

TECNICAS INTELIGENTES PARA LA RECUPERACION DE CONOCIMIENTO EN LA ENSEÑANZA UNIVERSITARIA

INTELLIGENT TECHNIQUES FOR KNOWLEDGE RECOVERY IN UNIVERSITY EDUCATION

Antonio Martín
toni@us.es

Carlos León
cleon@us.es

Universidad de Sevilla. Departamento de Tecnología Electrónica.

El desarrollo de sistemas de búsqueda que faciliten la gestión del conocimiento académico en un espacio distribuido como son las Bibliotecas digitales universitarias, es uno de los principales objetivos de instituciones y proveedores de información. Estos nuevos retos motivan a los investigadores y a la comunidad docente a buscar nuevos enfoques en la recuperación eficiente de la información. El presente estudio supone un esfuerzo en innovación educativa, y propone un enfoque pragmático en la aplicación de la recuperación del conocimiento en las bibliotecas digitales. Para ello utilizamos un enfoque ontológico y técnicas de la inteligencia artificial.

Palabras Claves: Sistemas Expertos, Protégé, Integración Semántica de la Información, Web Semántica, Ontología, Integración de Sistemas, extracción del conocimiento.

The main goal of the academic institutions and information providers is to development a search engine to retrieval information in a super distributed data space like digital university libraries. This begets new challenges to docent community and motivates researchers to look for intelligent information retrieval approach that search and/or filter information automatically. We make an effort in innovation education in this direction and we propose a semantic method for efficient information search. This paper suggests a pragmatic approach to the implementation of intelligent techniques and ontologies for efficient knowledge retrieval in the academic digital libraries.

Keywords: Expert Systems, Protégé, Integration of Semantic Information, Semantic Web, Ontology, Integration Systems, knowledge information.

1. Introducción.

Actualmente la WEB es un espacio preparado para el intercambio de información que consta básicamente de información representada en forma de texto. Los buscadores actuales realizan las búsquedas

mediante palabras claves que aparecerán en el código HTML, dando lugar a resultados que tratan o contienen los patrones solicitados y la mayoría de las veces con información insustancial para el usuario. En una Web Semántica estructurada ontológicamente, los buscadores de información dejarán de arrojar

millones de resultados indiscriminados, la mayor parte de ellos irrelevantes, y ofrecerán información cualitativa, muy parecida a la que puede seleccionar un especialista humano en un dominio concreto de búsqueda.

Las redes de aprendizaje compuestas por personas y agentes conectados a través de la información semántica hacen del aprendizaje algo más fácil y dinámico, posibilitando la comunicación de conocimiento entre personas que trabajan o investigan en un mismo tema. Todo ello gracias a las tecnologías de la Web semántica que ayuda a los investigadores a reunir información útil mucho más rápido y de forma mucho más eficiente. El aprendizaje y la organización educativa será más fácil de organizar a través de agentes, que de forma automática hacen un seguimiento de los recursos, servicios, colecciones, eventos, etc.

Para ello es necesario adoptar medidas que establezcan normas, vocabularios, orientaciones y reglas adecuadas que faciliten la integración de datos de distintas procedencias y posibiliten un intercambio efectivo de información. El conocimiento de la Web debe representarse de forma que sea legible por los ordenadores, esté consensuado y sea reutilizable. Las ontologías proporcionan la vía para representar este conocimiento a través de conceptos, relaciones, funciones, instancias y axiomas. En este entorno se plantea una Web Semántica académica que codifique el significado de la información, que soporte además el diseño de robots de búsqueda que «entiendan» los documentos y realicen procesos «inteligentes» de selección, extracción y tratamiento de la información que realmente es relevante para el usuario (Guo, 2006). La idea es que los sistemas inteligentes tengan en cuenta los contextos y situaciones, de

forma que puedan ayudar a resolver problemas y responder preguntas, dotando a la ciencia y la educación de nuevas formas de tratamiento del conocimiento.

El uso de ontologías e inteligencia artificial abre nuevas vías de investigación y provee nuevas expectativas. En nuestro estudio analizamos la relación entre ambos factores y ofrecemos una técnica eficiente de búsqueda desde una visión no individualizada de la enseñanza, que incluye capacidad para un aprendizaje colaborativo. Presentamos una propuesta para la representación de material y recursos docentes, que permiten al usuario adquirir conocimientos a través de sus propias estrategias de aprendizaje. En este trabajo analizamos las potencialidades de la Web semántica y la inteligencia artificial en el contexto de la educación y las ventajas que aporta a la hora de clasificar el conocimiento adquirido por los estudiantes, permitiendo al entorno Web adaptarse a las características de sus participantes, de acuerdo a un nivel de conocimientos y a un perfil de usuario determinado (Pérez i Garcías, 2002).

Con esta finalidad, nuestro artículo se estructura de la siguiente manera: en la primera sección se presentan las necesidades en las Bibliotecas Digitales actuales, en la segunda se presentan la necesidad de las ontologías y los sistemas expertos, en la tercera se formulan y responden las cuestiones que ayudan a determinar el dominio de conocimiento. En la cuarta sección se presenta nuestro prototipo. Finalmente, se presentan las conclusiones y los trabajos futuros.

