Ontologías borrosas para representar perfiles de usuario en una herramienta de gestión de objetos de aprendizaje

Mateus Ferreira-Satler[1], Victor H. Menéndez[2], Francisco P. Romero [3], Alfredo Zapata [2], Manuel E. Prieto [4]

[1]Escuela Superior de Informática, Universidad de Castilla La Mancha, msatler@gmail.com,
[2] Universidad Autónoma de Yucatán, Mérida, México, {mdoming, zgonzal}@uady.mx
[3] E.U. Ing. Tec. Industrial de Toledo, Univ. de Castilla La Mancha, FranciscoP.Romero@uclm.es
[4]Dept. de Tecnologías y Sistemas de Información, Univ. de Castilla La Mancha, ManuelE.Prieto@uclm.es

Resumen

Este trabajo propone una aplicación específica de la utilización de ontologías borrosas con el fin de representar perfiles de usuario, que permitan mejorar el rendimiento en tareas como el filtrado y la recuperación de información. La construcción de estas ontologías se realiza de forma automática utilizando para ello técnicas relacionadas con la lógica borrosa. Con el fin de probar la validez de esta aproximación, se han integrado en una herramienta de gestión de objetos de aprendizaje, en la cual, el perfil de cada uno de los usuarios se construye a partir de los objetos de aprendizaje publicados por él mismo. Según los experimentos realizados, el método de construcción y representación se adecua con gran precisión y exhaustividad a lo requerido por el usuario.

Palabras Clave: Ontologías Borrosas, Perfiles de Usuario, Objetos de aprendizaje.

1 INTRODUCCIÓN

La proliferación de sitios y repositorios dedicados a ofrecer recursos destinados a la educación es una prueba del constante desarrollo del e-learning en la actualidad. Es por ello que numerosas instituciones transforman y publican recursos educativos para que sean aprovechados y reciclados.

A pesar de los esfuerzos de normalización y estandarización de los recursos, un problema latente en toda aplicación e-Learning y especialmente en los repositorios de recursos educativos (LCMS, Learning Content Management System) es poder ofrecer los recursos apropiados a las necesidades de los usuarios [4]: profesores, estudiantes, diseñadores instruccionales, etc. Cada uno de estos usuarios posee metas específicas y está dentro de un contexto distinto cuando busca alguna información. La personalización de los procesos de filtrado y recuperación de la información contribuirá a atenuar la sobrecarga de información, ajustando los resultados presentados a las necesidades individuales de cada usuario y mejorando la eficiencia en la gestión de los recursos educativos a su cargo.

Uno de los medios que soportan la mejora del proceso de filtrado y recuperación de información son las ontologías. El uso de ontologías ha recibido gran atención en la organización y distribución del conocimiento porque agrega un nivel conceptual en la recuperación de información, en contraste con la recuperación meramente sintáctica. Además, los perfiles de usuario basados en ontologías vienen siendo ampliamente aplicados en la representación del contexto y en la personalización de aplicaciones, de manera que éstas se ajusten a los requisitos del usuario.

En este trabajo se propone la integración de los perfiles de usuario y las ontologías como herramientas de soporte a los procesos de gestión de objetos de aprendizaje dentro de un LCMS, para la mejora tanto del filtrado y recuperación de la información como la clasificación y gestión de estos elementos. La plataforma AGORA [5] pretende apoyar a los profesores en el desarrollo de objetos de aprendizaje acordes a estrategias de diseño educativo y empleando un enfoque de recomendación. Dentro de este contexto, se ha abordado la definición, aplicación y validación de la construcción automática de perfiles de usuario basados en ontologías borrosas.

El resto del artículo está organizado de la siguiente manera. En la siguiente sección se presenta la herramienta AGORA, base de nuestro estudio experimental. La sección 3 está dedicada a la representación del perfil del usuario mediante ontologías borrosas y el proceso de construcción de este perfil. En la sección 4 se presentan los resultados del caso de estudio realizado sobre el uso de AGORA, dedicando la última sección a las conclusiones y al trabajo futuro.

