, proporciona a los usuarios finales una mejor visión de la estructura de la investigación en los centros de universitarios que permiten a los usuarios llevar a cabo la comunicación cruzada entre los profesores y grupos de estudio mediante el uso de la búsqueda semántica de información (Hidayat, Yahya, Azman, & Noah, 2012).

Finalmente, existen muchas instituciones educativas que vienen usando portales semánticos para la gestión del conocimiento de investigación de los distintos grupos

de investigación existentes en sus campus. Estos grupos conformados por profesores, investigadores, programadores, estudiantes de Doctorado, Maestría y aquellos que tienen proyectos financiados externamente y las principales actividades y conjuntos de datos para estos grupos de investigación

son los documentos, cursos, presentaciones, eventos, conjuntos de datos propios de la investigación, software, hardware, datos de personas y financiación; actualmente ya no tienen la limitación de obtención, mantenimiento y actualización de los datos (Ding y otros, 2010).



Figura 1. Ecosistema de servicios y herramientas (Parastatidis, Viegas, & Hey, 2009)

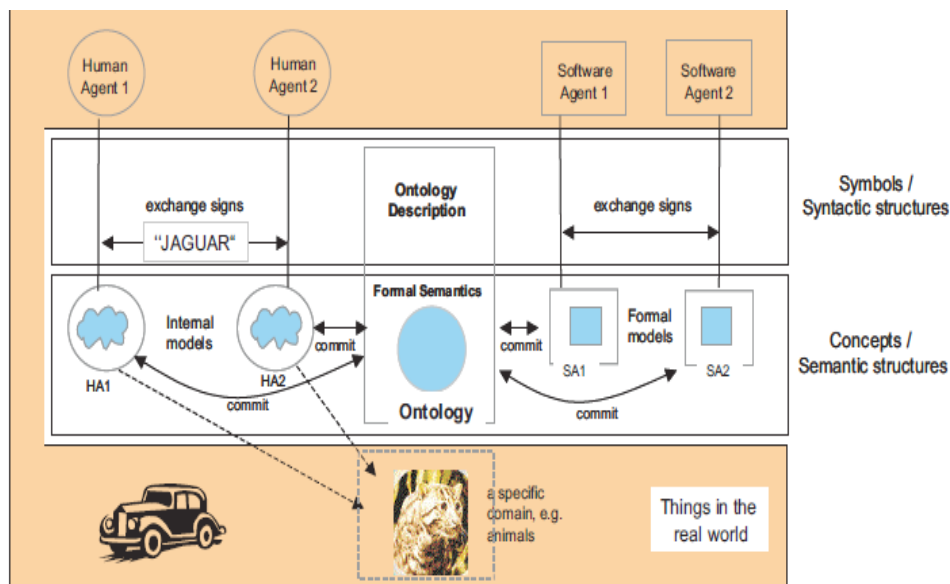


Figura 2. SEAL Communication between human and/or software agents

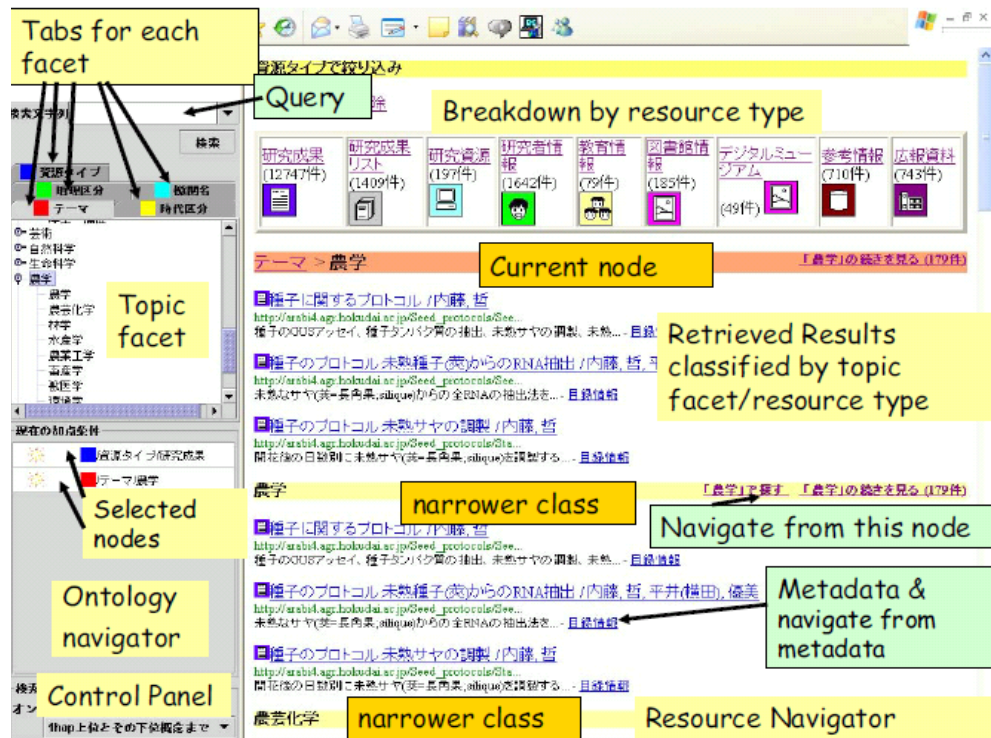


Figura 3. J uNii Plus (Kando, Kanazawa, & Miyazawa, 2006)