2. Fuentes semánticas de información y conocimiento.

Los sistemas actuales de recuperación de información académica tienen limitaciones desde el punto de vista del modelo de datos,

puesto que toman un único punto de vista del mundo. Describe los objetos o instancias de interés, pero bajo una sola posible interpretación. Si uno quiere reutilizar algún término, se hace evidente que el término puede tener diferentes interpretaciones dependiendo del contexto. Para poder reutilizar el conocimiento complejo es necesario tener en cuenta diferentes aspectos y significados de la información. Esto puede resolverse teniendo un entendimiento compartido que unifique los diferentes puntos de vista y establezca unos propósitos de comunicación e interoperabilidad entre sistemas. Para dar sentido a la información que contiene la Web es necesario basarse en la capacidad de asociar semánticamente datos independientes entre sí, pero relacionados en un determinado contexto.

Las tecnologías semánticas permiten un mejor aprovechamiento del espacio informacional, a la vez que facilitan la gestión del conocimiento y búsqueda de información desde distintas puntos de vista:

? Ayudando a la comunidad educativa en el desempeño de sus tareas de forma más flexibles. Permitiendo disponer de servicios educativos en línea más eficientes y menos aislados. Incluyendo tareas de desarrollo de enseñanza en línea, servicios al alumno, mejora de la obtención del conocimiento, manejo y administración de los recursos, etc.

? Ayudando a las personas que cumplen diferentes roles (alumnos, tutores y proveedores de contenido), a desempeñar sus funciones de forma más efectiva y eficientemente en espacios de aprendizaje grandes, distribuidos, basados en recursos de procedencia multiplataforma, multi-dominios, creando ambientes inteligentes de aprendizaje centrados en el usuario y adaptables a sus necesidades.

La Web semántica define y describe los datos contenidos en la Web, de tal forma que puedan ser utilizados y comprendidos por las máquinas sin necesidad de intervención humana. Admite una nueva manera de concebir la red, donde los contenidos pueden ser pre-procesados por agentes, que garanticen la significatividad de la información obtenida. En primer lugar, lenguajes de marcado que establecen normas acerca de cómo declarar la meta-información, como son las ontologías y en segundo lugar, un software inteligente o “agentes”, que sean capaces de hacer uso de estas normas de meta-información como son los sistemas expertos. De este modo, la Web se puede convertir en un espacio navegable y comprensible, donde es posible la relación entre términos independientes de una misma búsqueda que ofrece como resultado sólo aquellas páginas Web, que realmente se adaptan a los requisitos especificados por el usuario. Es decir únicamente aquellos términos que cumplen la condición de búsqueda, así como aquellos que tienen el mismo significado o uno equivalente. De esta manera, se dota una nueva Web de más significado, donde las máquinas no sólo son capaces de presentar toda la información contenida en ella (la sintaxis), sino que además puede entender y gestionar de forma inteligente su significado. La Web semántica propone una nueva arquitectura basada en 7 capas (Hagino, 2100) (véase la Figura 1)

Esta arquitectura permite el desarrollo de servicios Web semánticos y el diseño de agentes inteligentes, que facultan al usuario de una nueva forma de petición de servicios. Los agentes interactúan de forma autónoma entre ellos y se encargan de ubicar la información requerida, transformarla en conocimiento accesible y presentable al

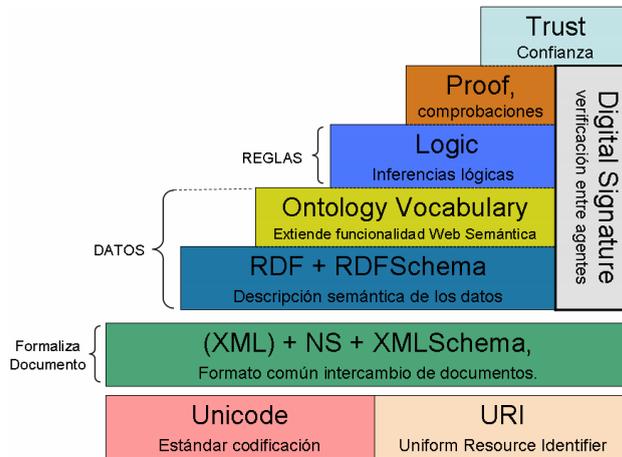


Figura 1. Arquitectura de la Web Semántica

usuario. De esta manera el proceso de arriba/abajo realizado hasta ahora para la presentación de un resultado de búsqueda determinado, se transforma en un proceso de ensamblado dinámico de subsistemas y servicios independientes, que en un momento dado se requieren (Taniar, 2006). De tal manera que con el uso de los servicios Web semánticos y de los agentes inteligentes, evolucionamos desde un entorno de revisión de documentos, a un entorno de obtención de información y resultados (véase la Figura 2).

Este enfoque abre un nuevo camino en el desarrollo de motores de búsqueda en el dominio educativo y en la gestión del conocimiento contenido en la Web académica. Permite la planificación y el desarrollo de sistemas inteligentes, mediante el diseño de una porción o subconjunto del dominio de aplicación pero bajo un esquema colaborativo y distribuido.

