

SOAF: un sistema de indexado semántico de OA basado en las anotaciones colaborativas

Doina Ana Cernea¹, Esther Del Moral², Emilio Labra³

¹Departamento de Matemáticas – Universidad de Oviedo

²Departamento de Ciencias de la Educación – Universidad de Oviedo

³Departamento de Informática – Universidad de Oviedo
{cerneadoina,emoral,labra}@uniovi.es

Abstract. Actualmente, la producción de OA (Objetos de Aprendizaje) y materiales educativos en formato digital, así como su almacenamiento en repositorios cada vez más extensos, es una práctica generalizada en todas las instituciones dedicadas a la formación. Sin embargo, las herramientas de búsqueda y localización efectiva de dichos contenidos, en base a su etiquetado, todavía no poseen una gran consistencia. Este artículo propone una arquitectura para un sistema capaz de indexar semánticamente los OA de un repositorio, SOAF (Semántica de OA basada en *Folksonomías*), la cual combina las técnicas de extracción automática de información con las tecnologías de etiquetado colaborativo. De modo que, la meta-información de los OA, así obtenida, aporte el valor añadido derivado de la práctica de las comunidades de usuarios, capaces de compartir sus experiencias a través de anotaciones específicas que sirven para identificar cada OA y contribuyen a optimizar su reusabilidad en nuevos contextos de aprendizaje.

Keywords: Indexado semántico, *Gap* semántico, Objetos de Aprendizaje, anotaciones de los usuarios, aprendizaje colaborativo, *folksonomías*.

1 Introducción

Para garantizar la reutilización de los objetos y recursos educativos en nuevos contextos de aprendizaje se precisa desarrollar sistemas, cada vez más sofisticados y basados en el uso de semántica, con objeto de facilitar la gestión y categorización de las extensas colecciones de OA y recursos de aprendizaje contenidos en la multiplicidad de repositorios que existen, tomando como punto de partida las descripciones cualitativas estimadas por los usuarios, para que respondan con mayor fiabilidad a lo que en realidad quieren referirse.

Sin embargo, la disponibilidad de tecnologías para la descripción de los recursos *Web* basada en conceptos se encuentra todavía en un estado poco avanzado. La pluralidad de formatos de recursos de aprendizaje exige técnicas específicas que hagan posible su descripción detallada.

Si bien, el análisis de recursos de carácter textual puede llevarse a cabo a través de la asignación consensuada de *palabras claves* para facilitar sus búsquedas posteriores [31] en el caso de los recursos multimedia, se percibe un gran desajuste, llamado *gap semántico*, entre las características de bajo nivel extraídas de forma automática (como parámetros de color, texturas, contornos,...) y la descripción de su contenido basada en conceptos o características de alto nivel. Por esta razón, es necesario recurrir a nuevos métodos de indexado semántico más apropiados para este tipo de recursos.

Existen numerosos intentos para añadir descripción semántica a los recursos multimedia [22,30]. Una de las primeras incursiones para utilizar el formato MPEG-7 con la finalidad de almacenar metadatos fue realizada por Hunter en 2001 [15], que desarrolló inicialmente una ontología RDFS enriquecida con DML-OIL, más tarde enriquecida con OWL. Como una parte del proyecto aceMedia [6], se desarrolló una ontología para descriptores visuales de características de bajo nivel. Estas características fueron incorporadas en la parte visual de MPEG-7.

Algunas técnicas propuestas para la extracción automática de semántica parten de la división de una imagen en regiones y la asignación de una etiqueta a cada una de las regiones utilizando métodos estadísticos [10,27], algoritmos EM [9,27] o recientes modelos probabilísticos como Cross Relevance Model [16]

Por otra parte, algunos sitios Web como Amazon [19], implementaron algoritmos de filtrado colaborativo que tienen en cuenta la semejanza entre usuarios para ofrecer recomendaciones de productos. Otros sitios Web permiten a sus usuarios el etiquetado colaborativo de los distintos tipos de recursos, método que ha adquirido gran popularidad recientemente como sistema de categorización [7,20].

Entre las fórmulas que apuestan por un tipo de etiquetado colaborativo se encuentran aplicaciones de internet como *Delicio.us*, *Tecnocrati* y *Flickr* [20], las cuales permiten a los usuarios la anotación de un recurso concreto como una página *web*, un *blog post*, una imagen, etc., a partir de un conjunto de palabras libremente elegidas ("*tags*"), lo cual facilita su reutilización y uso compartido.