2 AGORA: GESTIÓN DE OBJETOS DE APRENDIZAJE

Un objeto de aprendizaje es una unidad básica de información en e-Learning, que combina uno o más recursos educativos con los metadatos que los describen. Esta estructura permite su intercambio y reutilización entre distintos sistemas e-Learning sin que exista pérdida de su funcionalidad o características.

Los objetos de aprendizaje pueden ser cualquier recurso que el profesor considere relevante para la instrucción de sus alumnos. Como ejemplos de estos recursos se pueden citar los objetos de aprendizaje basados en documentos con información textual, que fueran utilizados como fuente de información para construir el perfil de usuario en este trabajo; y los objetos de aprendizaje basados en multimedia, como videos o audios.

La gestión de objetos de aprendizaje implica un conjunto de actividades requeridas para controlar su ciclo de vida. Su finalidad es el desarrollo de objetos de aprendizaje de calidad que puedan ser reutilizados en distintas soluciones e-Learning. Dicha gestión puede realizarse por medio de un entorno que integre las características de edición, transformación, almacenamiento y colaboración presentes de forma individual en las herramientas existentes para el desarrollo de objetos.

La plataforma AGORA (Ayuda para la Gestión de Objetos Reusables de Aprendizaje) [5] surge como una solución para los procesos de gestión de objetos de aprendizaje realizados de forma integral. Uno de sus objetivos es emplear un enfoque de asistencia y recomendación para la ejecución de procesos como el etiquetado (generación de metadatos), la búsqueda y la secuenciación de objetos de aprendizaje. Para ello, se consideran factores como el diseño instruccional, la gestión del conocimiento, su historial de acciones, etc. Un papel clave para lograr la máxima eficiencia en esta labor de gestión es la definición concreta y correcta del perfil del profesor, punto en el cual se centra este trabajo.

AGORA está formado por cinco módulos, cada uno de ellos especializado en una función concreta: Un módulo de gestión del conocimiento y su representación basada en ontologías instruccionales y de dominio; un módulo de gestión de objetos de aprendizaje para la administración de los recursos y metadatos propios; un módulo de metabúsqueda para buscar objetos de aprendizaje almacenados en otros repositorios, un módulo extracción de conocimiento para el análisis de acciones realizadas por el usuario para la generación de perfiles y modelos de comportamiento. Por último, el sistema cuenta con un módulo de recomendación que utiliza los módulos anteriores para dar respuesta a las necesidades de los usuarios.

Las principales características de AGORA pueden resumirse en:

- Una red de repositorios distribuidos en las instituciones participantes.
- Permite gestionar cualquier tipo de recurso utilizado en un LMS – Learning Management System (presentaciones, documentos, videos).
- Incluye un generador-etiquetador de objetos, un meta-buscador, un mecanismo de evaluación de la calidad y un entorno de presentación de contenidos.
- Emplea asistentes y recomendadores para la realización de las actividades.
- Fomenta la colaboración, al ofrecer mecanismos de comunicación e intercambio de conocimiento entre los usuarios.

Dentro de AGORA, los usuarios pueden crear, publicar, agregar, manipular, descubrir y almacenar objetos de aprendizaje. También pueden realizar una búsqueda sobre los objetos publicados. Para optimizar estas búsquedas un factor clave es el aprovechamiento del conocimiento inherente al entorno y al usuario. Con este propósito, se pretende incorporar el perfil de cada uno de los profesores. Para ello, la construcción automática de este perfil se implementa dentro del módulo de extracción de conocimiento. El módulo utiliza un componente que permite definir el perfil de cada usuario a partir de la base que viene conformada por los objetos de aprendizaje que él haya publicado en el sistema. De esta manera, se puede disponer de una representación del perfil de acuerdo con los intereses de trabajo del propio profesor.

Dentro de este contexto, es importante considerar la incorporación de características semánticas al perfil, mediante técnicas de identificación de preferencias de usuario y organización de la información. Las ontologías y la lógica borrosa proporcionan los medios adecuados para representar la información en formato digital y permiten describir una vista general del dominio de interés tratado.

