

Uso de un servicio web semántico para apoyar la gestión de la información aplicado a los procesos de selección de estudiantes en una institución de educación superior

Use of a semantic web service to support the information management applied to the selection process of students in a Higher Education Institution

Alexander Guerrero A.
Universidad Francisco de Paula Santander
Correo: aguerreroav@ufpso.edu.co

Sandra C. Sanguino
Universidad Autónoma de Bucaramanga
Correo: ssanguin@unab.edu.co

Fecha de recibido: 10/06/2016 y Fecha de aprobación: 21/07/2016

Resumen

La integración de la web semántica y los servicios web poseen el potencial para darle a la web actual la infraestructura para construir los servicios, las herramientas y aplicaciones que les permitan a los computadores interactuar de manera inteligente y transformar la información en conocimiento. El proyecto de investigación propuesto consiste en aplicar la construcción de la ontología (modelo conceptual explícito y formal) enfocada en procesos que permitirán además de estandarizar y clasificar la información, ser replicada en otras áreas de la organización a medida que el conocimiento sea asimilado. Este trabajo presenta una visión general de las metodologías para usar ontologías y presenta un marco de trabajo enfocado en el desarrollo de un prototipo de un servicio web semántico que se integra con un repositorio de documentos relacionados en el proceso de admisión de estudiantes de una universidad, que además de realizar búsquedas de información, el resultado es clasificado y relacionado según el dominio especificado en la ontología, dando el punto de partida para implementar una solución como es la gestión de la información.

Palabras Clave: *Ontología, Web Semántica, Servicio Web, Recuperación de información, Administración de información.*

Abstract

The integration of the Semantic Web and Web services have the potential to give to the current web, the infrastructure to build services, tools and applications that allow computers to interact intelligently and transform information into knowledge. The proposed investigation project consists in applying the construction of ontology (formal and explicit conceptual model) focused on processes that will further standardize and classify information, be replicated in other areas of the organization as knowledge is assimilated. This work presents a general vision about the methodologies in order to use ontologies and show a frame of working focused on the development of a prototype of a semantic web service that integrates with a repository of documents related to the admission process of students from a university, in addition to information searches, the result is classified and related to the specified domain in the ontology, providing the starting point to implement a solution such as the management of information.

Keywords: *Ontology, Semantic Web, Web Service, Recovery of information, Management of Information.*

1. Introducción

Desde la aparición de la Web ha incrementado la cantidad de documentos y por ende los tiempos para la recuperación y clasificación de la información [1]. La arquitectura de la información aumentó su relevancia y surgió la necesidad de poder extraer la información útil y de manera eficiente [2]. Con el apoyo de la tecnología se busca recuperar los enlaces a las páginas web que tengan significancia, transformando así la web en un espacio de conocimiento que proporcione información relevante sobre un tema en particular sin intervención humana [3] (ver Figura 1).

Inicialmente, para la recuperación de información los motores de búsqueda usaron sistemas basados en índices temáticos, que luego evolucionaron para integrar el uso de palabras claves. Por ejemplo, el buscador de Google usa la técnica de relevancia de contenidos, esta última por los resultados en sus búsquedas cuenta con la aceptación de los usuarios; aun así, para mejorar el resultado de las búsquedas se han realizado múltiples propuestas para mejorar la interoperabilidad, la sobrecarga por los millones de sitios de información y la heterogeneidad de las fuentes en los resultados, búsqueda de la que hoy carecemos y que con la ayuda de tecnologías y conceptos como la web semántica trabajando de manera integrada pueden ser explotados [4]. Se introduce entonces una descripción detallada de un marco para permitir la interacción inteligente de los servicios web y la web semántica.

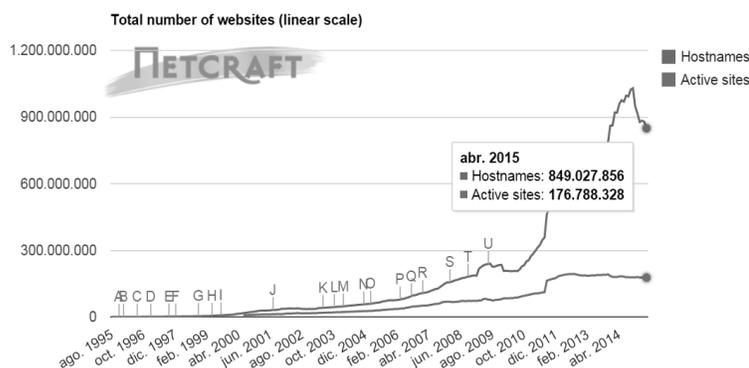


Figura 1. Número de sitios web en el mundo, tomada de www.netcraft.com, consultado Junio 2015.