Recientemente se ha propuesto una combinación entre el trabajo colaborativo y la recuperación y clasificación de imágenes, introduciendo el concepto de *Navigation Map* [1,2], que vincula usuarios, etiquetas y datos. Este sistema es capaz de dotar de distintos significados a la misma etiqueta haciendo uso de características visuales extraídas de forma automática de una imagen.

En el presente artículo, se propone una arquitectura para un sistema SOAF que integra las anotaciones colaborativas de los usuarios en la descripción semántica de los OA y

de los recursos de aprendizaje que los componen, mejorando de este modo la semántica emergente de estos recursos.

SOAF se define como un sistema que combina las técnicas de recuperación de información basadas en la extracción automática de semántica, con las anotaciones específicas realizadas por los diferentes usuarios que componen las comunidades de aprendizaje que los usan.

2 Redes sociales dentro de la Web Semántica

Las redes sociales son representaciones de las relaciones que existen entre individuos y grupos de individuos dentro de una comunidad. El análisis de los vínculos sociales y las estructuras que se crean ha encontrado aplicaciones en distintos dominios científicos como sociología, psicología, informática, etc.

El concepto de red social designa una estructura de nodos, generalmente personas u organizaciones (llamados también *actores*), y aristas que indican el tipo y el rango de la conexión que existe entre ellos [4].

Las aplicaciones basadas en Internet que utilizan las redes sociales, conocidas como redes sociales virtuales, proporcionan nuevas formas de socialización para encontrar amigos, socios de negocios o, simplemente, poner en contacto a personas con aficiones comunes. Ese es el caso de *Friendster*, una red social que utiliza las técnicas propias de los Círculos de Amigos generados dentro de las comunidades virtuales, la cual pretende demostrar que todas las personas del mundo pueden estar relacionadas entre sí a través de una cadena social de conocidos relativamente corta.

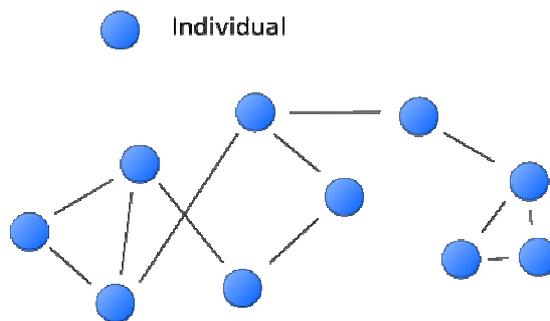


Fig. 1. Esquema de la estructura de una red social

2.1 Tecnologías de la Web Semántica

La Web Semántica surge como una visión de futuro de la Web que implica dotar de significado los recursos publicados para que sus contenidos puedan ser leídos por los ordenadores, no sólo para su presentación sino también para la automatización de tareas, integración y reutilización a través de varias aplicaciones [5].

Actualmente, la información publicada en Internet está concebida para el entendimiento humano, los navegadores siendo capaces de reconocer el lenguaje de marcado HTML para visualizar el contenido de los recursos Web en pantalla. Pero la interpretación del contenido de estos recursos y su gestión queda una tarea muy difícil de automatizar.

La Web Semántica implica la creación de tecnologías para la descripción de los datos contenidos dentro de las páginas Web. Un vocabulario propuesto es el RDF, que habilita la descripción de los recursos y de las relaciones que existen entre ellos.

Ejemplo: código RDF para la descripción semántica de un OA sobre Geometría Fractal:

```
<rdf:RDF xmlns:rdf="..."
... >

<rdf:Description rdf:about="http://.../geofractal/estruFractal">
  <rdf:type rdf:resource="http://...#Introduction"/>
  <dc:title>Introducción a la Estructura Fractal</dc:title>
  <dc:creator>
    <lom:entity> <vCard:FN>J. Garrido</vCard:FN></lom:entity>
  </dc:creator>
</rdf:Description>

</rdf:RDF>
```