3. La educación y la inteligencia artificial.

La Web semántica gira sobre la existencia de mecanismos que permiten definir de una forma homogénea y consensuada la información semántica. Para entender el lenguaje humano, el software necesita de normas, agentes y una meta-información normalizada, que sea capaz de discernir entre distintas situaciones y hacer referencia a diferentes contextos. Para los científicos y educadores existe una necesidad creciente por organizar estructuralmente el contexto de aprendizaje en la red (Cabero et al, 2004). Para ello se están desarrollando diferentes maneras de representar el conocimiento en diferentes grupos de trabajo compuestos por investigadores, desarrolladores, científicos, ingenieros, y la W3C. Sin embargo, todavía no posee un único entorno capaz de clasificar todos los contenidos. En el campo de la inteligencia artificial, “lo que existe es aquello que puede ser representado”.

La descripción del conocimiento en la inteligencia artificial se conoce con el nombre

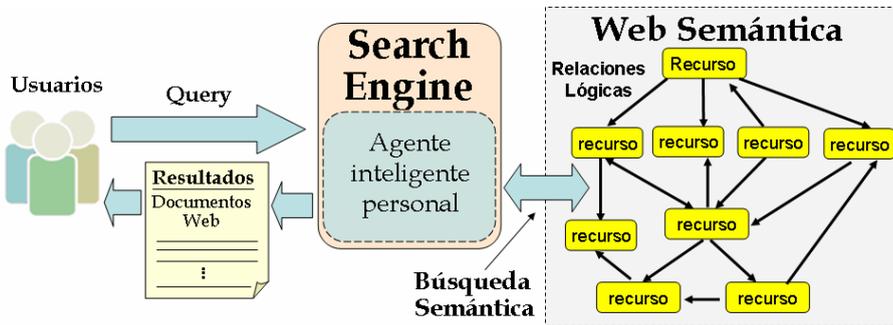


Figura 2. Proceso semántico de obtención de información

de representación del conocimiento. Dentro de este campo de estudio existen diferentes métodos de representación del conocimiento, como son las redes semánticas: Redes Is-As, Redes de Marcos, Ontologías, etc. Hoy en día, las ontologías se están convirtiendo en una de las técnicas más utilizadas para la representación del conocimiento. Una máquina no puede comprender la información que procesa en ningún sentido profundo de la expresión; sin embargo si los datos que procesa están estipulados semánticamente mediante ontologías, es posible que la Inteligencia Artificial simule la forma de razonar e inteligencia humana. Las ontologías definen de forma explícita y estructurada en taxonomías un dominio concreto de conocimiento que permitirá, simular que un ordenador manipula la información de manera funcionalmente parecida a como un humano entiende el lenguaje.

Una ontología permite ver como un modelo de datos de conocimiento que especifica una conceptualización, una forma de ver el mundo que contiene definiciones que proveen del vocabulario necesario para referirse a un dominio. Factores determinantes para la aplicación de las ontologías en las bibliotecas

digitales son:

- Posibilidad de existencia de ontologías múltiples para combinar dos o más taxonomías que introducen conceptualizaciones específicas.
- Abstracción de las ontologías para la creación de redes ontológicas que posibiliten una descripción completa del mundo.
- Multiplicidad de la representación. Un concepto puede ser caracterizado y representado de muchas formas, por lo que pueden coexistir múltiples descripciones de un mismo concepto.
- Mapeo de ontologías. Establece relaciones entre los elementos de una o más ontologías, para permitir conexiones, especializaciones, generalizaciones, etc.

4. Definición del conocimiento.

Implementar una Web Semántica requiere añadir metadatos semánticos que describan los datos, etiquetas como “Tipo Recurso”, “Autor”, “Título”, “Proveedor”, etc. que definen y caracterizan un recurso determinado. Esto permite a las computadoras procesar de manera efectiva los datos basados en la información semántica, los metadatos

que la describen. Cuando no hay suficiente información, las computadoras pueden hacer inferencias con estos metadatos acerca de los datos, su significado y cómo se relaciona unos con otros a fin de obtener los resultados esperados. Para conseguir que los computadores entiendan los datos, es necesario expresarlos en un formato uniforme, de forma que, por ejemplo, un campo llamado «titulo_publicacion» presente siempre el mismo formato y contenga el mismo tipo de información en dominios distintos.

Las ontologías buscan capturar y representar el conocimiento de forma consensuada, para ser reutilizado y compartido tanto por aplicaciones docentes, como por grupos de personas pertenecientes a la comunidad universitaria. Una ontología consiste en clases y sus atributos, las relaciones entre las clases, las propiedades de las relaciones y los axiomas que permiten restringir las interpretaciones de definiciones, de acuerdo al significado pretendido de los conceptos.

En el ámbito educativo ya existen diferentes propuestas de ontologías para describir el contenido de los recursos educativos, aunque creemos que en lugar de esperar una Web educativa semántica universal, es más argumentable decir que habrá diferentes sistemas para diferentes ámbitos de conocimiento que tienen que encontrar sus propias maneras de acercarse al mundo de la información semántica, hasta que llegue el día en que quizás se unan todas.

La ontología que hemos desarrollado define un vocabulario común para estudiantes e investigadores que necesitan compartir información en un dominio concreto, en nuestro caso la biblioteca digital de la Universidad de Sevilla. El dominio ontológico proporciona los conocimientos presentados

en forma adecuada para que el usuario pueda adquirir el conocimiento requerido y la capacidad de generar preguntas al sistema según el significado de las mismas.