Los medios digitales han creado nuevos escenarios y retos, en especial para determinar la dependencia y las relaciones de los objetos digitales, lo que ha dado pie para la adopción de lenguajes de modelado, beneficios gracias a la web semántica [5], motores Web de consulta, la integración de la información de múltiples fuentes, la localización y la integración de la información [3].

La conservación o preservación de la información es otro punto de importancia, pero de nada nos sirve si no podemos transmitirla con el objetivo de retroalimentar, gestionar y perfeccionar los conocimientos de la organización.

A. La Gestión de Conocimiento

La Gestión del conocimiento en la actualidad es fuertemente relacionada con el ámbito empresarial [6] y con la unión de esfuerzos para reunir la información de cada uno de los procesos que posee la organización tanto dentro de ella, como en su relación con el entorno. Las destrezas, experiencias y habilidades ofrecen también fuentes adicionales de conocimiento que son relevantes y que son conocidas como gestión del conocimiento empresarial [7], pero si no se cuenta con las herramientas necesarias para el manejo de toda esta documentación recolectada será difícil de asimilar.

Para suplir esta necesidad se han aplicado nuevas tecnologías que reducen la complejidad del desarrollo de esas herramientas y que ofrecen las características fundamentales que se persiguen en las aplicaciones empresariales [8], como son:

- Seguridad y fiabilidad.
- Adaptación al tipo y número de usuarios a los que irá dirigida la aplicación.
- Componentes que ofrezcan la posibilidad de reutilización.
- Mantenimiento sencillo.
- Almacenamiento de la información eficaz, y preferiblemente un modelo de datos unificado en una sola base de datos.
- Inteligencia en los procesos de negocio.
- Tiempo mínimo de inactividad de las aplicaciones.
- Alto rendimiento en las aplicaciones en la red (minimizar la latencia de la WAN, la gran carga de los servidores y la gestión de las conexiones con poco ancho de banda).
- Interoperabilidad con otras aplicaciones.

La gestión de conocimiento y su importancia en la nueva estructura empresarial ofrece la posibilidad de encontrar las fuentes del conocimiento que son relevantes para el problema actual y fortalece el pilar de la innovación, esencial en estos tiempos para lograr la eficiencia de los recursos, ya que si el conocimiento se transmite de manera rápida y eficaz, generará nueva información [7] y por consiguiente la posibilidad de generar nuevo conocimiento, que en un futuro transformarán y mejorarán los procesos, productos, servicios; *la organización*.

B. La Ontología en la Gestión de Conocimiento

La ontología es una de las herramientas que permite definir conceptos y relaciones de un dominio específico de forma compartida y expresados en un lenguaje entendible para las computadoras, el cual hace posible el modelado de conocimiento, pudiendo así extraer la información para ser agrupada y analizada de la mejor manera, siendo este el punto inicial en la interpretación del conocimiento y todo gracias a que las expresiones y relaciones entre los elementos se realizan de forma estructurada.

En sus inicios el estudio de las ontologías por múltiples disciplinas provocó ambigüedad en la especificación de sus componentes y que no se definan con claridad cada uno de sus elementos [9], es entonces que Thomas Gruber en su publicación "*It Is What It Does: The Pragmatics of Ontology (2003)*" decide realizar un consenso y estandarización de los componentes sobre la representación y composición de las ontologías, los cuales se explican a continuación [10]:

- Conceptos: Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc. Las clases en una ontología se suelen organizar en taxonomías a las que se les pueden aplicar los mecanismos de herencia.
- Relaciones: Representa la interacción y enlace entre los conceptos del dominio. Suelen formar la taxonomía del dominio.
- Funciones: Son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de la ontología.
- Instancias: Se utilizan para representar objetos determinados de un concepto.
- Axiomas: son teoremas que se declaran sobre relaciones que deben cumplir los elementos de la ontología.