Otra tecnología clave en el desarrollo de la Web Semántica es la asociada a la representación del conocimiento de un dominio concreto, es decir, la definición de los conceptos y sus relaciones. OWL se ha desarrollado para proporcionar un lenguaje común para la definición de las ontologías y está basada en la lógica descriptiva. El desarrollo descentralizado de ontologías facilitará la integración de datos de muy diversas proveniencias y distintos tipos de aplicaciones. Mediante OWL es posible declarar, por ejemplo que una explicación (*Introduction*) es un tipo de Lección (*lecture*) y que puede contener varios ejercicios:

```
http://...#Introduction  $\subseteq$  http://...#Lecture
http://...#Introduction  $\subseteq \exists$  http://...#contains http://...#Exercise
```

2.2 Sistemas de etiquetado colaborativo: *folksonomías*

Las *folksonomías* se definen como sistemas colaborativos para la categorización no jerárquica de los recursos de Internet a través de etiquetas compartidas por una comunidad o red social.

Al realizar una *folksonomía*, varios usuarios colaboran en la descripción de un espacio informativo para establecer una lista de palabras clave, o un vocabulario “no controlado”, que puede servir para categorizar o etiquetar determinados recursos. La diferencia entre una taxonomía convencional y una *folksonomía* estriba en que esta última se realiza a partir del intercambio de opiniones y matizaciones de varios usuarios no de un único experto, con el consiguiente enriquecimiento en la descripción del objeto del que se trate.

El *social tagging* (de *tag*, marca o etiqueta social) genera una navegación basada en la exploración que realizan los usuarios al interactuar con determinados recursos u objetos de aprendizaje. Los *social tags* asignados a los recursos sirven para clasificarlos, ordenarlos, buscarlos y encontrar información de distinto tipo vinculada a ellos. Este sistema habilita a los diferentes usuarios para añadir palabras clave, -suficientemente descriptivas-, a recursos de internet como páginas *web*, imágenes, vídeos,...y, en particular, a recursos de aprendizaje, para que puedan ser reutilizados y compartidos por otros.

Desde este punto de vista, la utilización de etiquetas colaborativas, generadas por los usuarios de una misma comunidad virtual, mejora la gestión de los recursos debido a la estructura social que se crea, al tiempo que se producen nuevas modalidades de comunicación social, intercambio y oportunidades de minería de datos.

Los sistemas de etiquetado colaborativo permiten que los usuarios compartan las etiquetas de los recursos y, a su vez, cada etiqueta puede enlazar a nuevos recursos adicionales que otros usuarios hayan etiquetado de la misma forma [20]. Así pues, estos sistemas se van construyendo sobre unas estructuras sociales emergentes que comparten los mismos parámetros conceptuales y lingüísticos de la comunidad de usuarios a la que pertenecen.

En este sentido, las etiquetas adoptadas por los usuarios establecen un vínculo entre éstos y los recursos. Paralelamente, los recursos pueden estar conectados entre sí a través de enlaces (sus propias etiquetas sirven de enlace con otros recursos etiquetados del mismo modo) y, a su vez, los usuarios pueden responder a ciertas afiliaciones sociales, al proceder de comunidades virtuales que comparten los mismos intereses... Esta compleja red de relaciones se va incrementando a partir de los nuevos nexos que surgen al compartir recursos y etiquetas.

La siguiente figura pretende ser la representación conceptual de una *folksonomía*, donde los usuarios asignan etiquetas a recursos concretos y donde pueden establecerse todo tipo de asociaciones.

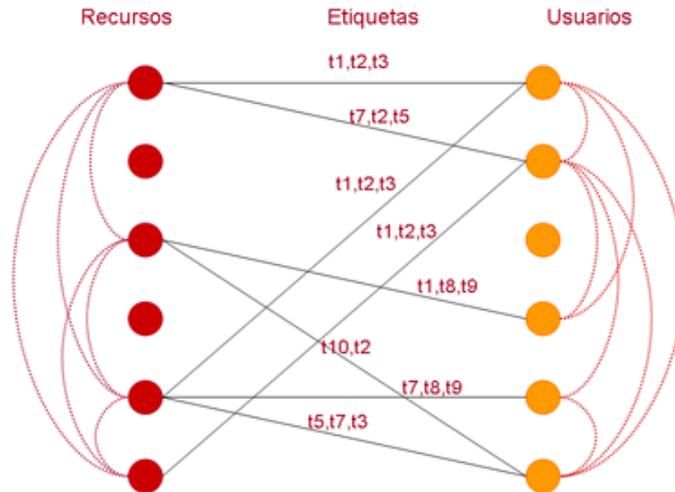


Fig. 2. El modelo conceptual de un sistema de etiquetado colaborativo [20]

2.3 Aprendizaje colaborativo potenciado por la Web Semántica

En el contexto creado por la Web 2.0 surge el nuevo paradigma de la colaboración denominado *software social*, cuyas aplicaciones como las redes sociales, *wikis* y *folksonomías* han encontrado una gran aceptación dentro de las comunidades de *e-Learning*, potenciando el aprendizaje colaborativo, basado en la construcción conjunta del conocimiento.