4.1. Descripción de la ontología del sistema.

Considerando la cantidad, variedad de datos y de fuentes de información diferentes existente en el dominio educativo, es necesario recurrir a lenguajes y tecnologías que los clasifiquen sobre la base de sus propiedades y sus relaciones con otros datos. Aquí es donde aparecen las tecnologías de Web Semántica como RDF, RDFS y OWL. Estos lenguajes permiten mediante relaciones taxonómicas, crear una jerarquía de conceptos. Las clasificaciones semánticas se añaden a la red como meta-información para que sea leída por máquinas. Por lo tanto, necesitan normas, como la Web Ontology Language (OWL) o el Resource Description Framework (RDF), lenguajes legibles por las máquinas para describir formalmente contenidos multimedia. Aunque durante los últimos años se han propuesto varios lenguajes para la definición de ontologías, en Febrero del 2004 el W3C ha estandarizado OWL como lenguaje para la descripción semántica de recursos en la Web Semántica (Heflin, 2010). OWL se diseñó teniendo como objetivo ofrecer un mecanismo común de representar el conocimiento en la Web, de forma que fuera posible procesarlo de forma automática. Por lo tanto, OWL está pensado para que sea “leído” por máquinas y no por humanos. Este lenguaje cuenta con numerosas ventajas: posibilidad de compartir ontologías “públicamente” accesibles, permitir la evolución y compatibilidad de ontologías, capacidad de integración de ontologías que representan un mismo concepto de formas diferentes, detección de inconsistencias,

equilibrio entre expresividad y escalabilidad, etc. De este modo, OWL es el lenguaje utilizado para desarrollar nuestra ontología. Asimismo para la construcción y modelado de la misma hemos utilizado la herramienta de edición de ontologías de libre distribución llamada Protégé (PROTÉGÉ, 2010). Este editor ontológico ha sido desarrollado por la universidad de Stanford, basado en el lenguaje Java y en forma de plugins (Horridge et al 2004).

La ontología construida, incluye la definición de los conceptos y relaciones necesarios para desarrollar servicios eficientes de búsqueda de información, en el dominio de las bibliotecas universitarias. Debido a la gran cantidad de términos identificados en nuestro dominio, la ontología se organiza en varios espacios de nombres (o sub-ontologías) que incluyen las propiedades, con sus vocabularios correspondientes (Kim, 2005). Se distinguen dos ontologías:

- **Ontología de usuarios.** Define los grupos de personas y sus relaciones con otros individuos. La interacción entre grupos de usuarios durante las tareas de aprendizaje colaborativo, la semántica relacionada con los objetos del aprendizaje y finalmente, la semántica implícita en lenguajes orientados al diseño del aprendizaje. El modelo identifica

los elementos necesarios para describir las características de un potencial usuario, permite además especificar determinadas preferencias de usuario. Incluye las propiedades y clases relacionadas directamente con la caracterización de los usuarios del sistema. La ontología permite disponer de mecanismos inteligentes de guía al usuario, en un entorno asistido de aprendizaje. Se identifican varios conceptos o clases que pueden ser denominadas de primer nivel y que representan los distintos tipos de usuarios de búsquedas:

- Externo: Usuarios que aunque no pertenecen a la comunidad universitaria, pueden circunstancialmente hacer uso de los recursos y servicios: alumnos y profesores visitantes, personal investigador externos, etc.
- PAS: Personal de Administración y Servicios de la Universidad.
- Estudiantes: Alumnos de primero y segundo ciclo de carrera.
- Investigador: Personal docente e investigador. (Véase la figura 3).

- **Ontología de Servicios.** Esta segunda sub-ontología localiza los servicios y recursos ofrecidos en un determinado espacio o área de trabajo. Forma parte de la Web semántica

Perfil PDI

Nuestras colecciones

- **Catálogo Fama**
- **Recursos-e**
- **Fondos digitales**
- **Busque en nuestras colecciones**

Docencia e Investigación

- **RefWorks**
- **Repositorios institucionales**
- **Factor de impacto**
- **Enseñanza virtual**

Servicios

- **Acceso desde casa**
- **Propuesta de compra**
- **Préstamo , reserva y renovación**
- **Préstamo interbibliotecario y obtención del documento**

Ayuda

- **Pregunte al bibliotecario**
- **Horario de las bibliotecas**
- **Mis cuentas**

Figura 3. Servicios y recursos relacionados con el perfil PDI

de carácter bibliográfico-cultural de nuestra Universidad. Permite integrar términos que describen los recursos y servicios ofrecidos desde las distintas plataformas disponibles (Ortiz, 2005). Comprende las entidades que imparten suministran recursos dentro del dominio específico de nuestra biblioteca: el Catálogo Colectivo del Patrimonio Bibliográfico, la Web de la Biblioteca, la Intranet, el Fondo Antigo, Fondo Tesis y Fondos Fototeca. Las propiedades y clases definidas especifican los datos correspondientes a los recursos y servicios, como puedan ser: tipo de proveedor, datos identificativos, estructura de la información, recursos ofrecidos, etc.

5. Arquitectura del sistema.

Las nuevas tendencias en la arquitectura de sistemas distribuidos, así como las emergentes tecnologías de servicios Web semánticos, agentes autónomos inteligentes, metadatos de la Web semántica y motores lógicos de inferencia, posibilitan la construcción de metasistemas que permiten

las búsquedas de conocimiento desde el punto de vista de su significado. En este sentido hemos considerado su aplicación en el área educativa, centrándonos en el desarrollo un meta-buscador o metasistema de recuperación de conocimiento relacionado con la docencia y la investigación (Toussaint, 2006). Este metabuscador ha sido concebido como un sistema de información de lógica acoplable, con una estructura computacional independiente, que sigue los estándares de la Web semántica y que por tanto es posible hacerlo interactuar de forma colaborativa con otros sistemas de análogas características.

