



MODELO PROCESAMIENTO SEMÁNTICO EN SISTEMAS DE RECUPERACIÓN DE INFORMACIÓN

Temática : El mobil learning y la educación virtual ubicua.

MSc. Hubert Viltres Sala ^{1*}, Ing. Paúl Rodríguez Leyva², MSc. Aneyty Martín García ³,
Ing. Lenny Amel Pons Flores⁴

¹Universidad de las Ciencias Informáticas, Cuba, hviltres@uci.cu

²Universidad de las Ciencias Informáticas, Cuba

³Universidad de las Ciencias Informáticas, Cuba

⁴Universidad de las Ciencias Informáticas, Cuba

Resumen

El avance de las tecnologías para el empoderamiento y la participación han revolucionado la forma de impartir las clases. El surgimiento y evolución de la enseñanza en línea donde el estudiante se convierte en el principal protagonista plantea la necesidad de ayudarlos a encontrar la información que le permita adquirir más fácil los conocimientos. En este contexto la web semántica genera gran perspectivas en el sector de la educación ya que aportar significado a la información disponibles en los repositorios institucionales y en Internet. La presente investigación propone un modelo para desarrollar sistemas de recuperación de información con anotación semántica que permiten obtener información en la web con una alta calidad y valor para potenciar la educación. En la investigación se describen los componentes del modelo y su impacto en la educación.

Palabras clave: Recuperación de información, web semántica, educación

Introducción

El avance de Internet y el surgimiento de las TEP han permitido que se desarrollen herramientas para mejorar el acceso al gran volumen de información disponible en la web. En la actualidad existen más de 4.5

millones de usuarios de internet que acceden a más de un billón de sitios web buscando información que se ajuste a sus necesidades. Demasiada información limita nuestra capacidad de procesar e identificar la de mayor calidad y relevancia para suplir nuestra necesidad. Para acceder a esta información los usuarios utilizan diferentes herramientas que les ayudan a disminuir el tiempo de búsqueda y les muestra información que en ocasiones no satisfacen sus necesidades debido a que no logran comprender la necesidad real del usuario cuando realiza una consulta en lenguaje natural. Un usuario cuando realiza una búsqueda de información pregunta en lenguaje natural lo que necesita y espera obtener información útil. Para lograr esto los sistemas de recuperación de información emplean técnicas de procesamiento semántico de información que permiten analizar el contexto de la pregunta y asociarlo al documento que más semejanza y relevancia tenga para el usuario. La recuperación de información con anotación semántica pretende entender la intención de búsqueda del usuario para ofrecerle información relevante según el significado contextual de los documentos, mejorando la precisión y exhaustividad en la recuperación de información.

La Web Semántica está cambiando la forma de obtener información en internet, es una de las tecnologías que más impacto ha generado para los usuarios de internet por la calidad de la información que obtiene. Según Berners-Lee (2011) la Web Semántica es "...una extensión de la Web actual, en la cual la información tiene un significado bien definido, facilitando a las computadoras trabajar mejor en cooperación con los humanos" y su objetivo principal ha sido permitir que los datos almacenados en la Web puedan ser procesados por las máquinas de manera inteligente, facilitando a las personas la búsqueda, integración y análisis de la información disponible. La web semántica tiene como principio el procesamiento de información de forma automática mediante la utilización de inteligencia artificial utilizando una gran variedad de algoritmos. Pretende además comprender la necesidad expresada por el usuario en una consulta realizada y dotar la búsqueda de un significado, identificando y brindando información confiable.

La web semántica es una tecnología que por sus potencialidades se esta



empleando en la educación con un amplio impacto en la forma de aprender adquirir el conocimiento. Según plantea Martínez (2011) satisfacer las necesidades de información no es suficiente garantía de éxito en la actualidad, sobre todo en entornos profesionales donde se maneja mucha información y hay que saber distinguir la que es relevante. En este contexto la recuperación de información ayuda a los usuarios a localizar, sintetizar y asimilar información relevante y ajustada a sus necesidades permitiéndoles generar conocimiento.

Con el objetivo de recuperar información relevante para los usuarios se plantea el desarrollo de un buscador semántico que “entienda la necesidad el usuario y analice la información disponibles en la Web mediante el la utilización de algoritmos que simulan comprensión o entendimiento”. La propuesta está compuesta por tres componentes (Rastreo-Indexación, Procesamiento y Presentación de la Información) que permitirán identificar la necesidad de información del usuario mediante el procesamiento, selección y posterior presentación de la información recuperada. A continuación se describen cada uno de los tres componentes (ver figura 1).

Rastreo-Indexación

Permite recopilar las páginas durante el proceso de rastreo, creando un índice con los datos de las páginas indexadas, se emplean técnicas para realizar la extracción semántica de información. Además se estandariza la información para su posterior procesamiento. Se emplea el procesamiento en lenguaje natural y una ontología de dominio específico para establecer las relaciones entre los elementos.