Cada vez son más frecuentes las propuestas de diseño de entornos o comunidades virtuales que favorecen el aprendizaje desde una perspectiva colaborativa, introduciendo formas de trabajo alternativas que subrayan la dimensión social del conocimiento. Por ello, en estas comunidades se favorecen los procesos de interacción y de resolución conjunta de problemas, convirtiéndose en espacios sociales [25] colaborativos, donde el uso de los OA de un repositorio puede contribuir a contextualizar el aprendizaje y a dotarle de mayor significado.

Dentro de estas comunidades virtuales se establecen fuertes vínculos entre los miembros que las constituyen, orientados al logro de objetivos comunes que las hacen muy poderosas, afianzando sus relaciones internas y materializándose en la creación de importantes *redes sociales*, que adquieren un valor cada vez mayor dado que se apoyan en las ventajas que aportan cada uno de los miembros que las componen [26].

En este marco, el aprovechamiento de las sinergias, -derivadas de la integración de las anotaciones colaborativas de los usuarios-, para etiquetar e identificar los OA que comparten, será la clave para favorecer su reusabilidad. Esta es, precisamente, la filosofía del “etiquetado social” que define a las *folksonomías*, y el principio que sustenta la arquitectura del sistema que se describirá más adelante.

3 Indexado semántico de los recursos Web educativos

Desde la perspectiva de la Web Semántica, los recursos de Internet dejan de ser unidades de información, permitiéndose la descripción de los mismos con un grado más fino de granularidad. Por ejemplo, de la descripción semántica del OA “Introducción a la Geometría Fractal” se podría extraer la definición del concepto de estructura fractal o la propiedad de recursividad. Así, en el contexto de la Web Semántica los recursos *Web* en general y los OA en particular, deben estar acompañados de las descripciones semánticas de sus contenidos.

En la actualidad, el proceso de añadir semántica a los recursos *Web* requiere de las valoraciones más precisas de las personas, sobre todo tratándose de recursos multimedia como imágenes, vídeos, sonidos, etc. Cuya descripción puede verse afectada por la carga subjetiva que cada usuario pueda incorporar. De ahí que se precisen soluciones que suplan esta eventualidad. La nuestra pasa por definir un sistema específico que permita automatizar el proceso de indexado semántico de los OA, combinando la semántica emergente con las anotaciones colaborativas de los usuarios.

3.1 OA en la Web Semántica

Los OA definidos como unidades mínimas con significado por si mismas, constituidas por recursos de información multiformato y con carácter interactivo, orientados al logro de un único objetivo de aprendizaje, y reutilizables en distintos contextos educativos, los cuales poseen además una estructura granular flexible [24]:

En un primer nivel los apartados (Objetos de Información)

En un segundo nivel los recursos que los componen (texto, multimedia, etc.)

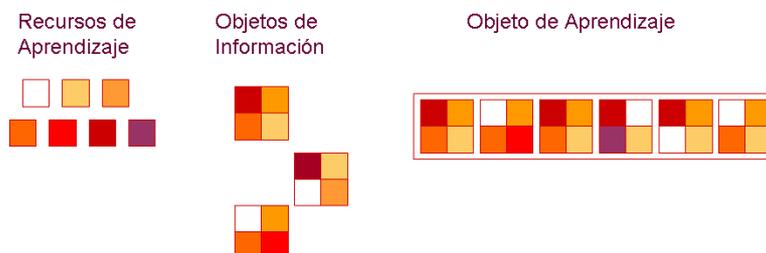


Fig. 3. Niveles de granularidad de un OA [8]

En el caso de los recursos textuales, la complejidad asociada al lenguaje natural cobra especial relevancia cuando se necesita recuperar información sobre su contenido. De este modo, cada documento se tiene que describir mediante un conjunto de términos que facilite su identificación.