3 PERFILES DE USUARIO BASADOS EN ONTOLOGÍAS BORROSAS

La propuesta presentada en este trabajo incluye varias etapas de procesamiento de datos para generar automáticamente perfiles de usuario basados en ontologías borrosas, a partir de un conjunto de documentos previamente seleccionados. Para lograr dicho objetivo, el proceso de construcción ha sido dividido en cinco etapas: pre-procesamiento lingüístico, indexación de términos (lo que llamamos pre-ontología), extracción de términos relevantes, construcción de la ontología del usuario y actualización del perfil del usuario. La Figura 1 muestra la estructura del proceso.

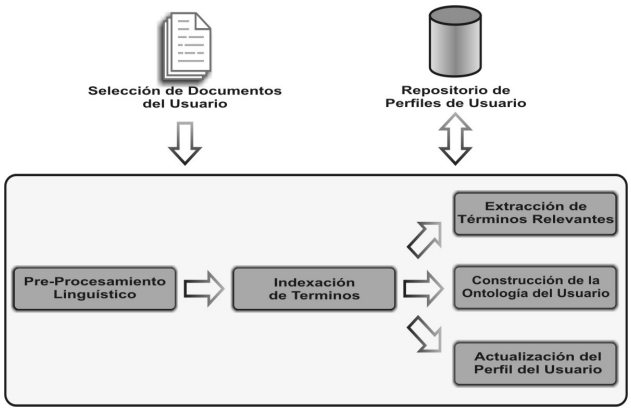

Figura 1: Proceso de Generación Automática de Perfiles de Usuario basados en Ontologías Borrosas.

En las siguientes secciones se describen en detalle cada una de las etapas del proceso.

3.1. PRE-PROCESAMIENTO LINGÜÍSTICO

El principal objetivo de esta etapa es transformar los documentos de los objetos de aprendizaje colocados por un usuario de AGORA en una representación apropiada, para que los mismos puedan ser procesados en las próximas fases.

Primero, se realiza la transformación previa de los documentos para tratar los diferentes tipos de formato (.doc, .txt, .pdf, etc.). Después, se eliminan todas las informaciones no textuales como dígitos, fechas, signos de puntuación, entre otros (análisis léxico) y se identifica el idioma de los textos, a través de una técnica que utiliza diccionarios de palabras frecuentes [1]. También se realiza la corrección ortográfica para sustituir las palabras que no estén escritas correctamente, utilizando diccionarios de corrección automática de ortografía.

El siguiente paso tiene por objeto reducir el vocabulario y generar una representación textual más significativa. Se utilizan dos técnicas conocidas como *Stoplists* [3] y *Stemming* [2]. Estas técnicas, aplicadas de manera adecuada, reducen el tamaño del vocabulario y muchas veces mejoran el desempeño del proceso.

En esta fase, una secuencia de caracteres define un término. Los términos están separados unos de los otros por un delimitador. El proceso prevé la existencia de términos formados por una o más palabras. De cada documento propiedad del usuario se extrae un conjunto de términos. Para que un término sea considerado como válido, el mismo debe aparecer por lo menos en tres documentos. Se utiliza esta estrategia para eliminar las palabras que aportan poca carga semántica al contexto del dominio. Solamente los términos válidos son procesados en las etapas posteriores.

3.2. INDEXACIÓN DE TÉRMINOS

En esta etapa se construye una estructura de índice, llamada *pre-ontología*, que contiene información sobre todos los términos generados en la fase anterior. La *pre-ontología* está formada por un conjunto de objetos del tipo *"term"*. A su vez, cada objeto *"term"* posee un identificador, que en este caso es el propio término, y una lista de objetos que contienen informaciones relacionadas con los documentos en los cuales el término aparece.

Cada objeto de esta lista de documentos contiene las siguientes informaciones: *Identificador del Documento; Número de Ocurrencias del términos más frecuente en el documento; Lista de posiciones del término en el documento.* La lista de posiciones contiene la posición física del término en el documento. La Figura 2 muestra la estructura de la *pre-ontología*.