En este sentido hemos trabajado en nuestro prototipo, creando un dominio semántico y un agente inteligente de búsqueda. Un buscador inteligente basado en la información semántica ontológica que busca estratégicamente la información relevante, para la resolución de los patrones de búsqueda y desestima todos los recursos que no se ajusten al significado perseguido. En la Figura 4 se presentan las distintas entidades que forman parte de la arquitectura del sistema.

MOTOR DE BUSQUEDA

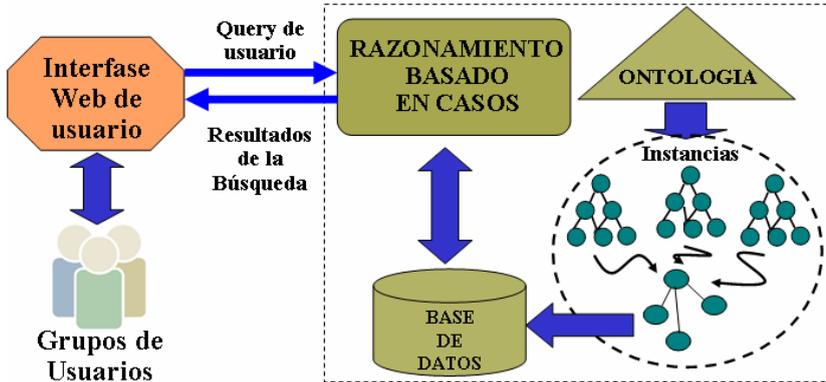


Figura 4: Arquitectura del sistema OntoBUS

nuestro sistema contiene la taxonomía de todos los conceptos que existen en nuestro dominio académico y que servirá como base para la realización de las búsquedas. La estructura del caso definido por el experto deberá ser consistente con la ontología y tendrá una correspondencia directa con las clases y propiedades definidas. Los recursos estarán representados por las instancias de cada concepto en la ontología. Un usuario podrá no solo consultar la información disponible gracias al concepto por el cual se preguntó, sino que también podrá sugerir otros conceptos asociados y términos más generales o más específicos de exploración, para ampliar o precisar sus búsquedas. Además, si el usuario busca por un concepto alternativo, podrá realizar la consulta de la misma forma que si hubiera preguntado por el concepto preferente.

Un elemento imprescindible asociado a la ontología y que forma parte del Onto-Buscador es la Base de Conocimiento o Base

de casos. Esta entidad almacena todos los casos o instancias del sistema experto definidos a través de las sucesivas búsquedas expertas. Incluye toda la información recogida e inferida por el sistema tanto de los recursos y servicios ofrecidos, como de los propios usuarios. La organización del conocimiento en los casos provistos, es realizada por un experto y dotada de una estructura común definida por el ingeniero del conocimiento. La representación de los casos utiliza las denominadas “tripletas” formadas por: asunto, predicado y objeto. Por ejemplo el recurso “Computer Science” (asunto) está en la dirección (predicado) <http://bib.us.es/recursoelectronico/computerscience.pdf> (objeto).

Una clasificación manual de estas instancias en las clases establecidas dentro de la ontología sería una tarea sumamente costosa, dada la bastedad de conocimiento existente en el dominio. Por tanto, se realizó una primera carga de información procedente desde las

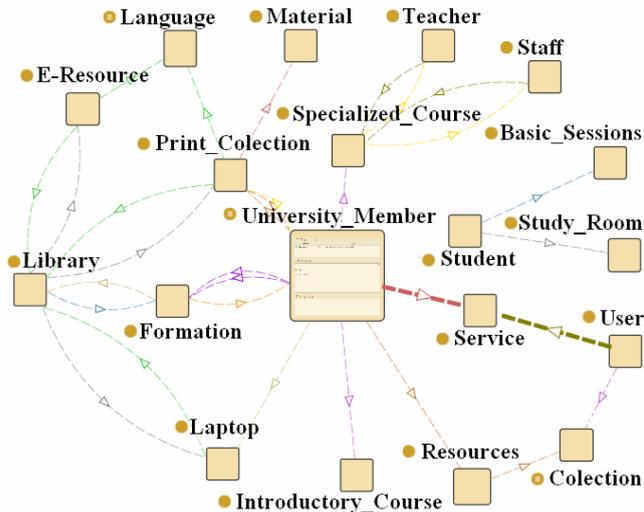


Figura 5. Clases definidas en la ontología del sistema.

distintas fuentes existentes en la Biblioteca. Esta información se transformó en declaraciones OWL, que acepta las búsquedas complejas utilizando un lenguaje de consultas sobre instancias de ontologías (véase figura 5).

La Base de datos que integra el sistema, dará soporte físico a la base de casos. Es decir, servirá como implementación del alojamiento de forma permanente de la base de casos. Además contendrá las tablas de definición de perfiles de usuarios, para crear los distintos entornos de búsqueda y para la habilitación de políticas de acceso al sistema (Dumais et al, 2003). Incluye también el control del acceso a los mismos: edición de casos, gestión de usuarios, autenticación de usuarios, etc.