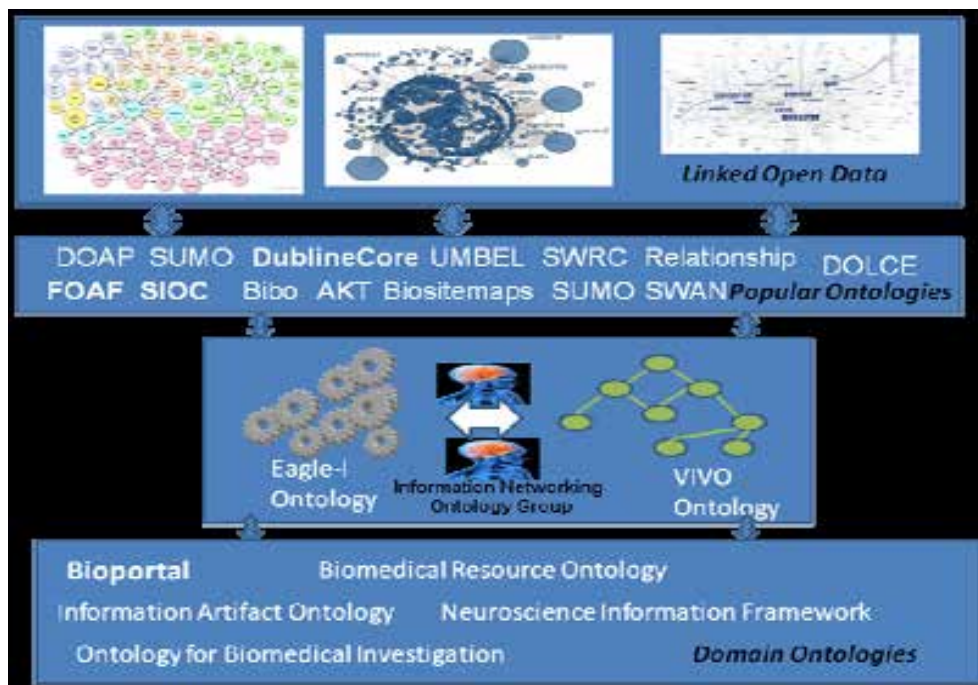


Figura 4. Information Networking Ontology Group framework



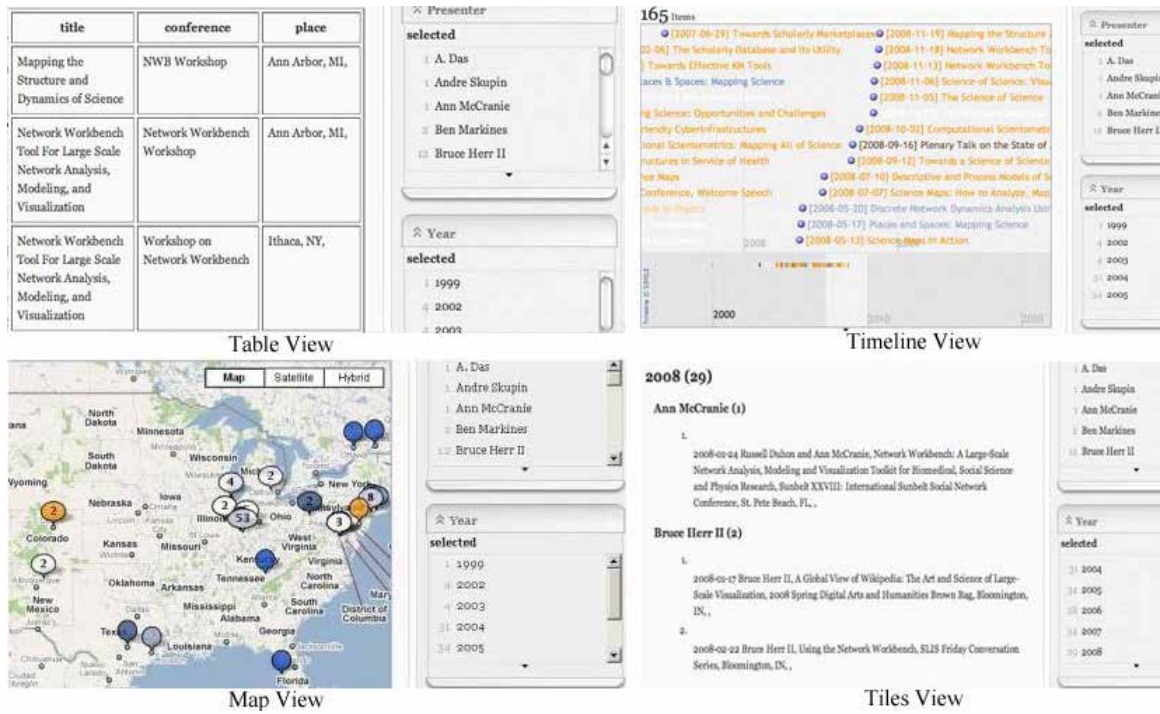


Figura 5. Screenshots of SWP's semantic visualization

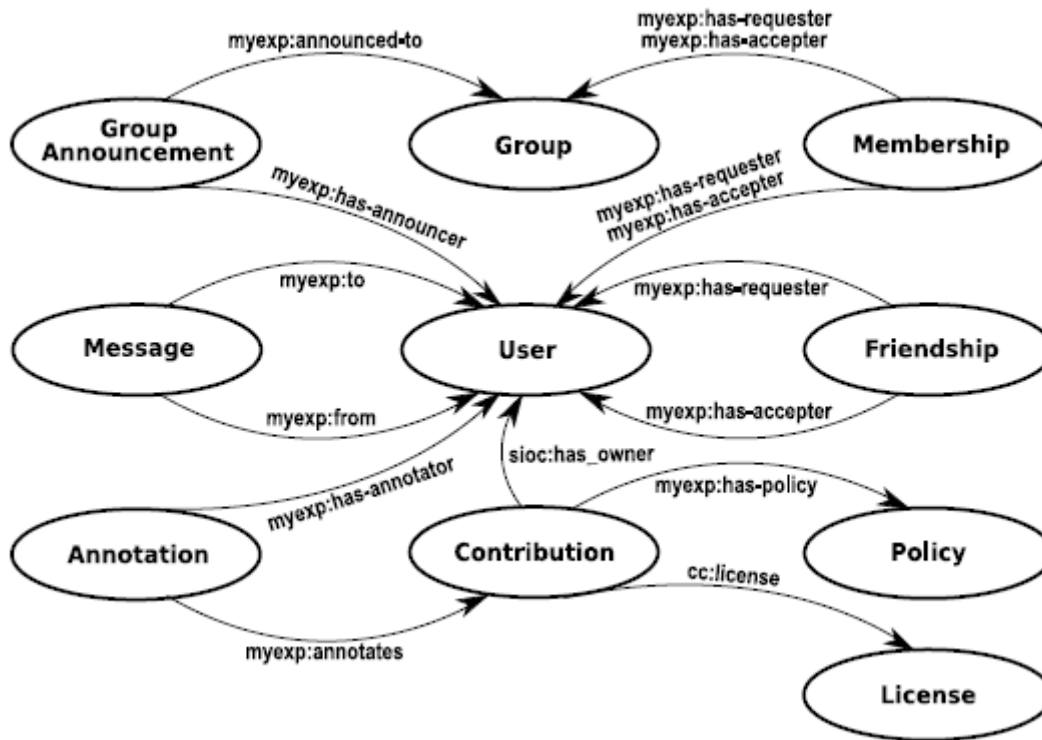


Figura 6. My Experiment Ontología Base

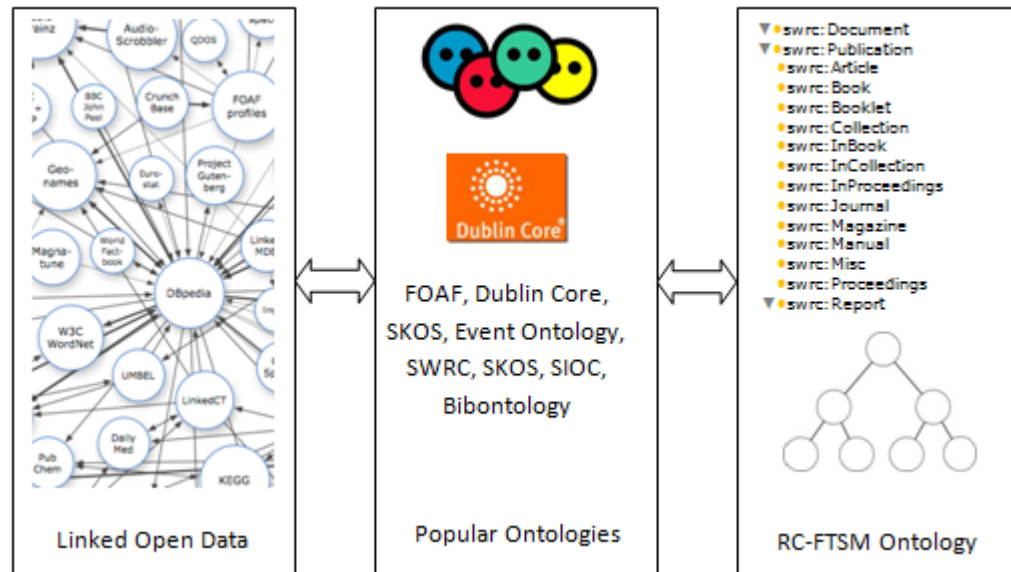


Figura 7. The Ontology Framework Information Network

The screenshot shows the InfoVis Lab website at Indiana University Bloomington, IN. The page features a navigation menu on the left with categories: People, Research, Teaching, and Publications. The main content area is titled "Publications (Group By Category)" and includes a "Download" button for "All References (EndNote file)".

**Submitted/In Press**

- Albert, Paul, Kristi Holmes, [Katy Börner](#), and Mike Conlon. "Research Discovery through Linked Open Data". *ACM/IEEE - CS Joint Conference on Digital Libraries 2012*.
- Skupin, Andre, Joseph R. Biberstine, and [Katy Börner](#). "Visualizing the Topical Structure of the Medical Sciences: A Self-Organizing Map Approach".
- [Börner, Katy](#), Robin M. Wagner, Nianli Ma, Joseph R. Biberstine, Rediet Berhane, Hong Jiang, Susan E. Ivey, Katrina Pearson, and Carl McCabe. "[Introducing the Science of Science \(Sci2\) Tool to the Reporting Branch, Office of Extramural Research/Office of the Director, National Institutes of Health](#)". *Workshop on the Science of Science Measurement, December 2-3, 2010, Washington D.C.*
- [Börner, Katy](#), [Richard Klavans](#), Michael Patek, [Angela Zoss](#), Joseph R. Biberstine, Robert Light, Vincent Larivière, and Kevin W. Boyack. "[Design and Update of a Classification System: The UCSD Map of Science](#)".