En la ontología cada componente tiene un alcance en la gestión del conocimiento, los conceptos, las instancias y las relaciones pueden representar el conocimiento tácito de los integrantes de una organización. Las funciones son muy usadas para describir los pasos y para desarrollar un proceso. Los axiomas, permiten hacer inferencias, lo que es de gran utilidad para la toma de decisiones.

La web semántica como idea general propuesta por Tim Berners-Lee creador de la WWW, propuso lo que sería una nueva red donde los documentos estarían dotados de relaciones para enriquecer su contenido [1] y proporcionar una herramienta para facilitar la búsqueda inteligente [4].

C. Los Servicios Web

Los servicios Web (*Web Services*) son componentes software que permiten que varias aplicaciones se comuniquen con el fin de intercambiar información [8].

Para llevar a cabo esta tarea se hace uso de protocolos estándares de Internet. En particular se utilizan principalmente tres protocolos ():

- WSDL (protocolo XML utilizado para la descripción de los servicios Web)
- SOAP (protocolo basado en XML empleado para comunicar clientes y servidores que exponen servicios Web).
- UDDI (protocolo de descubrimiento de servicios Web).

Los servicios Web ofrecen a la Web tradicional una plataforma más apropiada para la integración de información, debido a que los datos tienen significado bien definido, el software, en lugar de los seres humanos, puede utilizar la gran variedad de fuentes y generar información, para mejorar la adquisición, el almacenamiento y recuperación de conocimientos de la organización de manera muy significativa [7].

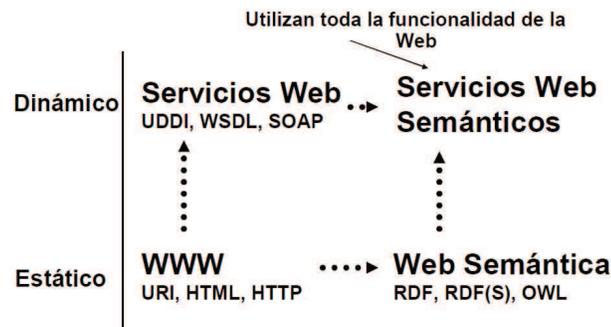


Figura 2. Visión de las tecnologías participantes en los SWS, consultado Junio 2015.

D. La Web Semántica

“La web semántica no constituye una web independiente, sino una ampliación de la actual; en la cual la información está dotada de significados bien definidos, con el fin de permitir un mejor trabajo en cooperación entre humanos y computadoras” [1].

Según la definición del W3C, “la Web Semántica es una Web extendida, dotada de mayor significado en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida” [11].

La web semántica en el futuro además de ayudar en la estandarización [1], [12] de las fuentes de información también disminuye la sobrecarga de información, ayudando en el proceso de la búsqueda por medio de la combinación de deducciones lógicas, aportando de esta manera el punto inicial para la creación de la base de conocimiento.

E. Orígenes y Evolución de la Web Semántica

Analizando los orígenes de la web semántica en todo su proceso de construcción [13], los métodos o ciclos de vida evolutivos se han basado en transformaciones sucesivas [14] para construir una ontología formal para un sistema que no siempre supone una mejora y que en muchas ocasiones es suficiente con una representación menos compleja [15]”genre”.”Journal article (Print/Paginated (ver Figura 3).

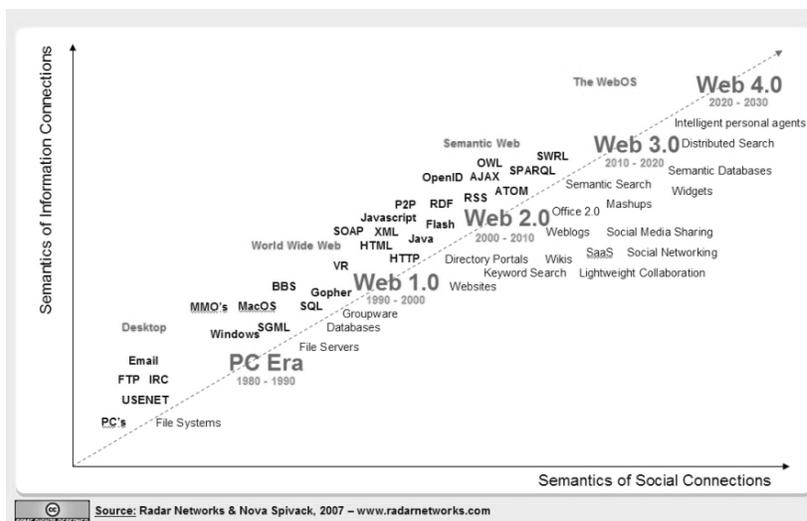


Figura 3. Evolución de la Web, tomada de www.radarnetworks.com, consultado Junio 2015.