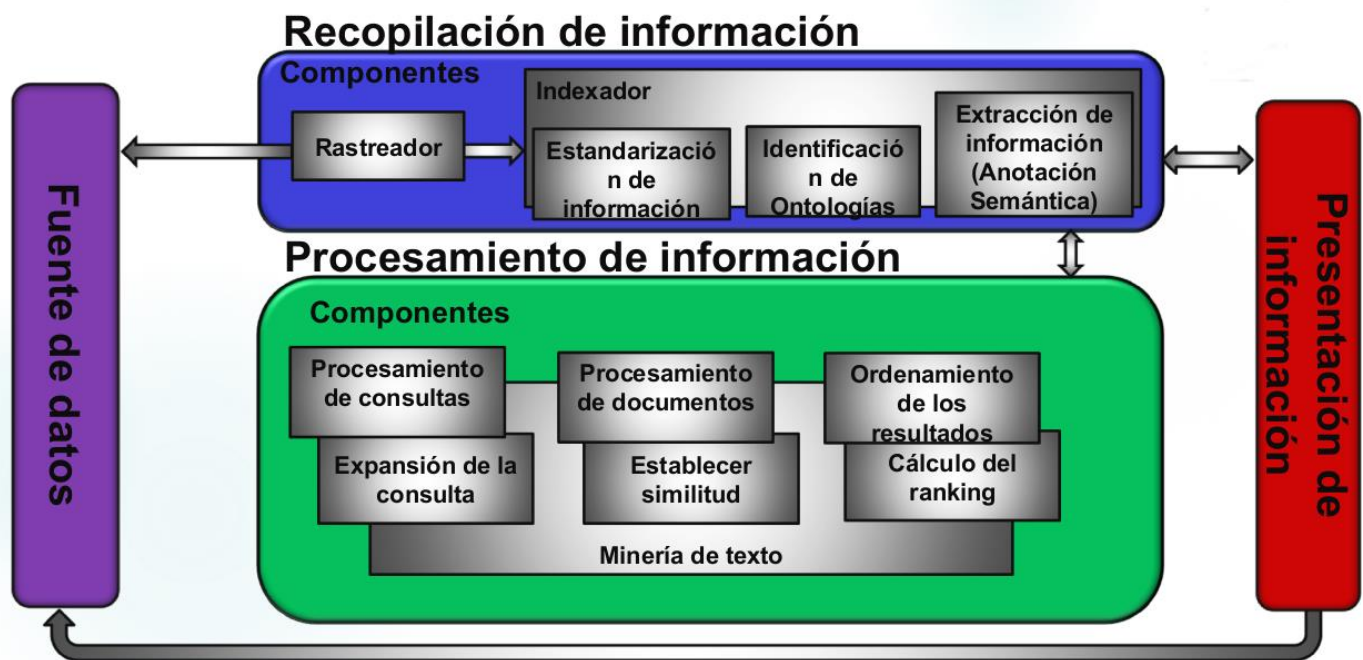


Ilustración 1: MODELO PROCESAMIENTO SEMÁNTICO EN SISTEMAS DE RECUPERACIÓN DE INFORMACIÓN

Procesamiento

Se encarga de procesar y analizar textos en lenguaje natural asociando cada sentencia de un texto a una representación semántica empleando como base una ontología con miles de palabras, donde las palabras se categorizan según los distintos significados que tienen y donde se definen las relaciones entre ellas. Una ontología según define Gruber (1993) “es una especificación explícita de una conceptualización” que permiten añadir un sentido a la información que se necesita procesar. Este componente procesa la consulta del usuario y selecciona los elementos que son más similares entre todos los documentos almacenados. Para los usuarios autenticados en el sistema se comprueba sus perfil de preferencia y se le añade un factor de relevancia para mejorar el resultado de búsqueda. Posteriormente se ordena cada elemento según la relevancia semánticamente que tiene para el usuario y le muestra una cantidad que sea fácil de asimilar.

Presentación de la Información

Permite mostrar la interfaz al usuario para que pueda realizar la consulta y es donde se le muestra los resultados.

Conclusiones

La aplicación del modelo para la recuperación de información permite mejorar la relevancia y calidad de la información mostrada a los usuarios.

Los usuarios pueden realizar preguntas en lenguaje natural y obtienen información relevante y en cantidades asimilables.

Se genera conocimiento mediante la información pública en la web, permitiendo potenciar las habilidades para adquirir nuevos conocimientos.

Referencias

Ávila, E. (2016) Bibliotecas digitales académicas y web semántica: elementos para la innovación de los servicios bibliotecarios y de información.

Bibliotecas. Vol 34, N° 1, enero-junio, 2016, pp. 1-11.