**2012**

- Schamhorst, Andrea, [Katy Börner](#), and [Peter van den Besselaar](#). 2012. "[Foreword](#)". In *Models of Science Dynamics: Encounters Between Complexity Theory and Information Science*. Springer Verlag.

Figura 8. Portal semántico de investigación de la Universidad de Indiana



### Proceso de investigación

Los procesos que soportan la investigación en las distintas unidades de investigación en la FIA, difieren entre unidades, laboratorios, áreas, escuelas e institutos; sin embargo, el proceso de producción académica sigue actividades similares que se muestran en la Figura 9.

Esta estructura estándar que se ha esbozado a base de las entrevistas realizadas, desde el punto de vista de la obtención y generación de documentación que sustentan los logros y avances, en los proyectos de investigación, alrededor de las situaciones de interés de cada grupo de investigación. Los grupos identificados articulan su conocimiento, de una manera estructurada, identifican, acopian y estructuran según su interés y publican los resultados que van obteniendo, en su proceso, en diversos medios.

Los grupos de investigación tienen sus propios canales para que el resultado de sus esfuerzos sea utilizado, analizado y criticado por otros grupos de interés. Esta característica obliga al investigador a usar e integrar herramientas que faciliten su accionar en las distintas fases del proceso de investigación aplicada como se muestra en la Figura 9.

Desde su fase inicial, el proceso exige la necesidad de publicar información para ser compartida con la comunidad científica,

desde escribir un artículo de marco conceptual en donde se hace énfasis en la identificación, descripción del problema y el aporte a la solución. Continuando con un artículo del estado del arte, en donde la investigación da énfasis a una revisión actualizada de métodos y conceptos usados para obtener un conocimiento profundo de ellos, de sus fortalezas y debilidades, en que se puedan identificar un problema conceptual o metodológico. Si el investigador desea y se motiva, puede proponer una serie de hipótesis alternativas de solución, evaluando su viabilidad y pertinencia con criterios sólidos; en este caso, se puede obtener un artículo con una propuesta de investigación, que puede orientar la ejecución de un proyecto así como sustentar su financiación, en caso lo requiera.

Una vez que se tiene una hipótesis de trabajo seleccionada y un enfoque de solución, si la complejidad tecnológica del proyecto lo demanda, puede escribir un artículo de análisis de alternativas tecnológicas, con el que se puede determinar la plataforma tecnológica en la que se apoyará el proyecto o los productos obtenidos en la investigación. Una vez que el proyecto se inicia, se pueden elaborar artículos de caso de estudio. Además, los investigadores pueden identificar la demanda de un producto o servicio, para lo cual se pueden escribir documentos relacionados con ello.

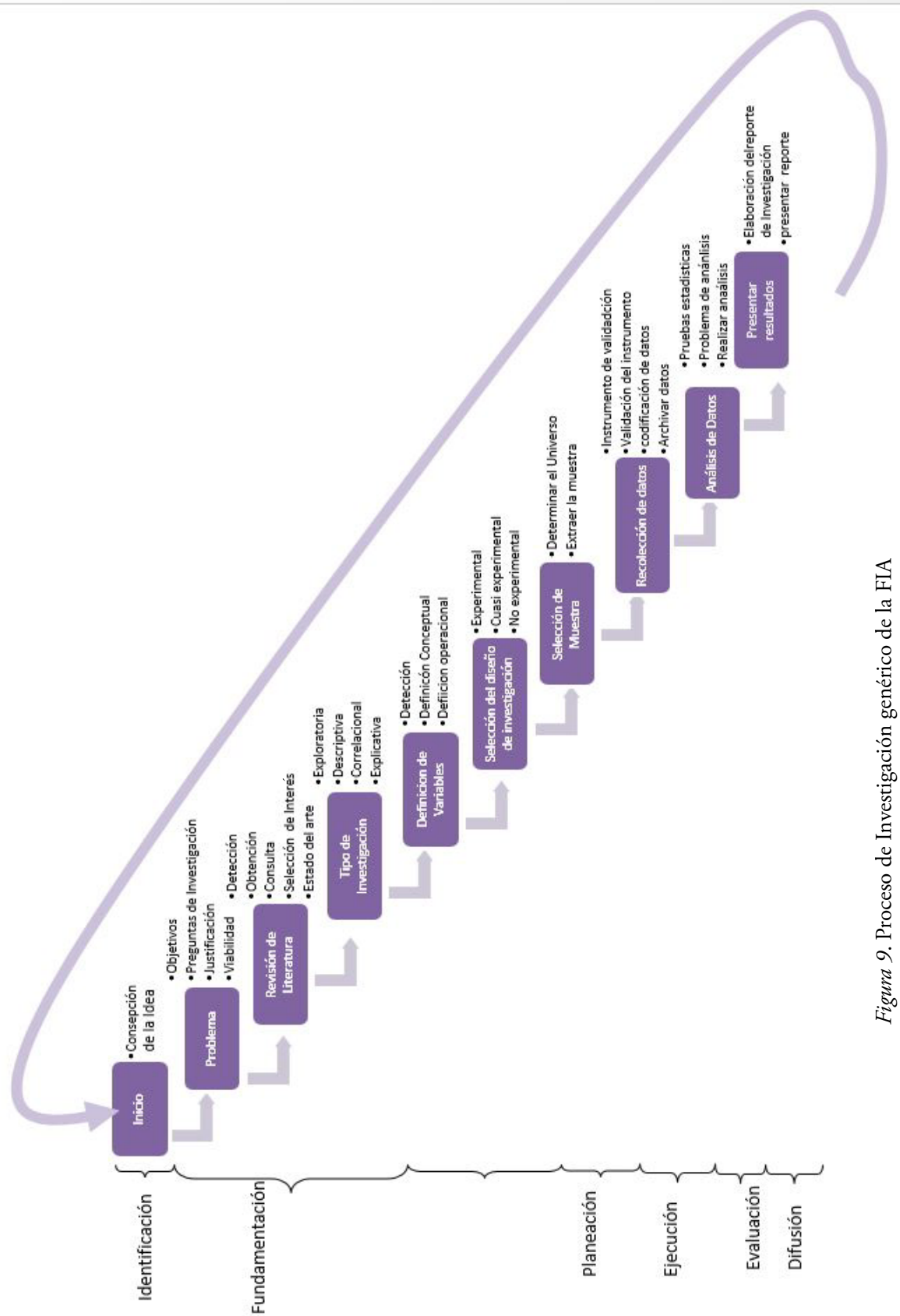


Figura 9. Proceso de Investigación genérico de la FIA

Se puede observar que, en las distintas fases del proceso, existen diferentes actividades involucradas. Estas actividades no son estructuradas, sino que se relacionan entre ellas, por lo que se pueden representar estas

relaciones en un mapa temático conceptual, propio de la investigación aplicada que se realiza como se muestra en la Figura 10 o de la investigación básica como se muestra en la Figura 11.