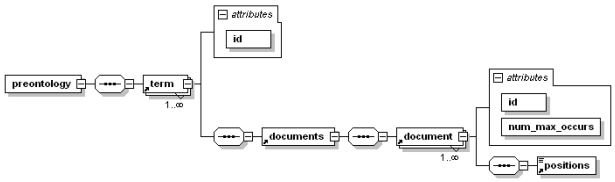

Figura 2: Estructura de la Pre-Ontología.

La *pre-ontología* sirve como base, tanto para la extracción de términos relevantes como para la construcción de la ontología. También sirve como soporte al proceso de actualización de resultados.

3.3. EXTRACCIÓN DE TÉRMINOS RELEVANTES

Con el fin de generar una representación semántica del contexto del usuario, se deben extraer los términos más relevantes dentro de los documentos que ayudan a definir el perfil. Con este fin, es necesario establecer un método para cuantificar el valor de representatividad de un término con relación a un conjunto de documentos.
En este trabajo han sido implementadas dos medidas de relevancia diferentes para calcular el peso de los términos. La primera de ellas es una ampliación de la estrategia propuesta en [6] basada en la medida clásica *tf-idf* con el fin de obtener el peso relativo (pr_i) de cada término dentro de los documentos del perfil:

$$pr_i = \frac{docs(i) + \sum_{j \in D} \frac{ft(i,j)}{|mf_j|}}{|D|} \qquad (1)$$

Donde se utiliza el número de documentos en que el término *i* aparece (*docs(i)*) y su frecuencia de ocurrencia en estos documentos (*ft(i,j)*). La normalización del valor se obtiene utilizando el número de documentos procesados ($|D|$) y el número de ocurrencias del término más frecuente del documento ($|mf_j|$).

La segunda medida de relevancia es una variación de la medida propuesta en [8] llamada KF-IDF. De acuerdo con esta fórmula, el peso de un término crece logarítmicamente inversamente proporcional al número de pre-ontologías de usuarios en las que aparece. En otras palabras, se considera un término como relevante si el mismo ocurre más frecuentemente que otros términos en una determinada pre-ontología de usuario, y ocasionalmente dentro de las restantes. Como se muestra en la fórmula (2), KF-IDF considera la relevancia de un término en base a un grupo o conjunto de usuarios. La nueva medida del peso, llamada *kmod*, se define de la siguiente manera:

$$kmod_i = \sum_{j \in D} tf(i,j) \times \left(1 + \frac{docs(i)}{|D|}\right) \times Ln\left(\frac{|users|}{users(i)} + 1\right) \quad (2)$$

donde:

- $tf(i,j)$ = frecuencia del término i en el documento j.
- $docs(i)$ = número de documentos en que el término i aparece.
- $|D|$ = número total de documentos.
- $|users|$ = número total de usuarios.
- $users\ (i)$ = número de usuarios cuya pre-ontología tiene indexado el término i.

En este caso, el valor calculado no se ofrece directamente como relevancia, sino que se normaliza dividiéndolo por la media de los valores obtenidos. De este modo, el perfil del usuario contiene los términos con mayor relevancia (del 75% al 100%).

3.4. CONSTRUCCIÓN DE LA ONTOLOGÍA DEL USUARIO

En esta etapa, la construcción automática de la ontología del usuario tiene por objeto generar una clasificación jerárquica de los términos indexados en la pre-ontología. La ontología, en este contexto, puede ser considerada como un conjunto de árboles relacionados donde cada nodo representa un ítem. En estos árboles, el nodo *A* es descendiente del nodo *B* si el ítem *A* es *"más concreto que"* (*narrower than*) el ítem *B*, o si el ítem *B* es *"más general"* (*broader than*) el ítem *A*. La propuesta para generar automáticamente ontologías borrosas, se basa en el algoritmo propuesto en [7]. Este algoritmo permite calcular medidas borrosas del grado de generalidad (*GD – Generality Degree*) entre parejas de palabras extraídas de una colección de documentos.