Esta evolución ha permitido contribuir al proceso e implementación de aplicaciones informáticas facilitando la caracterización por medio de componentes, proporcionando a los usuarios determinar el dominio de una ontología, su estructura jerárquica, así como visualizar su datos estadísticos y propiedades.

F. Usos de la Web Semántica

Los usos de la web semántica son muy diversos, en investigaciones médicas como apoyo en los aspectos clínicos de la malaria [16], para evaluar las definiciones lógicas que se han desarrollado para la Ontología de mamíferos [17], integración de aplicaciones y servicios institucionales de educación superior mediante SOA [18], prototipos para mapear, interconectar y visualizar ontologías [19], STAGE un agente semántico para representar la información asociada del egresado [20].

SABIOS [21] es una aplicación web donde su objetivo es convertir la información en conocimiento teniendo como base la representación, almacenamiento y organización para tener éxito en la recuperación de la información.

Son tan amplias las posibilidades de integración de la web semántica que ya está siendo combinados con algoritmos de búsquedas para apoyar la recuperación de información, especialmente aplicando técnicas en la transformación del lenguaje natural [22], o la combinación de algoritmos con la retroalimentación del usuario según sus intereses [23] para obtener un mejor resultado en la búsqueda.

El uso de la web semántica para mejorar y mantener la disponibilidad de la información es cada vez mayor, su fácil adaptación con otras tecnologías emergentes y otras ya fortalecidas, proponen nuevos retos y nuevas posibilidades de integración [18].

2. Metodología

Para el diseño y construcción de la ontología, siendo algunas de las más relevantes para el presente trabajo se analizaron las siguientes metodologías [24]:

- Guía para crear ontologías – UNIVERSIDAD DE STANFORD: Metodología simple para la ingeniería del conocimiento aplicada a la creación de ontologías.
- METHONTOLOGY: Metodología desarrollada en el laboratorio de Inteligencia Artificial de la Universidad Politécnica de Madrid. Permite la construcción de ontologías de conocimiento e incluye la identificación del proceso de desarrollo de la ontología, un ciclo de vida basado en la evolución de prototipos y técnicas particulares a usar en cada paso.
- TOVE: Se basa en la experiencia de desarrollar la ontología que lleva su mismo nombre; se utilizó para con-

- construir la ontología TOVE (TOronto Virtual Enterprise).
- ENTERPRISE: Este esquema metodológico constituye la base de muchos de los métodos propuestos y usados en la actualidad. El esquema está constituido por cuatro pasos y considera además un conjunto de guías o recomendaciones de diseño que se deben tener presentes en cada paso del método.

Para el desarrollo del prototipo del servicio web semántico fue diseñado un método que tiene como finalidad realizar las búsquedas de documentos en un repositorio donde previamente se han catalogado y archivado; además el servicio web semántico usará una ontología también previamente diseñada para relacionar y clasificar los documentos recuperados en la búsqueda.

El servicio web integrará un repositorio y una ontología. Este componente será desarrollado en PHP y será encargado de recuperar y mostrar los enlaces para acceder a la información recuperada.

Para el diseño del prototipo se propone un modelo de desarrollo donde se evidencia por separado cada una de las fases del desarrollo del servicio web semántico (ver Figura 4).