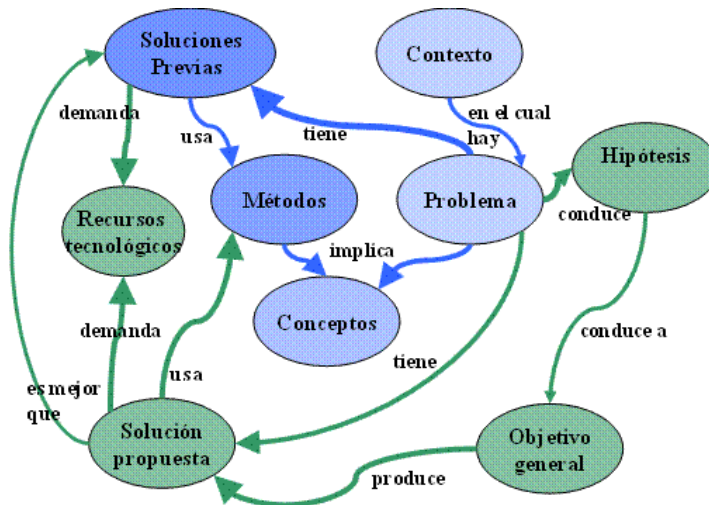


Figura 10. Elementos conceptuales o estructurales que intervienen en el proceso de una investigación aplicada en Ingeniería (Suaza y Ortega, 2007)

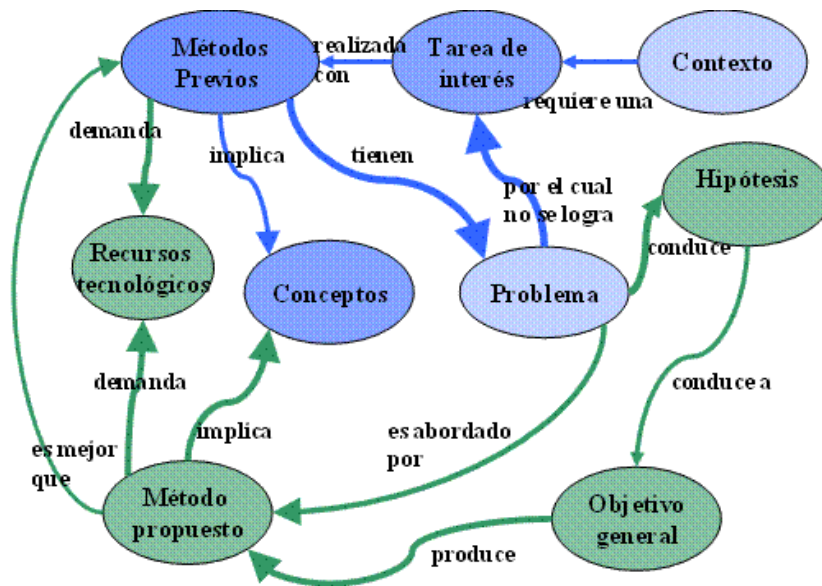


Figura 11. Elementos conceptuales o estructurales que intervienen en el proceso de una investigación básica en Ingeniería

Por tanto, se ha ordenado la producción analizada en la Tabla 1 Momentos para la escritura de artículos en el marco de un proceso de investigación, en donde se indica que

para cada tipo de artículo existe un momento apropiado de escritura como resultado del avance en el proceso de investigación (Suaza y Ortega, 2007).

Tabla 1

*Momentos para la escritura de artículos en el marco de un proceso de investigación*

Artículo	Identificación	Fundamentación	Formulación	Planeación	Evaluación Ejecución	Difusión
Identificación de problema	X					
Identificación de tarea de interés	X					
Estado del arte	X	X				
Propuesta de investigación			X	X		
Caso de estudio	X				X	X
Evaluación empírica					X	X
Módulo de negocios	X	X				
Módulo técnico			X			
Análisis de alternativas tecnológicas			X	X		

Esta producción es difundida en diversos medios, que se detallan en la Tabla 2. Los medios de difusión son sistemas de transmisión de mensajes por parte de un emisor con destino a uno o varios receptores. Dentro del entorno académico-universitario, los emisores son los profesores, los investigadores y otros profesionales relacionados con las unidades de investigación y la universidad. Los receptores de la información son la sociedad en su conjunto, la comunidad académico-universitaria, la comunidad científica y el alumnado.

Las formas de transmisión pueden ser orales, impresas o digitales. En el ámbito académico-universitario, la transmisión de la producción científica, docente o de apoyo (gestión, formación, calidad, etc.) se suele hacer a través de los siguientes medios: publicaciones, producciones audiovisuales, exposiciones, encuentros científicos, registros de la propiedad industrial e intelectual e internet (Rosado et al., 2008). Este trabajo hace énfasis en el uso de Internet como medio.

Tabla 2

*Producción de Investigación a través de Internet*

<b>Producción académica y científica</b>	<b>Tipos</b>	<b>Subtipos</b>	<b>Difusión por internet</b>
<b>Publicaciones</b>	Seriadas	Revista, anuario.	
	No seriadas	Libro, folleto, guía.	x
<b>Medios de comunicación audiovisual</b>	Producciones visuales	Producción gráfica, producción fotográfica (Negativo, copia fotográfica, diapositiva, imagen digital, copia digital, imágenes médicas ),	x
	Producciones sonoras	Producción radiofónica, musical.	x
	Producciones audiovisuales	Producción cinematográfica, televisiva, multimedia (científica, documentales, debate, informativas y experimentales )	x
	Artísticas		x
<b>Exposiciones</b>	Científicas y tecnológicas		x
	Congresos y jornadas	Congresos (conferencias, mesas redondas, cátedras), jornadas.	x
<b>Encuentros científico-profesionales</b>	Seminarios, mesas redonda y talleres	Seminario, mesa redonda y taller.	x
	<b>Propiedad industrial e intelectual</b>	Propiedad industrial	Patente, modelo de utilidad, topografía de productos semiconductores, diseño industrial, dibujos y modelos industriales.
Propiedad intelectual		Diseños arquitectónicos y de ingeniería, gráficos, mapas y diseños topográficos, Programas de cómputo, Bases de datos, Modelos de medición de intangibles.	

### **Barreras de investigación**

Las barreras de investigación identificadas para la obtención y difusión de la información en las diversas unidades de investigación, laboratorios, institutos, áreas de la FIA se concentran en dos grupos, barreras de difusión y barreras de almacenamiento y recuperación de la información como

se describe en la Tabla 3, donde se puede observar, en la descripción, que muchas de esas barreras se apoyan en herramientas existentes, actualmente, en la web. El problema es la no existencia de herramientas de apoyo, la no integración de esas herramientas o la estandarización del manejo de la información.

Tabla 3

*Barreras para la obtención y difusión de la información*

<b>Barreras de difusión</b>	
Comunicación por diversos correos electrónicos	Uso de mucha mensajería software, mensajería electrónica, lista de distribución, servicio de news, servicio de intercambio en tiempo real (Chat).
Uso compartido de la red	Uso de muchas bases de datos y recursos a través de la red.
Conjunto de páginas hipertextuales dispersas	Diversos formatos de los sitios, foros de discusión, blogs, wikis, portal, biblioteca digital, repositorios (artículos, tesis, informes, entrevistas, audiovisuales, contenido multimedia) usados.
Información en mundos virtuales sin integración	Diversos software que no interactúan entre sí, procesos, juegos, sitios entre otros.
<b>Barreras de almacenamiento y recuperación</b>	
Almacenamiento y recuperación de información con diversos procesos.	Muchos motores de búsqueda para muchos un mismo índice índices o directorios.

Estas barreras identificadas se apuntalan si observamos de manera muy sucinta la existencia de otros tipos de barreras como:

Barreras académicas, en donde el reglamento actual de la actividad docente no favorece que profesores e investigadores puedan obtener horas de investigación dedicada.

Barreras administrativas: todos los docentes son contratados por la Unidad de Servicios Académicos, a solicitud de la escuela.

Barreras económicas: las unidades de investigación no cuentan con presupuestos dedicados y no tienen mecanismos para conseguirlos, más allá de la buena voluntad de sus integrantes.

Barreras de la organización: no existe un centro único de investigación que articule los centros especializados de investigación.

Barreras de la cultura organizacional: se confunde el trabajo operativo con el trabajo e investigación.