Para este tipo de recursos de aprendizaje, SOAF propone el indexado de colecciones de documentos mediante técnicas de procesamiento del lenguaje natural, muy utilizadas por los buscadores de Internet, que suponen la división del texto en una lista de términos que, más tarde, son normalizados utilizando técnicas de radicación, para eliminar las ambigüedades lingüísticas y las palabras vacías de contenido, minimizando el grado de arbitrariedad así [31].

Las colecciones de OA y recursos multimedia de diversa índole que componen los múltiples repositorios institucionales se han multiplicado exponencialmente. Por ello, la identificación de dichos recursos multimedia adquiere una gran importancia, y requiere de técnicas de extracción de información semántica específica, es decir, precisa de unos metadatos suficientemente descriptivos relativos a los conceptos que éstos incluyen, que les convierta en fiables. Y, esta tarea es algo compleja.

A diferencia de los recursos textuales, los multimedia (imágenes, vídeos, animaciones, sonidos,...) son más difíciles de analizar mediante métodos automáticos, por ello, se propone la utilización de un sistema que integre las tecnologías de las *folksonomías* que complementen las descripciones de los OA.

3.2 El Gap Semántico

Existe un alejamiento entre los métodos automáticos de análisis de los recursos multimedia y la semántica que contienen, conocido como *gap semántico*, definido como la ausencia de coincidencias entre la información extraída automáticamente por ordenador, y la derivada de la percepción humana de las imágenes basada en *conceptos de alto nivel (high-level concepts)* [14]. El estado del arte de los sistemas de análisis de recursos multimedia limita su alcance a descripciones apoyadas en las

características de bajo nivel (low-level features) propias de los contenidos, relativas a sus atributos de color, texturas, bordes, etc.

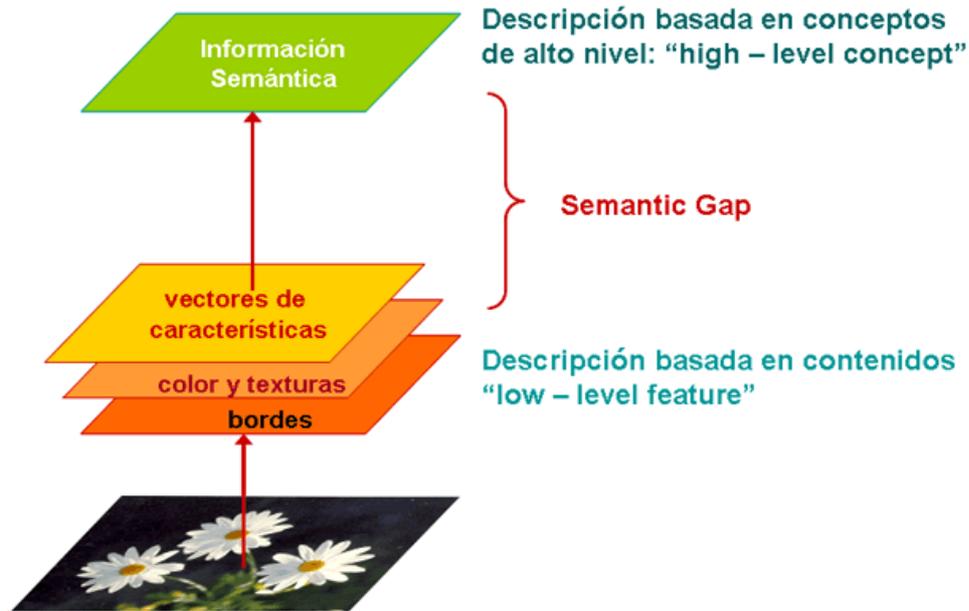


Fig. 4. El Gap Semántico

Las investigaciones orientadas a minimizar el *gap semántico* refieren sus resultados a las anotaciones y al marcado semántico de recursos textuales fundamentalmente, dejando la extracción automática de semántica de los multimedia como un problema abierto, dada su complejidad. Por ello, más adelante, se propone un sistema que intenta disminuir esta brecha haciendo uso de las anotaciones de los usuarios.