Formalmente, se define $C = (d_1, d_2, ..., d_n)$ como el conjunto de documentos d_js, donde cada documento $d = (t_1, t_2, ..., t_m)$ está formado por un conjunto de términos t_i. Dado un término t_i, la ocurrencia de t_i en un documento d se representa por $ocurr(t_i,d)$ y su valor se obtiene de acuerdo con la siguiente fórmula:

$$ocurr(t_i,d) = \frac{f(t_i,d)}{f_{\max}(d)} \quad (3)$$

donde:

- $f(t_i,d)$ = frecuencia de ocurrencia del término t_i en el documento d.
- $f_{max}(d)$ = numero de ocurrencias del término más frecuente del documento d.

La relación $GD(t_i, t_j)$ expresa que t_i es *"más concreto"* que t_j. El grado de esta relación está definido por:

$$GD(t_i,t_j) = \frac{\sum_{d \in D} ocurr(t_i,d) \otimes ocurr(t_j,d)}{\sum_{d \in D} ocurr(t_i,d)} \quad (4)$$

donde $\otimes$ se refiere a un operador de conjunción borroso. En este trabajo han sido implementadas seis t-normas conocidas que ejercen esta función. Son las siguientes:

1) Mínimo(*Mi*): $\min\{x,y\}$
2) Producto de Hamacher (*PH*): $\dfrac{x \cdot y}{x + y - (x \cdot y)}$
3) Producto Algebraico (*PA*): $x \cdot y$
4) Producto de Einstein(*PE*): $\dfrac{x \cdot y}{2 - [x + y - (x \cdot y)]}$
5) Producto Acotado(*PC*): $\max\{0, x + y - 1\}$
6) Producto Drástico(*PD*): $\begin{cases} x & if\ y = 1 \\ y & if\ x = 1 \\ 0 & otherwise \end{cases}$

Para construir la ontología de usuario, primero se calcula el grado de generalidad (*GD*) para una pareja de términos distintos de la pre-ontología, es decir, $GD(t_i, t_j)$ y $GD(t_j, t_i)$. A continuación, se elimina el valor más pequeño. El valor mayor se utiliza para generar la relación que se va a incluir en la ontología. El valor más pequeño se descarta para eliminar relaciones redundantes entre los términos. Con esta estrategia de decisión, se elije la instancia positiva del término si la diferencia entre $GD(t_i, t_j)$ y $GD(t_j, t_i)$ es grande, o se elije la relación más fuerte si la diferencia es pequeña. El proceso se repite para cada pareja de términos de la pre-ontología de todos los usuarios. Finalmente, se eliminan de la ontología las relaciones cuyo valor de GD es menor que un determinado umbral, dado que se considera información no relevante. El valor del umbral se ha definido empíricamente en 0.6, tras análisis de las relaciones generadas.

La ontología borrosa generada contiene una descripción taxonómica de las relaciones *GD*, y su estructura se muestra en la Figura 3.

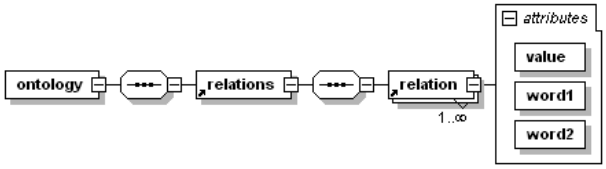

Figura 3: Estructura de la Ontología.

Como se puede observar en la Figura 3, la ontología está compuesta por un conjunto de relaciones. Cada relación tiene dos identificadores (word1 y word2) que representan los términos que se relacionan, y un cuantificador (value) que indica el grado o valor de la relación.

3.5. ACTUALIZACIÓN DEL PERFIL DE USUARIO

Con el objetivo de que el proceso de construcción del perfil de usuario sea más dinámico, es necesario permitir que se pueda incorporar información de nuevos objetos de aprendizaje seleccionados o creados por el usuario directamente en el perfil, si éste ya existe. Dicho procedimiento debe reproducir los mismos resultados que se obtendrían si los objetos de aprendizaje estuvieran disponibles en el inicio del proceso de construcción del perfil.