**Estrategias frente las barreras de investigación**

Las estrategias de investigación nacen como respuesta a las barreras de obtención y difusión de la información como se muestra en las estrategias, en donde se plantea para cada barrera una estrategia de difusión y obtención de la información.



Tabla 4

*Barreras y estrategias*

Barreras	Estrategias
Comunicación por diversos correos electrónicos	Uso de componentes de Data Ingestión (DI) para facilitar la conversión de los datos de entrada en varios formatos en RDF. Ofrece diferentes plantillas y componentes para manejar varios formatos de datos comunes, tales como archivo de texto, bases de datos relacionales y hojas de Excel.
Uso compartido de la red	<p>Uso de Componente Gestión de Ontología (OM) cuya función principal es permitir la creación fácil de la ontología en línea, la edición, navegación, mapeo y anotación.</p> <p>Componentes para administración de sitios web corporativos.</p> <p>Organización taxonómica automatizada de datos y documentos.</p> <p>Construcción e implementación de agentes inteligentes.</p> <p>Administración de derechos y control de accesos a lugares privados.</p>
Conjunto de páginas hipertextuales	<p>El uso de componentes para la fácil navegación (FB) permitirá al usuario, explorar complejas estructuras de manera fácil y significativa gracias a la flexibilidad del modelo de datos RDF. Estos componentes permiten que el navegador pueda tener múltiples facetas, con filtro, por ejemplo, para un portal de grupo de investigación. Los usuarios pueden navegar por cualesquiera de todas las presentaciones existentes por un grupo de investigación o solo aquellos que dentro de un año específico y en una determinada ubicación; un portal del centro de salud, el médico puede conocer el número de pacientes que tienen diabetes y viven en el distrito tal de la provincia tal.</p> <p>Uso de Componentes de Visualización Semántica (SV): Permite mostrar o visualizar los datos de RDF, en la pantalla, la línea de tiempo como el mapa de Google y formatos de tabla. También permite la visualización de la faceta de modo que los usuarios puedan visualizar todos los miembros del grupo de investigación, o solo los miembros de los grupos que comparten intereses comunes de investigación.</p> <p>Uso de componentes de colecciones multimedia.</p>
Mundos virtuales	Uso de meta versos semánticos y ontología de los lenguajes virtuales (que se utilizan en Internet, tanto para su configuración técnica, como para su manejo) para integrar mediante agentes que interpretan máquinas e investigadores, indistintamente, en diversos grados y niveles que sirven de mediadores radiales de una determinada información. Su pretensión es representar un “metamundo” con un universo determinado, a su vez, por un “metalenguaje”. La artificialidad y artificiosidad de metamundos y metalenguajes, dan lugar a una teoría específica de significados, a una semántica virtual de un orden distinto a la de los lenguajes naturales que están acostumbrados (Moreno, 204).

<p>Barreas de almacenamiento y recuperación de información</p>	<p>Implementación de un Buscador Semántico (SS) integrado con ontologías pre-definidas del portal que proporcionará la base de las búsquedas. Por ejemplo, si los usuarios ingresan la palabra “web semántica” como consulta de búsqueda recibirán los recursos RDF, que contienen la cadena “web semántica”, en donde estos recursos se clasifican además como persona, proyecto, publicación, presentación, y el evento. Las subclases de un grupo de personas pueden clasificarse en personal académico o estudiante. El buscador puede ser desplegado en cualquier dominio o aplicación para entrada, salida, visualización y búsqueda de datos semánticos.</p> <p>Implementación de Reglas de Categorización para mejorar la búsqueda.</p>
--	---

### Metodología

Esta investigación constituye un estudio de caso porque se enfoca al análisis de procesos de investigación que se desarrollan en las unidades de investigación de la FIA, Ingeniería Electrónica, Computación y Sistemas, Industrial, Civil, Arquitectura e Industrias Alimentarias. Se pretende elaborar un análisis comprensivo de las acciones que constituye la realización de investigación. El propósito es documentar los procesos de investigación desde la visión y actuación de sus actores dentro del contexto en que se desenvuelven. La información recogida permitirá aportar explicaciones al proceso de investigación respecto al manejo de la información en las diferentes etapas del proceso de investigación.

La opción metodológica para recoger y sistematizar los datos fue cualitativa porque se ha recogido la información a base de testimonios, acciones y significados que los investigadores entrevistados otorgan al proceso de investigación en la FIA. El interés no es verificar relaciones causales entre variables o componentes que intervienen en el proceso, sino explicativo y comprensivo de esos procesos. Las estrategias de acopio de infor-

mación fueron la entrevista, observación etnográfica no participante y las historias académicas.

La entrevista realizada es no estandarizada, en que se pretende explorar, recoger y caracterizar el proceso por el que transita la información en sus distintos formatos, es decir, identificar los tipos de datos que los investigadores necesitan y las estructuras de almacenamiento que usan para que la data persista en la investigación en curso.

La observación que se realizó en los espacios de investigación fue del tipo “panorámica-no participante” (Ruiz e Ispizúa, p. 89). Por este método se registraron las formas de interacción entre investigadores, practicantes y tesis, apoyados con audio grabación.

### Implementación del portal la UI EPICS

Se analizaron los datos relacionados con los ítems anteriores y se realizó el cuadro de vinculación de estrategias de investigación a fin de que la implementación de un portal semántico esté fundamentada en la necesidad de información inteligente de las distintas unidades de investigación de la FIA.

Tabla 5

*Estrategias soportadas por el portal semántico*

Estrategias propuestas	Características de difusión y registro en los ranking que asegura el portal semántico
Crear componentes de Data Ingestión (DI) para facilitar la conversión de los datos de entrada en varios formatos en RDF.	Visibilidad Creación de contenidos de flujo constante y rápido. Uso de estándares para enriquecer los sitios web.
Crear Componente Gestión de Ontología (OM). Componentes para administración de sitios web corporativos.	Visibilidad Creación de contenidos de flujo constante y rápido. Registro e indexación de recursos.
Organización taxonómica automatizada de datos y documentos.	Interconexión Presentación de contenidos en múltiples lenguajes
Construcción e implementación de agentes inteligentes.	Presentación de ficheros ricos y multimedia
Administración de derechos y control de accesos a lugares privados	Archivado y persistencia
Construcción de componentes para fácil navegación (FB)	Diseños amigables para los motores de búsqueda Capacidad de popularidad y estadísticas
Construcción de Componentes de Visualización de Semántica (SV). Meta versos semánticos, ontología de los lenguajes virtuales (que se utilizan en Internet, tanto para su configuración técnica como para su manejo).	Agentes de computación ubicua.
Implementación de un Buscador Semántico (SS).	Diseños amigables de los motores de búsqueda.
Implementación de Reglas de Categorización para mejorar la búsqueda.	

La propuesta de implementación de un portal semántico de investigación se realiza con el fin de usar las estrategias propuestas en las barreras de investigación. El portal semántico por sus características descritas en la Tabla 5 permitirá el intercambio y colaboración entre los docentes, investigadores y

estudiantes a usar y a difundir investigación entre las comunidades que lo requieran.

Esta es una gran oportunidad para que el conocimiento que se produce en las distintas unidades de investigación de la FIA pueda presentarse a la comunidad científica y aca-

démica, de manera inmediata e ingresar a los parámetros de medición de instituciones que se ocupan de medir la producción de investigación y académica de las universidades.

Por tanto, se puede usar información que está almacenada en bibliotecas de documentos que permitirán ver los archivos, editarlos en cualquier momento y lugar desde su equipo. También se admiten las wikis, los blogs, listas de discusión, encuestas, calendarios listas de contactos, listas de discusiones, tareas y listas personalizadas, publicación de documentos, clasificación de información no estructurada tales como casos y soluciones, espacios para la colaboración y debate de un casuística de investigación, espacios personales, disponer de fácil acceso a un gran volumen de información proveniente de disímiles fuentes, el manejo de usuarios a partir de su rol, entre otras. Las búsquedas no se limitan a ubicar una palabra, puede buscarse todo un concepto y la herramienta responde

de manera muy rápida. La personalización es muy simple y se integra de manera completa a las herramientas de gestión de proyectos e indicadores de gestión.