Por último el sistema Experto Basado en Correspondencia de Casos, CBR es la entidad encargada de realizar las inferencias sobre los casos almacenados (Golbreich, 2004). La presentación de los datos obtenidos, como solución óptima del caso propuesto al sistema experto (catalogación, búsquedas, presentación de resultados, visualización,

etc.) así como la administración de las bases de datos necesarias. Este agente CBR se ha implementado mediante la plataforma de desarrollo jCOLIBRI, que permite crear aplicaciones CBR, de una forma sencilla. jCOLIBRI es un Shell de desarrollo software que incluye una interfaz gráfica de usuario (GUI), que posibilita de forma tutelada la construcción de sistemas CBR (Recio-García, et al 2006). El uso de esta herramienta como entorno de construcción de sistemas CBR, se justifica por ser un sistema implementado en Java, de código abierto, fácilmente adaptable y/o extensible a entornos específicos.

6. Interfaz gráfica de usuario.

Nuestro sistema trata de agilizar el acceso a los recursos pertenecientes a la biblioteca digital, optimizando para ello el resultado de las búsquedas realizadas sobre la ontología y haciendo uso de distintos perfiles inteligentes de búsqueda. El usuario interactúa con el sistema de forma eficiente a través de la interfase, que permite a cada tipo de usuario

Figura 6. Resultados de perfil Estudiantes

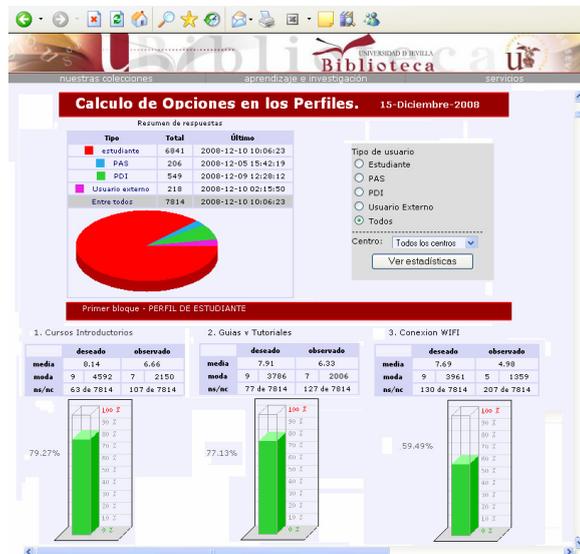


Figura 7. Análisis matemático de las opciones de un Perfil

seleccionar un perfil de búsqueda individual o colectivo. Cuando se accede por primera vez a nuestra aplicación, aparecerá por defecto una primera aproximación de búsqueda, con distintos recursos seleccionados para un determinado perfil colectivo. Nuestro interfaz permite caracterizar el comportamiento de un grupo de usuarios y el de los individuos que lo componen a través de un conjunto de atributos o etiquetas. El agente facilitador utiliza estos atributos que introducen los usuarios, para ofrecer sugerencias y consejos con el objetivo de mejorar la interacción dentro de cada grupo (Martín, 09) (véase la figura 6).

Esta primera aproximación al perfil se obtiene a partir de un método de clasificación automático, que genera las opciones de búsqueda por defecto. El mejor resultado se obtiene de las opciones utilizadas y seleccionadas anteriormente por los usuarios en sus distintos perfiles de búsqueda. El

usuario, puede cambiar las posibilidades de búsqueda iniciales, ajustando los distintos valores a través de la interfaz Web, que ofrece la posibilidad de seleccionar materias, estudios o titulación, sobre la que queremos centrar nuestra consulta. Tras realizar los primeros ajustes de búsqueda, podemos proceder a la obtención de los resultados.

La presentación de las opciones iniciales en los distintos perfiles, es dinámica. Es decir se va modificando en función de las distintas búsquedas efectuadas por usuarios anteriores. Pueden ir apareciendo nuevas opciones o desapareciendo algunas que había. Para la construcción de un nuevo perfil se realiza un análisis de las operaciones realizadas por los usuarios expertos, a través de una aplicación software realiza una serie de cálculos matemáticos y estadísticos. (Véase la figura 7).

Con esta B.D. se establece un análisis de

uso y preferencias de búsquedas aportadas por los distintos colectivos que utilizan el OntoBuscador y facilita la creación dinámica del perfil de búsqueda correspondiente a cada tipo de usuario.

7. Evaluación del sistema

En esta sección exponemos el método seguido para la evaluación de nuestro onto-buscador respecto de un buscador tradicional como es por ejemplo Google. Usuarios pertenecientes a los distintos colectivos definidos en el sistema, han realizado una batería de consultas relacionadas con los recursos y servicios existentes en nuestra Biblioteca Digital. Las consultas realizadas abarcan consultas con una única palabra y consultas con varias palabras, de las cuales

al menos una resulta por sí misma ambigua, pero tiene sentido con el contexto de la consulta.

Para evaluar los resultados de las búsquedas tendremos en cuenta el orden de los términos buscados, que deben aparecer en el mismo en el que se formuló la consulta y no solo las palabras por separado. Es decir si la expresión a buscar es: “Libro de Ingeniería de la Programación”, se puntuaría que esa sea la frase que aparece en la página buscada y no en otro orden.