Para realizar dicha tarea, los nuevos documentos son previamente sometidos a las rutinas de pre-procesamiento lingüístico e indexación, de acuerdo con los procesos descritos en las secciones 2.1 y 2.2. Como resultado, se genera una nueva pre-ontología. El próximo paso consiste en calcular el peso de los términos de la pre-ontología, de acuerdo con las métricas descritas en la sección 2.3 y, posteriormente, actualizar los términos relevantes del usuario. Por último, se actualiza la ontología del usuario siguiendo un procedimiento semejante al de su construcción. Primero se calcula el Grado de Generalidad (*GD*), relacionando cada uno de los términos de la nueva pre-ontología con todos los demás términos, lo que genera como resultado un conjunto de relaciones borrosas. Posteriormente, se usan estas relaciones para actualizar la ontología, de acuerdo con las siguientes reglas: 1) Si la relación generada no existe en la ontología del usuario, entonces se incorpora esta nueva relación. 2) Si la relación ya existe, entonces se actualiza su valor o se excluye la relación, si este es el caso.

4 PERFILES DE USUARIO EN AGORA

En este apartado se presenta el estudio experimental realizado con el fin de representar el perfil de usuarios de la plataforma AGORA, a partir de los documentos proporcionados por los usuarios durante los últimos meses en el servidor situado en la Universidad de Castilla-La Mancha. El sistema cuenta con algo más de 200 usuarios de distintos niveles educativos, instituciones y países.

Únicamente se ha construido el perfil de los usuarios con un número de documentos aportados suficientemente significativo, en concreto más de 5, ya que los propios administradores de AGORA consideran que con una aportación menor, los perfiles generados no representarán correctamente a un usuario concreto. Para este caso se seleccionaron 40 usuarios, quienes han aportado 180 documentos de diversa naturaleza y formato (HTML, PDF, etc.).

En la Tabla 1 se puede observar el conjunto de términos obtenidos como relevantes por parte del algoritmo explicado en el apartado anterior sobre los documentos asociados al usuario 174. Este usuario es el que posee más documentos colocados en el sistema (20 documentos).

Tabla 1: Términos más relevantes para un usuario.

Término	***Peso***	***Término***	***Peso***
aprendizaje	15,42	aplicación	7
dato	10,64	biblioteca	5,68
objeto	9,51	estudiante	5,5
información	9,1	instructor	4,83
objeto aprendizaje	7,22	patrón	4,79
sistema	7,1	alumno	,47

Este usuario ha indicado que los términos seleccionados por el algoritmo incluyen a todos los relevantes a su juicio (100% de exhaustividad) aunque también incluyen algún término demasiado general, por ejemplo, *dato* o *información* (75% en precisión). Los resultados son similares (90% en exhaustividad y 70% en precisión) en cuanto a las relaciones más importantes entre conceptos que se reflejan en la Figura 4. Para el resto de usuarios los resultados han sido similares, con un alto grado de exhaustividad (89,1%) y un grado algo menor de precisión (74,8%).

Por último, en la Tabla 2 se muestran las estadísticas sobre el número de relaciones de las ontologías de los usuarios más significativos generadas según los diferentes operadores de conjunción borrosos. Aquí ***Mi*** es el Mínimo, ***PH*** es el Producto de Hamacher, ***PA*** es el Producto Algebraico, ***PE*** es el Producto de Einstein, ***PC*** es el Producto Acotado y ***PD*** es el Producto Drástico.

Tabla 2: Estadísticas de las Ontologías de los Usuarios.

Usuario	***Mi***	***PH***	***PA***	***PE***	***PC***	***PD***
139	253	187	0	0	0	0
154	946	789	0	0	0	0
158	15952	10539	250	169	72	68
166	9662	3619	321	218	104	92
170	1212	745	26	23	20	20
171	5089	2484	134	124	107	104
172	872	341	81	75	39	38
173	12562	5485	339	278	169	165
174	18688	6658	430	342	221	186