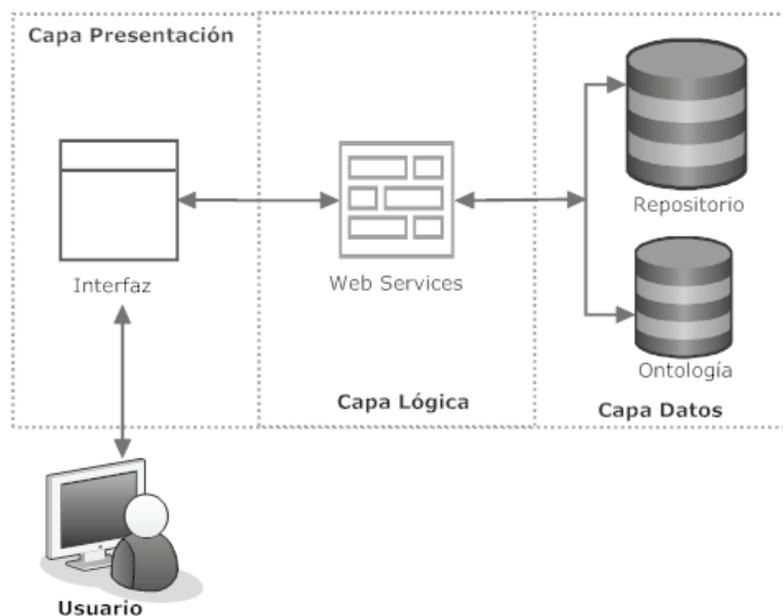


Figura 4. Arquitectura de capas del prototipo, Elaboración propia, Junio 2015.

3. Método propuesto para el desarrollo de servicios Web semántico

Para la ejecución del proyecto se propuso un método que concentra y hace énfasis en las fases representativas simplificándolas para ser entendido por los desarrolladores (ver Figura 5):

- 1) Revisión de Literatura: Durante esta fase se realizó una búsqueda de literatura al igual que proyectos relacionados con la recuperación de información, servicios web y web semántica que soportan la investigación.
- 2) Construcción de la ontología: Durante esta actividad se evaluaron los diferentes tipos de ontologías, los procesos de diseño y las metodologías para la construcción de ontologías, además fue recolectada la información necesaria para la definición de las clases, relaciones y características de la ontología requerida para el proceso de selección del área de Registro Académico.
- 3) Instalación del repositorio: Esta actividad tuvo como propósito preparar el servidor y realizar la instalación del software. Adicionalmente, se analizaron y seleccionaron las aplicaciones que permitieron la catalogación y archivado de archivos.
- 4) Desarrollo del Servicio Web Semántico: En esta fase del proyecto se definieron los componentes, módulos y servicios requeridos para el desarrollo del servicio web semántico.
- 5) Prueba del Servicio Web Semántico: En esta etapa se analizaron y seleccionaron las aplicaciones que permitieron

diseñar ontologías. Así mismo, se diseñaron las funcionalidades del sistema y se efectuó la carga de la información de muestra y la prueba funcional del prototipo.

- 6) Evaluación del Servicio Web Semántico: Durante esta actividad se analizaron los datos recogidos, el impacto en el proceso de selección de estudiantes y, se realizaron comparaciones con trabajos relacionados de la literatura.

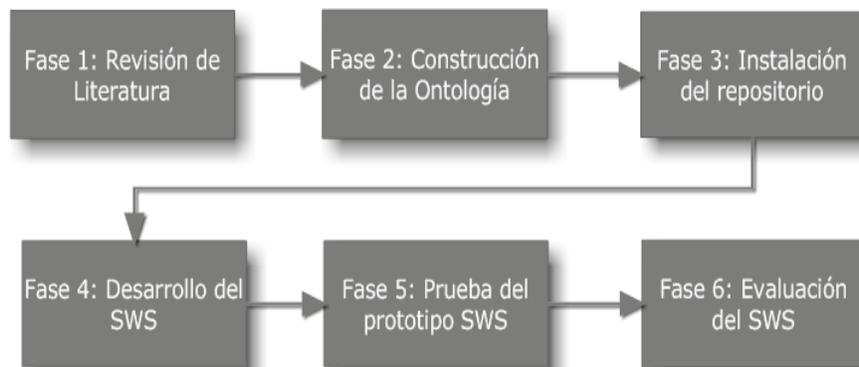


Figura 5. Fases de la Metodología, Elaboración propia, Junio 2015.

Como instrumento de medición, fue utilizado un formato con el registro de las peticiones, donde además se realizaron las observaciones pertinentes a medida que se fueron presentando, para la realización y puesta en marcha del prototipo. Los equipos usados para este proyecto están conectados a la red mediante la infraestructura de la organización con el fin de realizar mediciones en un entorno de tráfico de red real.