La capacidad de crear ontologías de manera simple con lo cual se puede integrar en un único sitio, todas las actividades relacionadas con cada grupo de investigación e incluir las capacidades de búsqueda y recuperación de la información almacenada (desde los distintos sitios creados para respaldar unas conferencias, seminario, charla, entre otras actividades realizadas por los integrantes de investigación). En la Figura 12, se muestra cómo la estructura jerárquica de investigación dentro de las escuelas no impide la difusión de la información, en el nivel inferior se encuentran los sitios creados a partir de sitios de investigación que, a su vez, pueden tener sub-sitios orientados a problemas particulares de cada grupo de investigación.



Figura 12. Estructura del Portal de la Unidad de Investigación EPICS

Las variables que podemos destacar, en este prototipo de implementación de un portal llevado a cabo en la unidad de investigación de la Escuela de Ingeniería de Computación y Sistemas son las siguientes: ante las páginas medidoras de la producción académica que existen en la web.

### **Características de difusión de la producción escrita de investigación**

#### **a. Visibilidad**

El portal semántico de investigación permitió integrarse al único URL de la USMP, por tanto, la producción de investigación conllevó a una citación del sitio y contribuyó a la visibilidad de la universidad, pues las citaciones del sitio en artículos, contribuyen con un porcentaje sobre la acumulación de puntos de evaluación de las universidades pues la atención se centra en las instituciones que están aportando al conocimiento y la ciencia a escala mundial (QS, 2012).

#### **b. Creación de contenidos de flujo constante y rápido**

Con el aumento del número de autores activos en la producción de artículos y diversos escritos, producto de sus investigaciones y con la facilidad de inserción e indexación al mundo científico que asegura el portal semántico, se contribuye de manera que asegure una gran presencia en la Web. La mejor forma de conseguir presencia en la web es permitir a una gran proporción del personal, investigadores o estudiantes la posibilidad de convertirse en autores potenciales. Y el acceso a bibliotecas, centros de documentación y bibliográficas, tesis doctorales, pre-impresiones e informes, sitios web propios de cada individuo que participa en la investigación, conferencias vía web, los almacenes de programas, las sociedades

científicas y sus publicaciones, revistas electrónicas, entre otros, contribuirán con la visibilidad (CSIC, 2012).

#### **c. Registro e indexación de recursos**

Los recursos importantes que están disponibles en un formato no electrónico pueden ser convertidos a páginas web muy fácilmente. El portal lo registró como producción académica de manera inmediata.

#### **d. Interconexión**

El portal semántico asegura enlaces de manera incremental, basta crear la ontología que incluya el incremento de enlaces importantes ante una investigación para que la página Web de le USMP sume esos link y esto también es una medida por clasificar.

#### **e. Presentación de contenidos en múltiples lenguajes**

El portal semántico usa el estándar UNICODE, que es un estándar cuyo objetivo es proporcionar el medio por el cual un texto, en cualquier forma e idioma, pueda ser codificado para el uso informático. Por tanto, permite mostrar información en cualquier idioma y con la certeza de que no aparezcan símbolos extraños permitiendo que de manera automática la traducción de las páginas, artículos y otros documentos editables de conviertan, de manera automática, al idioma inglés, u otros idiomas.

Esto contribuyó con el parámetro de medición de los rankings de la audiencia web global.

#### **f. Uso de ficheros ricos y multimedia**

La web semántica permite el uso de documentos independientes de su forma-

to, de ingreso y salida gracias a los URI, que son cadenas que permiten acceder a cualquier recurso de la web. En la Web Semántica, las URI son las encargadas de reconocer objetos. Estos pueden ser identificados mediante una URI. Y sobre el URI corre el estándar XML (Extensible Markup Language) que ofrece un formato común para el intercambio de documentos, Namespaces (NS) proporciona un método para cualificar elementos y atributos de nombres usados en documentos XML asociándolos con espacios de nombre identificados por referencias URIs. XML Schema es un lenguaje que permite describir la estructura y restringir el contenido de documentos XML. Por tanto, permite formatos de archivos ricos como son el Adobe Acrobat pdf o el MS Word doc permitiendo una mejor distribución mejor de documentos y materiales multimedia como colecciones de videos, entrevistas, presentaciones, gráficos animados, e incluso fotos digitales que aseguran producción académica.

#### **g. Diseños amigables para los motores de búsqueda**

Los portales semánticos evitan el recargo excesivo de menús de navegación basados en Flash, Java o Java Script o de Directorios profundamente complejos que imposibiliten la búsqueda de los robots de los sitios de medición de producción web.

Los accesos a las bases de datos e incluso páginas muy dinámicas están asegurados en su acceso por la estructura semántica del portal.

#### **h. Capacidad de popularidad y estadísticas**

El portal mediante sus agentes captura no solo el número de visitas, sino que también permitirá monitorizar de dónde vienen, su distribución y la forma por las que llegan al sitio web. La mayoría

de los actuales analizadores de registros ofrecen una gran diversidad de tablas y gráficos que muestran los datos geográficos y demográficos relevantes, pero debe asegurarse de que existe una opción para mostrar a los referentes, las páginas web desde las cuales las visitas llegan, o el término o frase usada si la visita llegó desde un motor de búsqueda que permita llevar un ranking de las páginas más populares de investigación FIA y de directorios relevantes.

#### **i. Archivado y persistencia**

El uso de los metadatos permite mantener una copia de material viejo o desfasado en el sitio de manera obligatoria, esto gracias a la propiedad de los metadatos de describir, identificar y localizar contenidos en documentos de la web.

#### **j. Uso de estándares para enriquecer los sitios web**

El uso de títulos con significado y meta-etiquetas descriptivos incrementa la visibilidad de las páginas. La web semántica permite estándares que al ser utilizados añaden información de autoría, palabras clave y otros datos acerca de los sitios web.

XML: provee una sintaxis elemental para las estructuras de contenidos dentro de documentos.

XML Schema: es un lenguaje para proporcionar y restringir la estructura y el contenido de los elementos contenidos dentro de documentos XML.

RDF: es un lenguaje simple para expresar modelos de los datos, que refieren a los objetos “recursos” y de sus relaciones. Un modelo de RDF-based se puede representar en sintaxis de XML.



RDF Schema: es un vocabulario para describir propiedades y clases de recursos. RDF-based, con semántica para generalizar jerarquías de las propiedades y clases.

OWL: es un mecanismo para desarrollar temas o vocabularios específicos en los que podamos asociar esos recursos.

### Estructura lógica del portal

En su estructura lógica, este portal propuesto tiene las siguientes capas:

- a. de acceso a la información
- b. de procesamiento de información
- c. de integración de tecnologías

La capa de acceso a la información se encarga de la usabilidad y la capacidad del portal para la aceptación del usuario; la capa

de procesamiento de la información abarca la capacidad de procesamiento del conocimiento que comprende los del flujo de creación, publicación, organización al acceso de mantenimiento según se muestra en la semántica. La capa de integración de tecnología permite la integración con las características de las tecnologías que apoyarán a los niveles de las partes superiores.

Esta estructura de capas contribuye con la colaboración incluida, la evaluación, las características técnicas de elaboración y realización son inspeccionados para cada fase.

Es importante diferenciar las tecnologías de sistema de las tecnologías de la Web Semántica: la primera cubre los aspectos básicos de tecnologías portal web mientras que la segunda tiene que ver con las tecnologías que hacen que un portal de Internet convertirse en un Portal Web Semántico (Lausen et al., 2005).