4 Arquitectura SOAF: Integración de las anotaciones colaborativas de los usuarios con la semántica emergente

La arquitectura propuesta va orientada a clasificar los diferentes OA contenidos en un repositorio a partir de las etiquetas generadas colaborativamente por los usuarios. SOAF utiliza un sistema basado en *folksonomías* para capturar las anotaciones de los propios usuarios focalizadas en sus experiencias de aprendizaje con un determinado recurso. Estas etiquetas se procesarán para obtener un conjunto de términos conectados con una red de ontologías. SOAF viene así a incrementar la fiabilidad del etiquetado de los métodos automáticos, incorporando las matizaciones de los usuarios

de una comunidad virtual de aprendizaje como anotaciones específicas para cada recurso.

Las etiquetas generadas colaborativamente aportan a la semántica emergente de los recursos *web* un valor añadido, que constituye una solución óptima para lograr descripciones semánticas más operativas. Por una parte, las anotaciones de los usuarios vienen a completar y a afinar la semántica extraída de forma automática, y por otra, la semántica emergente matiza y resuelve algunos problemas de las etiquetas de los usuarios como los errores de ortografía, uso de homónimos, sinónimos, reduciendo la ambigüedad terminológica, etc.

4.1 Extracción automática de semántica con SOAF

El sistema SOAF propone una arquitectura para la extracción automática de descripción semántica de los recursos educativos multimedia basada en *Latent Semantic Indexing* [28] a partir de la representación de los recursos (texto o multimedia) en vectores del espacio vectorial \mathbf{R}^n , y aplicando análisis de matrices establece las asociaciones y las conexiones entre los diversos recursos de un repositorio.

A pesar de todo, los métodos automáticos de indexado semántico no garantizan la total fiabilidad en el etiquetado final, ya que pueden introducir sesgos en la información extraída. Por lo que, se hace preciso acudir a otros métodos complementarios que permitan matizar en profundidad dicha información semántica.

Desde esta perspectiva, el sistema SOAF contempla y procesa tres tipos de meta-información que pueden acompañar a los OA o a los recursos que los componen:

La semántica proporcionada de forma automática mediante indexado semántico a partir de las características de bajo nivel

Los descriptores de alto nivel proporcionados por los autores (en la mayoría de los casos se refieren a título, fecha de creación, nombre de los autores, tema, objetivos de aprendizaje,...)

La semántica aportada por los usuarios mediante anotaciones colaborativas.

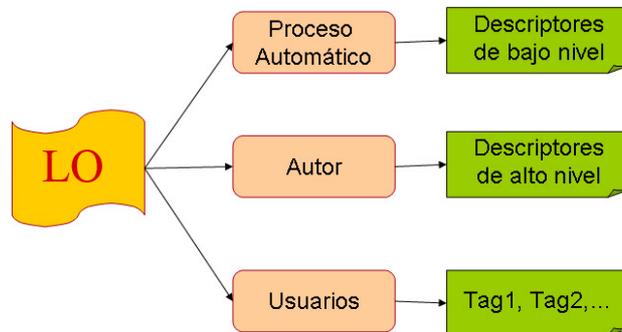


Fig. 5. Los tres tipos de meta-información de un OA, según su procedencia.

4.2 Etiquetas generadas colaborativamente

Sin duda, una poderosa fuente de información semántica para los OA de un repositorio son las descripciones aportadas por los propios usuarios de forma espontánea, mediante un sistema de etiquetado o anotaciones específicas que engloba nuevos conocimientos basados en sus experiencias reales compartidas. SOAF utiliza estas anotaciones colaborativas para completar y matizar el indexado semántico generado automáticamente por ordenador a partir de las características de *bajo nivel*. Aunque las anotaciones colaborativas son, en general, etiquetas libres, es decir, cadenas de caracteres elegidas libremente por los usuarios, el presente sistema sugiere al usuario etiquetas más o menos homogéneas o similares para favorecer la convergencia.

En los mencionados desarrollos potenciados por la Web Semántica como *Flickr*, los usuarios comparten recursos y etiquetas. Del mismo modo, la arquitectura SOAF partiendo de esa filosofía, propone la utilización de etiquetas sociales, que cobran consistencia y relevancia dentro de la comunidad de usuarios, al retroalimentarse constantemente, observándose que, con el tiempo, la frecuencia relativa de uso de una misma etiqueta tiende a ser un valor constante, lo que explica que el etiquetado colaborativo posee una mayor capacidad para coordinar las acciones de los usuarios y crear anotaciones convergentes, más coherentes para los recursos que comparten.