Para la recopilación de datos y la medición de los tiempos en la recuperación de información del servicio web se implementó una función dentro del software que se encarga de ejecutar y almacenar los tiempos de respuesta cuando el usuario realiza la solicitud.

Las siguientes variables basadas en el resultado de recuperación de información y en la satisfacción del usuario se usaron para evaluar el prototipo y emitir una evaluación del rendimiento:

- 1) Precisión: Porcentaje de documentos recuperados que resultan relevantes.
- 2) Exhaustividad: Ratio entre el número de documentos relevantes que el usuario espera encontrar y el número de documentos examinados en un intento de encontrar esos documentos relevantes.
- 3) Medidas promedio de la satisfacción del usuario: Medidas que pretenden medir la reacción de los usuarios ante el resultado de una búsqueda.
- 4) La tasa de fallo: Refleja el porcentaje de documentos recuperados no relevantes sobre el total de documentos no relevantes de la base de datos.
- 5) Promedio efectividad E-P: Medida de la capacidad del sistema para satisfacer al usuario en términos de la relevancia de los documentos recuperados.

El análisis de datos busca establecer gráficas comparativas de los tiempos de recuperación de información y el proceso actual de recuperación de información, comparando a partir del proceso de selección de estudiantes del área de Registro Académico.

4. Resultados

Se diseñó una ontología del proceso de selección de estudiantes que permite catalogar, y distribuir la información relacionada permitiendo estandarizar la terminología (ver Figura 6).

apoyo en la recuperación de información mejoró el desempeño de los usuarios del área en las actividades relacionadas con la admisión.

Con el análisis de los tiempos reflejados en las búsquedas y en las mediciones de satisfacción de los usuarios, producto del resultado aplicado a un escenario real, permiten avanzar como un trabajo futuro en el diseño de ontologías orientado a los procesos y se escale a todas las áreas y procesos de la organización.

Con el uso de la web semántica se ha potenciado a la web actual para ofrecer mejores resultados producto de la clasificación y estandarización de la terminología, pero el camino hasta ahora está comenzando, ya que solo es posible acceder a la información de dominios previamente elaborados.

Desde el origen del término de web semántica se han realizado proyectos enfocados a dotar de conocimiento la web actual, aunque con dificultad se ha tenido éxito en ambientes controlados; hasta el momento ninguno de ellos ha sido llevado al plano empresarial o comercial.

En este proyecto, con la integración de tecnologías para catalogar y clasificar la información, se busca mejorar el rendimiento de los procesos de la organización y mostrar un avance en la gestión de la información por intermedio de un servicio web semántico, proponiendo un punto de partida para, en el futuro, contar con un sistema que permita transmitir y gestionar conocimiento.

Referencias

- [1] T. BERNERS-LEE, J. HENDLER, and O. LASSILA, "The Semantic Web, A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities," *Scientific American*, vol. 284, n.º 5, pp. 34–43, May-2001.
- [2] R. García, F. Botella, and M.-C. Marcos, "Hacia la arquitectura de la información 3.0: pasado, presente y futuro. (Spanish)," *El Prof. Inf.*, vol. 19, n.º 4, pp. 339–347, Jul. 2010.
- [3] I. HORROCKS, "Ontologies and the Semantic Web," *Communications of the ACM*, vol. 58, 2008.
- [4] V. Dhingra and K. K. Bhatia, "Towards Intelligent Information Retrieval on Web." [Online]. Available: <http://cite-seerx.ist.psu.edu/viewdoc/download?doi=10.1.1.301.7956&rep=rep1&type=pdf>. [Accessed: 10-Mar-2015].
- [5] Y. Marketakis and Y. Tzitzikas, "Dependency management for digital preservation using semantic web technologies," *Int. J. Digit. Libr.*, vol. 10, n.º 4, pp. 159–177, May 2010.
- [6] E. G. Pemberty and E. R. Pineda, "Soluciones organizacionales a partir de ontologías," *Av. En Sist. E Informática*, vol. 8, n.º 1, p. 11–, 2011.
- [7] A. P. Soltero, M. B. Valenzuela, and G. G. A. S. Schmitz, "La web semántica como apoyo a la gestión del conocimiento y al modelo organizacional," *Ing. Informática*, n.º 12, p. 4, 2006.
- [8] P. Valledor Pellicer, "Servicios Web Semánticos," Universidad de Oviedo, 2006.
- [9] T. Gruber, "It Is What It Does: The Pragmatics of Ontology," Washington, D.C, Mar-2003.
- [10] R. Kishore and R. Sharman, "Computational Ontologies and Information Systems I: Foundations," *Communications of the Association for Information Systems*, vol. 14, 2004.
- [11] "Guía Breve de Web Semántica." [Online]. Available: <http://www.w3c.es/Divulgacion/GuiasBreves/WebSemantica>. [Accessed: 10-Mar-2015].
- [12] M. C. Périssé and J. Eterovic, "Gestión del conocimiento en una Web Semántica segura para la educación superior," presented at the XIV Congreso Argentino de Ciencias de la Computación, 2008.
- [13] R. Guns, "Tracing the origins of the semantic web," *J. Am. Soc. Inf. Sci. Technol.*, vol. 64, n.º 10, pp. 2173–2181,