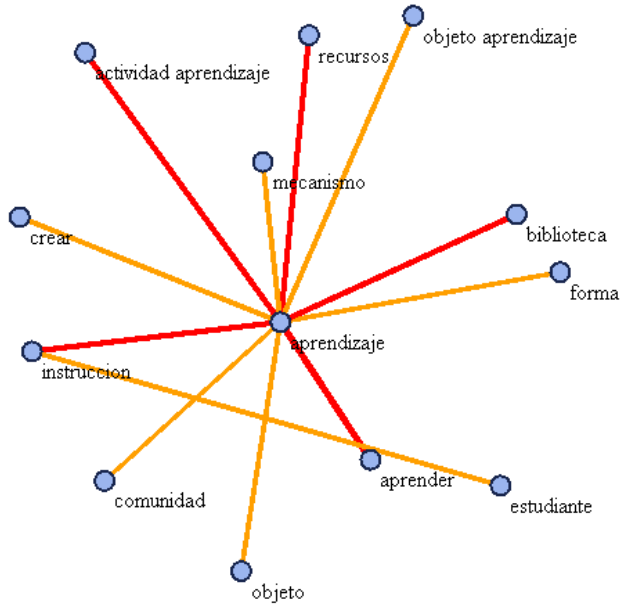

Figura 4: Relaciones más relevantes entre conceptos.

5 CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se ha propuesto la utilización de ontologías borrosas con el fin de representar perfiles de usuario. Esta aproximación ha sido probada en la herramienta de gestión de objetos de aprendizaje AGORA, cuyos usuarios y expertos han validado que el perfil representa de forma confiable sus temas de interés y preferencias de trabajo.

Como líneas a desarrollar en el futuro se propone la experimentación con un número mayor de usuarios y de documentos, dentro de la misma plataforma. Esto será posible ya que el sistema está siendo adoptado por varias universidades y redes de investigación en Iberoamérica.

Por otro lado, nos planteamos mejorar la calidad del perfil de usuario mediante la realización de procesos de poda que eviten la aparición de términos pocos significativos dentro del perfil, así como incorporar la información proporcionada por el usuario en su descripción.

Agradecimientos

La investigación realizada está financiada por los proyectos TIN2007-67494 del Ministerio de Ciencia e Innovación, PEIC09-0196-3018 de la Junta de Comunidades de Castilla-La Mancha y A/016625/08 de la Agencia Española de Cooperación Internacional para el Desarrollo (AECID), así como por el Consejo Nacional de Ciencia y Tecnología CONACYT, México y el Consejo de Ciencia y Tecnología del Estado de Yucatán CONCyTEY, México.

Referencias

[1] K.R. Beesley. Language Identifier: A Computer Program for Automatic Natural-Language Identification of On-Line Text. *Languages at Crossroads: Proceedings of the 29th Annual Conference of the American Translators Association.* Pág. 47-54, 1998.

[2] T. Korenius, et al.. Stemming and lemmatization in the clustering of finnish text documents, *in Proceedings of the thirteenth ACM international conference on Information and knowledge management.* ACM: Washington, D.C., USA. Pág. 625-633, 2004.

[3] R.R. Korfhage. Information storage and retrieval. *John Wiley & Sons, Inc.* Pág. 349, 1997.

[4] N. McLean, C. Lynch. Interoperability between Information and Learning Environments: Bringing the Gaps. *A joint White Paper on behalf of the IMS Global Learning Consortium and the Coalition for Networked Information.* Disponible en http://www.imsglobal.org/dlims_white_paper_publicd publ_1.pdf. Pág. 13, 2003.

[5] M.E. Prieto, et al.. A Recommender System Architecture for Instructional Engineering. *In Proceedings of the 1st World Summit on the Knowledge Society: Emerging Technologies and information Systems For the Knowledge Society.* LNIAI, vol. 52 (88). Springer, Heidelberg. Pág. 314-321, 2008.

[6] F. Romero et al.. Automatic extraction of the main terminology used in empirical software engineering through text mining techniques. *In Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement.* ACM: Kaiserslautern, Germany. Pág. 357-358, 2008.

[7] D.H. Widyantoro and J. Yen. Incorporating fuzzy ontology of term relations in a search engine. *In BISC International Workshop on Fuzzy Logic and the Internet.* Berkeley: University of California. Pág. 564-576, 2001.

[8] F. Xu et al. A Domain Adaptive Approach to Automatic Acquisition of Domain Relevant Terms and their Relations with Bootstrapping. *In Proceedings of LREC 2002, the 3rd International Conference on Language Resources an Evaluation.* Las Palmas, Spain. 2002.