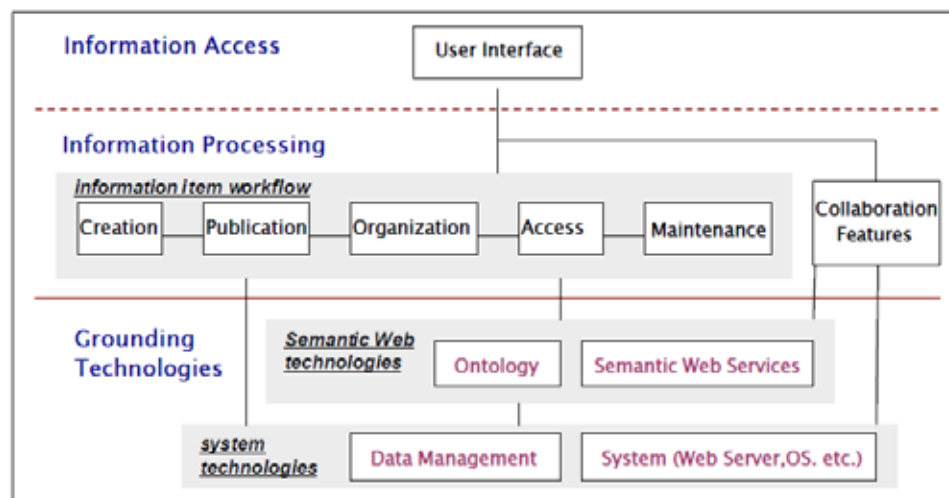


Figura 13. Capas de Portal semántico (Lausen et al., 2005)

El objetivo final de esta propuesta es que el portal web semántico ayude a la gestión de conocimientos, incluyendo la adquisición, representación, intercambio y la evolución del conocimiento de investigación a través de Internet. Lo que permite asegurar el registro y captura de la producción intelectual de las diversas unidades o centros de

investigación y ser objeto de captura de los centros de medición intelectual.

Estos conceptos están organizados en estructuras con formato (ontologías) que se mantienen en los portales web. Estas estructuras están abiertas al análisis de los trabajadores del conocimiento y de los agentes de

software. Por tanto, el portal web semántico proporciona un respaldo a la investigación como nueva clase de entorno de trabajo (Wang & Hai, 2011). El uso de los agentes en los portales semánticos asegura la inferencia. En la Figura 14, se puede observar que los agentes interactúan con los servicios web, la ontología y la meta data. La información o conocimiento ingresa al portal por la interface del portal. Este conocimiento se estandariza en la base semántica y es consumida por los servicios web semánticos, que entienden la ontología y alimentan la meta data y es entendida por los agentes quienes permiten devolver inferencia, sin importar los formatos en los cuales el conocimiento es almacenado.

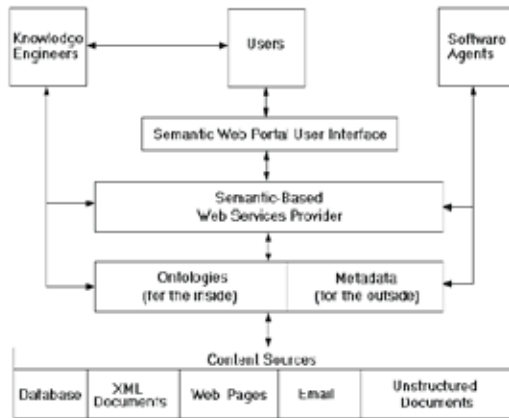


Figura 14. Semantic Web portal and knowledge Management (Wang & Hai, 2011)

Integrando el portal a la gran infraestructura en malla que significa la Internet, el portal es capaz de convertirse en un nodo de esa gran malla y su infraestructura debe ser capaz de soportar e integrar a la nueva plataforma en malla web; debe ser capaz de soportar transacciones e intercambio de información, los nuevos paradigmas para la Ingeniería de Software y Servicios, tales como el software de-asa-Service (SaaS) y el cloud computing, el intercambio de recursos de funcionalidad y la informática. Esta plataforma debe permitir a los desarrolladores externos crear aplicaciones adicionales (Webpart) que se integran en la aplicación principal y se alojarán en una infraestructura Cloud Computing (García-Sánchez et al., 2010).

En la Figura 15, se puede observar la estructura de implementación del portal semántico interactuando en la gran malla web, capaz de soportar la integración de datos, la colaboración y la futura inferencia de los mismos. Se pueden observar las capas de soporte de la infraestructura en la parte inferior, luego en la capa media, se muestra la integración de la data debidamente resguardada por las capas paralelas de gestión, virtualización y servicio (Haase, Matha, Schmidt, Eberhart, & Walther, 2010).

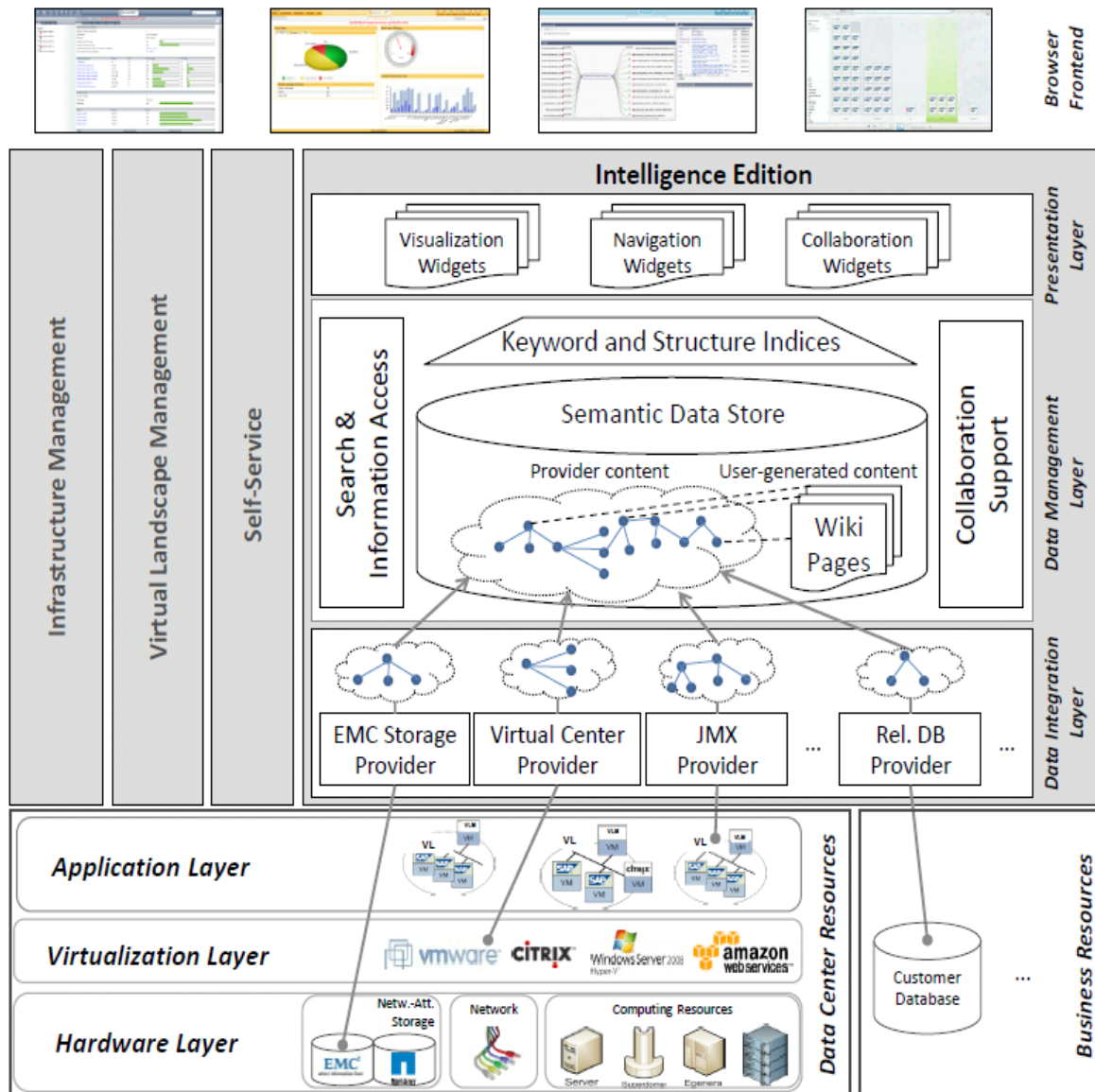


Figura 15. eCloudManagerArchitecture

Luego el acceso al portal tiene como acceso un portal similar al usado por el Gene Analysis Virtual Lab Experiment (Ahson & Mohammad, 2011) mostrado en la Figura 16.