García, D. G., Guña, E. J., y Quinatoa, E. E. (2016) Utopía o realidad de aplicaciones informáticas en la educación. Caso universidad ecuatoriana Revista Publicando, 3(9).2016,119-137.

Guix, E. (2016) USO DE APLICACIONES WEB 3.0 EN UN CICLO FORMATIVO DE GRADO MEDIO. VALORACIÓN DEL ALUMNADO Y PROFESORADO. Revista DIM / Año 11 - No 33 - Marzo 2016

Martínez, D.(2010) Inteligencia competitiva y web 3.0: aprendizaje de estrategias y destrezas informacionales en la enseñanza superior, Alicante, España.

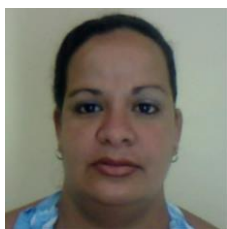
Sánchez, M., Prendes Espinosa, María Paz; Fernández Breis, Jesualdo Tomás (2013). Tecnologías semánticas para la evaluación en red: análisis de una experiencia con la herramienta OeLE. Revista de Investigación Educativa, 31 (2), 447-464.

Leguizamó, A M.(2016) La semántica como mecanismo integrador en entornos virtuales de formación No 01, Vol.I,pp. 61-77

Vaquerizo, M. B.(2012) ENSEÑANZA-APRENDIZAJE CON WEB 2.0 Y 3.0 Revista de Comunicación Vivat Academia, Año XIV NoEspecial, pp. 116- 121



Síntesis bibliográfica de los autores



Nombre y Apellidos: Aneyty Martín García

Edad: 32

Graduado de nivel: Universitario

Especialidad: Ingeniería en Ciencias Informáticas

Categoría docente: Asistente

Experiencia profesoral:

• Profesor de la asignatura Ingeniería de Software desde el curso 2007-2008 hasta el curso 2016-2017 en la carrera de Ingeniería en Ciencias Informáticas en la Universidad de las Ciencias Informáticas (UCI).

Actividad Científico Investigativa

Publicación Científica "Fundamentos para implementar y certificar un Sistema de Gestión de la Seguridad Informática bajo la Norma ISO/IEC 27001", en la serie científica en la Universidad de las Ciencias Informáticas.

Publicación Científica "Fuerza de Trabajo Calificada", en UCIENCIA 2008.

Publicación Científica "Buenas Prácticas de la metodología XP en el Centro de Informatización Universitaria", en UCIENCIA 2012.

Publicación Científica "Sistema Unificado de Gestión de Fuerza de Trabajo Calificada", en UCIENCIA 2012.

Publicación Científica "Compendio de ejercicios para el desarrollo de habilidades en la asignatura de Ingeniería de Software", en Informática 2016

Posgrados relacionados con las temáticas del evento:

La Tecnología Educativa como herramienta para generar engagement en el aprendizaje universitario.

- La tutoría en la Educación a Distancia
- La Tecnología Educativa. E-Learning, M-learning y U-Learning
- Fundamentos de Docencia Universitaria
- Cursos de la Maestría de Educación a Distancia Aprendizaje en la Edad Adulta
- Bases Teóricas de la EaD
- Diseño Curricular en la EaD
- Diseño de Módulos y la Organización tutorial
- La evaluación en la EaD
- Organización y Administración de centros de EaD
- Recursos Multimedia en EaD, Sociología de la EaD
- Tendencias Pedagógicas Contemporáneas y TIC en la EaD)

Síntesis bibliográfica de los autores



Nombre y Apellidos: Hubert Viltres Salas

Edad: 32

Graduado de nivel: Universitario

Especialidad: Ingeniería en Ciencias Informáticas

Categoría docente: Asistente

Experiencia profesoral:

- Profesor de la asignatura Problemas Sociales de la Ciencia y la Tecnología en el curso 2008-2009, profesor de la asignatura Desarrollo de aplicaciones web con Drupal desde 2008 hasta la fecha y profesor la asignatura Proyecto de Investigación y Desarrollo en la carrera de Ingeniería en Ciencias Informáticas en la Universidad de las Ciencias Informáticas (UCI).

Actividad Científico Investigativa

Publicación Científica “Componentes y funcionalidades de un sistema de recuperación de la información” en la Revista Cubana de Ciencias Informáticas.

Publicación Científica "Algorithm for calculating relevance of documents in information retrieval systems”, en International Research Journal of Engineering and Technology.

Publicación Científica “Personalización del sistema de gestión de contenidos Drupal” en II Simposio Informática y Comunidad. XIII Convención y Feria Internacional, Informática 2009.

Publicación Científica “Ubercart, módulo de Drupal para el comercio electrónico” en IV Taller Internacional de Comercio Electrónico. XV Convención y Feria Internacional, Informática 2011

Publicación Científica “La clase de problemas sociales de la ciencia y la tecnología como escenario de integración” en UCIENCIA 2010.