Otro factor que valoramos como resultado de calidad, es que el contexto donde se integra la frase tenga algo que ver con el sentido que se está buscando. Por ejemplo, si buscamos: “Recursos electrónicos sobre inteligencia artificial” esperamos obtener todos los accesos on-line ofrecidos por la biblioteca

The screenshot shows a comparison between two search methods: 'OntoFAMA (Ontology Search)' and 'Traditional Search (No ontology Search)'. The search query is 'computer science book in library of Seville University'. The 'Ontology Search' results are more precise, listing books like 'Manual de derecho informático' and 'Cómo programar en C#'. The 'Traditional Search' results are more general, including 'Allegory and other matters in the Libro de buen amor' and 'La biblioteca de gente 2'. The interface includes a search bar at the bottom with the query and buttons for 'Back' and 'Retain'.

OntoFAMA (Ontology Search)	Traditional Search (No ontology Search)
<p>Título: Manual de derecho informático Autor: Davara Rodriguez, Miguel Angel Biblioteca: Informática Publicacion: Cizur Menor (Navarra) : Thomson-Aranzadi, 2008 Materia: Informática -- Derecho</p> <p>Título: Cómo programar en C# Autor: Deitel, Harvey M. Biblioteca: Informática Publicacion: México [etc] : Pearson Educación, 2007 Materia: C# (Lenguaje de programación)</p> <p>Título: Manual imprescindible de HTML Autor: Galeano Gil, Germán Biblioteca: Informática Publicacion: Madrid : Anaya Multimedia, 2009 Materia: HTML (Lenguaje de marcas)</p> <p>Título: Aprender a programar : algoritmos y fundamentos de programación orientados a la ingeniería y ciencias Autor: Rodríguez Rancel, Mario Biblioteca: Informática Publicacion: [La Laguna, Tenerife] : Mario Rodríguez Rancel, 2006 Materia: Programación orientada al objeto (Informática)</p> <p>Título: Visual Basic 2008 : curso de iniciación Autor: D'Andrea, Edgar Biblioteca: Informática Materia: Visual Basic 2008 (Lenguaje de programación)</p>	<p><input checked="" type="checkbox"/> Título: Allegory and other matters in the Libro de buen amor Autor: Hart, Thomas R. Publicacion: London : Dept. of Hispanic Studies, Queen Mary, University of London, 2007 Materia: Ruiz, Juan, Arcipreste de Hita, n. 1283. Libro de Buen Amor Alegoria</p> <p><input checked="" type="checkbox"/> Título: La biblioteca de gente 2 [Archivo de ordenador] : Curso de español : materiales complementarios para tus clases / [idea original, Pilar Salamanca, Neus Sans, Sergio Troitño] Publicacion: Barcelona : Difusión, 2008 Materia: Español (Lengua) -- Estudio y enseñanza -- Alófonos</p> <p><input checked="" type="checkbox"/> Título: Neural computation [Recurso electrónico] Publicacion: Cambridge, MA : The MIT Press Materia: Redes neuronales (Informática) -- Publicaciones periódicas Revistas electrónicas</p> <p><input type="checkbox"/> Título: Guía de recursos de empleo en la ciudad de Sevilla Publicacion: Sevilla: Observatorio de la Calidad del Empleo, Ayuntamiento de Sevilla, 2006 Materia: Búsqueda de empleo -- España -- Sevilla</p> <p><input type="checkbox"/> Título: Neural computation [Recurso electrónico] Publicacion: Cambridge, MA : The MIT Press Materia: Redes neuronales (Informática) -- Publicaciones periódicas Revistas electrónicas</p> <p><input type="checkbox"/> Título: Guía de recursos de empleo en la ciudad de Sevilla Publicacion: Sevilla: Observatorio de la Calidad del Empleo, Ayuntamiento de</p>

Figura 8. Resultados de búsquedas en OntoBUS

digital referentes a "Inteligencia Artificial" y no un libro con el título siguiente "Recursos de la Inteligencia Artificial para recuperar documentos electrónicos" (figura 8).

Para medir la eficiencia de nuestro sistema hemos tenido en cuenta el tiempo de respuesta de la búsqueda, accesibilidad de la interfaz, servicios adicionales ofrecidos, etc. En nuestro caso hemos utilizado la relevancia y precisión de los resultados obtenidos. La relevancia de los documentos recuperados, indica en que proporción ha sido satisfecha la necesidad de información, de los usuarios que hacen la consulta. La relevancia puede ser un criterio subjetivo debido a que diferentes personas asignarían diferentes valores de relevancia a un documento. En nuestro estudio hemos asignado un peso a cada resultado según la relación y calida de los documentos recuperados, respecto al objetivo de su búsqueda (Tabla 1).

Los documentos obtenidos por nuestro buscador han obtenido una puntuación media de 2,82, superior a la del otro buscador evaluado, que han obtenido una puntuación media de 2,75.

8. Conclusiones y trabajos futuros.

Internet y las Bibliotecas Digitales se han convertido en la principal fuente de

información en muchos contextos educativos, donde es posible encontrar conocimiento y materiales de aprendizaje. Dado que es incalculable la cantidad de información que se encuentra disponible en estos dominios, es necesario obtener información útil de esa gran Web distribuida de una forma eficiente y rápida. Para facilitar las búsquedas, y orientarlas al significado de la información contenida, surge el concepto de Web Semántica, que tiene como objetivo incorporar elementos que den significado semántico a los documentos contenidos en Internet. Estos elementos a la vez, pueden ser procesados automáticamente por los ordenadores, permitiendo a éstos participar en el proceso de búsquedas por significados.