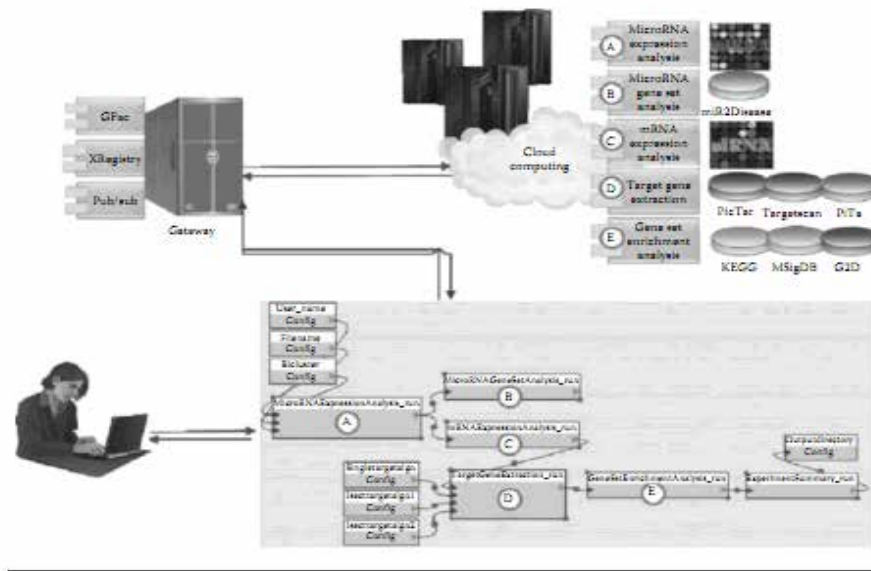


Figura 16. BioVLAB-microarrayarchitecture.

### Propuesta del modelo ontológico para la meta data del portal

El modelo ontológico de la meta data de investigación para que pueda ser implementada en el portal propuesto describe el Modelado de Procesos de Negocio Ontología (BPMO) que forma parte de un enfoque de modelado de los procesos de negocio a nivel semántico, la integración del conocimiento sobre el contexto de la organización, las actividades de flujo de trabajo y Servicios de la Web Semántica. Aprovechamos la representación del conocimiento y las técnicas de razonamiento para que los flujos de trabajo de procesos de negocio puedan ser expuestos y compartidos a través de descripciones semánticas.

### Conclusiones

Las principales barreras identificadas en la FIA para la implementación de un portal semántico se pueden convertir en estrategias de difusión y presencia de la universidad en la comunidad de investigación, la capacidad de las herramientas semánticas implementadas en un portal. Es una oportunidad para el incremento de nuestra producción intelectual no formal, que permitirá las mediciones y validaciones de los sitios medidores de Ranking de producción intelectual de centros de investigación al interior de una institución educativa superior.

Con esta investigación, se quiere proponer una alternativa usando tecnología para ascender en la percepción de producción intelectual, mientras se fortalece la organización formal de la investigación en la Universidad.

## Referencias

- Ahson, S.A. & Mohammad, I. (2011). *Cloud Computing and Software Services Theory and Techniques*. USA: Taylor and Francis Group.
- CSIC (2012). *Ranking Mundial de Universidades en la Web*. Recuperado de <http://www.webometrics.info/>
- Ding, Y., Sun, Y., Chen, B., Borner, K., Ding, L., Wild, D., ... Toma, I. (2010). *Semantic Web Portal: A Platform for better browsing and Visualizing Semantic Data*. Trabajo presentado en 6th International Conference on Active Media Technology, Toronto, Canada.
- Garcia-Sanchez, F., Fernández-Breis, E., Valencia-Garcia, R., Jiménez, E., Gómez, J., Torres-Niño, J., & Martínez-Maqueda, D. (2010). *WSEAS Transactions on Computer*, 9, 154-163.
- Haase, P., Matha, T., Schmidt, M.S., Eberhart, A., & Walther, U. (2010). *Semantic Technologies for Enterprise Cloud Management*. Trabajo presentado en 9th international semantic web conference, Shanghai, China.
- Hidayat, R., Yahya, Y., Azaman, S., & Noah, M. (2012). Semantic Web Portal in University Research Community Framework. *International Journal on Advanced Science, Engineering and Information Technology*, 2(6), 39-43.
- Kando, N., Kanazawa, T., & Miyazawa, A. (2006). *Retrieval of Web Resources Using a Fusion on Ontology-based and Content-based Retrieval with the RS Vector Space Model on a Portal for Japanese Universities and Academic Institutes*. Trabajo presentado en el 39th Annual Hawaii International Conference, Hawaii.
- Lausen, H, Ding, Y., Stollberg, M., Fensel, D., Lara, R., & Han, S.K. (2005). Semantic web portals: State-of-the-art survey. *Journal of Knowledge Management*, 9(5), 40-49.
- Maedche, A., Staab, S., Stojanovic, N., & Stojanovic, R. (2001). SEMantic portAL – The SEAL approach. En D. Fensel, J. Hendler, H. Lieberman, & W. Wahlster (Eds.), *Spinning the Semantic Web*. Massachusetts: MIT Press.
- Moreno, C. (2004). *Hacia una semántica Virtual*. Madrid, España.
- Newman, D.R., Bechhofer, S., & De Roure, D. (2009). *myExperiment: An ontology for e-Research*. Trabajo presentado en Semantic Web Applications in Scientific Discourse, Washington DC, US.
- Paradistis, S., Viegas, E., & Hey, T. (2009). Viewpoint a “Smart” Cyberinfrastructure for research: a view of semantic computing and its role in research. *Communications of the AMC*, 52(12), 33-37.
- QS (2012). *QS Latin University Rankings*. Recuperado de <http://www.topuniversities.com/university-rankings/latin-american-university-rankings/2012>
- Rosado, M., Román, A., Sanz, E., Berges, M., Gómez, A., García, F. y Aguillo, I. (2008). *Criterios de clasificación de los medios de difusión de la producción académica y científica universitaria* (Informe). Recuperado del sitio de Internet Agencia de Calidad, Acreditación y Prospectiva de las Universidades de Madrid: [http://www.aqu.cat/doc/doc\\_25874143\\_1.pdf](http://www.aqu.cat/doc/doc_25874143_1.pdf)

- Scientometrics (2012). *Ranking web de mejores universidades del mundo*. Recuperado de [http://webometrics.info/about\\_rank\\_es.html](http://webometrics.info/about_rank_es.html)
- Suaza, L. y Ortega, O. (2007). *Un camino a la investigación en ingeniería* [Entrada de blog] Recuperado de <http://minga.udea.edu.co/~oortega/researchhints/guiaelaboracionpropuesta/>
- Wang, H., Gibbins, N., Payne, T. Saleh, A., & Sun, J. (julio, 2007). *A Formal Semantic Model of the Semantic Web Service Ontology (WSMO)*. Trabajo presentado en la 12th IEEE International Conference, Auckland, New Zelanda.
- Wang, H., & Hai, W. (2011). Semantic Web Portal. En A. Tatnall (Ed.), *Encyclopedia of Portal Technologies and Applications: Vol. 2* (pp. 901-904). USA: Yurchak printing.