2013.

- [14] J. R. Hilera, C. Pagés, J. J. Martínez, J. A. Gutiérrez, and L. de-Marcos, "An Evolutive Process to Convert Glossaries into Ontologies," *Inf. Technol. Libr.*, vol. 29, n.º 4, pp. 195–204, Dec. 2010.
- [15] S. Sanchez-Cuadrado, J. Morato-Lara, V. Palacios-Madrid, J. Llorens-Morillo, and J.-A. Moreiro-González, "De repente, ¿todos hablamos de ontologías?," *El profesional de la información*, 11-Jun-2012. [Online]. Available: <http://eprints.rclis.org/17150/>. [Accessed: 10-Mar-2015].
- [16] R. García, F. Botella, and M.-C. Marcos, "Hacia la arquitectura de la información 3.0: pasado, presente y futuro. (Spanish)," *Inf. Archit. 30 Past Present Future Engl.*, vol. 19, n.º 4, pp. 339–347, Jul. 2010.
- [17] D. Gutierrez González, D. G. González, and E. R. Ricardo, "Aplicación informática para caracterizar ontologías," *Ser. Científica*, vol. 6, n.º 5, May 2013.
- [18] I. HORROCKS, "Ontologies and the Semantic Web.," *Commun. ACM*, vol. 51, n.º 12, pp. 58–67, Dec. 2008.
- [19] B. Heitmann, S. Kinsella, C. Hayes, and S. Decker, *Implementing Semantic Web applications: reference architecture and challenges*. 2009.
- [20] T. Vacharasintopchai, W. Barry, V. Wuwongse, and W. Kanok-Nukulchai, "Semantic Web Services Framework for Computational Mechanics.," *J. Comput. Civ. Eng.*, vol. 21, n.º 2, pp. 65–77, Apr. 2007.
- [21] Y. Marketakis and Y. Tzitzikas, "Dependency management for digital preservation using semantic web technologies.," *Int. J. Digit. Libr.*, vol. 10, n.º 4, pp. 159–177, Dec. 2009.
- [22] E. G. Pemberty and E. R. Pineda, "Soluciones organizacionales a partir de ontologías," *Av. En Sist. E Informática*, vol. 8, n.º 1, p. 11–, 2011.
- [23] "La Web Semántica como apoyo a la Gestión del Conocimiento y al Modelado Organizacional." 2006.
- [24] F. J. García Peñalvo, "Web Semántica y Ontologías," Universidad de Salamanca, 2005.

Sobre los Autores

Alexander Guerrero Avendaño. Docente investigador del Grupo de investigación Rotã de la Universidad del Cauca. Popayán, Colombia.

Sandra Cristina Sanguino Galvis. Docente del programa de Ingeniería de Sistemas de la Universidad Autónoma de Bucaramanga. Colombia.

Este artículo se cita:

- IEEE A. Guerrero A and S. C. Sanguino, "Uso de un servicio web semántico para apoyar la gestión de la información aplicado a los procesos de selección de estudiantes en una institución de educación superior," *Revista Colombiana de Computación*, vol. 17, pp. 42-60, 2016
- APA Guerrero A, A., & Sanguino, S. C. (2016). Uso de un servicio web semántico para apoyar la gestión de la información aplicado a los procesos de selección de estudiantes en una institución de educación superior. *Revista Colombiana de Computación*, 17(2), 42-60.