El objetivo de nuestro proyecto es satisfacer las nuevas expectativas y requisitos hacia los sistemas de aprendizaje y recuperación basados en la Web semántica, desarrollando para ello un sistema adaptativo e inteligente. Proponemos una visión semántica para la localización del conocimiento en un contexto docente, donde la Inteligencia Artificial (IA) favorece la mejora de los procesos de aprendizaje. Una herramienta complementaria de enseñanza/aprendizaje, que permite aumentar la calidad de éstos, acercando de forma transparente al usuario el conocimiento distribuido en distintas fuentes de

	Excelente	Bueno	Aceptable	Deficiente
Relevancia	4	3	2	1
Calidad	4	3	2	1

Tabla 1. Ponderaciones para medir los resultados obtenidos.

información. Especialmente al usuario no experto, facilitando así el proceso de construcción de sistemas de búsqueda tutorizados y equilibrando la relación entre el coste y la eficiencia de las búsquedas.

El presente trabajo muestra cómo por medio de las tecnologías asociadas a la Web semántica y la inteligencia artificial, es posible contribuir al mejoramiento en la recuperación de la información para un dominio específico como es el aprendizaje y la investigación. En este artículo se ha descrito OntoBUS, el cual contiene una ontología desarrollada en OWL, que integra recursos y servicios existentes en el dominio de las Bibliotecas Universitarias. La ontología se ha desarrollado utilizando técnicas de la Web semántica y la inteligencia artificial, proporcionando un modelo común para expresar los elementos existentes y que facilita el acceso a los mismos. Las clases y propiedades definidas pueden ser utilizadas para formular consultas complejas y recuperar de forma más estricta la información requerida por los usuarios. Comentar asimismo que se han realizado diferentes experimentos y pruebas sobre nuestro prototipo y que se han obtenidos resultados satisfactorios.

Para finalizar proponer como trabajos futuros la integración y explotación de información procedente de otras bibliotecas digitales y servicios, perfeccionar el motor de búsquedas CBR y agregar nuevas funcionalidades a la interfase de usuarios del sistema.

9. Referencias bibliográficas.

Cabero, Julio. & Morales, J. A., (2004). La red como instrumento de formación. Bases para el diseño de materiales didácticos. *Pixel-Bit. Revista de Medios y educación*, 22.

Dumais, S. T. & Cutrell, E. (2003). Stuff I've

Seen: A system for personal information retrieval and re-use. *Proceedings of SIGIR*

Golbreich, C. (2004). Combining rule and ontology reasoners for the semantic Web. in *RuleML (G. Antoniou and H. Boley, eds.), vol. 3323 of Lecture Notes in Computer Science, pp. 6–22, Springer.*

Guo, W. & Chen, D., (2006). Semantic Approach for e-learning System. *Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences.*

Heflin, J. (2010). OWL Web Ontology Language Use Cases and Requirements, *W3C Recommendation, Word Wide Web Consortium*, (<http://www.w3.org/TR/webont-req>) (05-06 2010).

Horridge, M. & Knublauch, H. (2004). A Practical Guide To Building OWL Ontologies Using The Protégé-OWL Plugin and CO-ODE Tools. *The University Of Manchester. Reino Unido.*

Kim, T. & Kim, M. (2005). On Employing Ontology to e-Learning. *Proceedings of the Fourth Annual ACIS International Conference on Computer and Information Science (ICIS'05).*

Martín, A. & León, C. (2009). Intelligent Retrieval in a Digital Library Using Semantic Web, *Proceedings of the 5th International Conference on Education, 77-81. International Association for the Development of Advances in Technology (IADAT).*

Ortiz, A., (2005). Interacción y TIC en la docencia Universitaria. *Pixel-Bit. Revista de Medios y educación*, 26.

Pérez i Garcías, A., (2002). Elementos para el análisis de la interacción educativa en los nuevos entornos de aprendizaje. *Pixel-Bit. Revista de Medios y educación*, 19.

PROTEGE. (2010). The Protégé Ontology Editor and Knowledge Acquisition System.

(<http://protege.stanford.edu/>) (5-09-2009)

Recio-García, J.A. & Díaz-Agudo, B. (2006). Ontology based CBR with jCOLIBRI. *Procs. of the 26th SGAI Int. Conference AI-2006. Springer-Verlang.*

Taniar, D. & Wenny, J. (2006). Web semantics and ontology. *Hershey, PA: Idea Group Pub.*

Toussaint, J. & Cheng, K. (2006). Web-based CBR (case-based reasoning) as a tool with the application to tooling selection. *International Journal of Advanced Manufacturing Technology.* 29(1-2), 24-34.

Hagino, T. (2010). What Semantic Web Different from Current Web, *W3C Recommendation, Word Wide Web Consortium.* (<http://www.w3.org/2001/09/21-orf/hagino-sw/>) (28-12-2010).

Fecha de recepción: 10-05-2010

Fecha de evaluación: 06-10-2010

Fecha de aceptación: 16-11-2010

Fecha de publicación: 01-01-2011