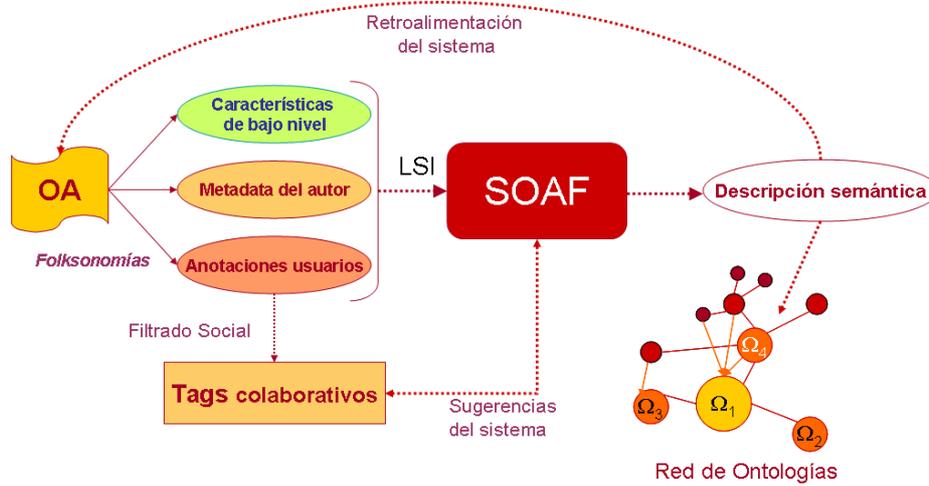


Fig. 6. Arquitectura del Sistema SOAF

4.3 Filtrado social de las anotaciones colaborativas de los usuarios

Finalmente, el conjunto de etiquetas asignadas por los usuarios se pueden filtrar aplicándoles un *test* de confianza basado en la similitud de perfiles para garantizar su fiabilidad, pertinencia y validez [29]. De modo que, el sistema crea y almacena los perfiles de los usuarios a partir de sus intereses y sus valoraciones subjetivas. Posteriormente, compara dichos perfiles y establece sus similitudes a través de fórmulas estándar de correlación estadística. Mediante el filtrado social el sistema “recomienda” una etiqueta si varios usuarios con perfiles similares coinciden en su asignación, es decir, se le dota de un mayor valor de confianza a la etiqueta más utilizada por usuarios con perfiles semejantes.

5 Conclusiones

La incorporación de semántica a los recursos educativos es, en la actualidad, un problema abierto. Este artículo propone una arquitectura SOAF que permite el indexado semántico de los OA de un repositorio combinando técnicas de extracción automática de información basado en Latent Semantic Indexing con las tecnologías de etiquetado colaborativo. El conjunto de etiquetas asignadas por los usuarios de una a de aprendizaje se filtran mediante un test de confianza a partir de la semejanza de perfiles con el objetivo de mejorar su fiabilidad, pertinencia y validez. El sistema, asimismo, recomienda una etiqueta si varios usuarios se consideran semejantes.

Mediante al tecnología que aquí se presenta, la meta-información de los OA obtenida incorpora el significado deducido de la práctica de las comunidades de usuarios, logrando una mejor identificación de los OA que contribuye a optimizar su reusabilidad en contextos de aprendizaje diversos.

6 Bibliografía

1. Aurnhammer M., Hanappe P., and Steels L. (2006): "Augmenting Navigation for Collaborative Tagging with Emergent Semantics". In *5th International Semantic Web Conference*, Athens, GA, USA, November 5-9, 2006, LNCS 4273.
2. Aurnhammer, M., Hanappe, P. and Steels. L. Integrating Collaborative Tagging and Emergent Semantics for Image Retrieval. Proceedings WWW2006, Collaborative Web Tagging Workshop, May 2006.
3. Barnard K., Duygulu P., Freitas N., Forsyth, D., Blei, and Michael I. Jordan, "Matching Words and Pictures", *Journal of Machine Learning Research*, Vol 3, pp 1107-1135, 2003.
4. Barnes, J.A.: "Class and Committees in a Norwegian Island Parish" *Human Relations*.1954; 7: 39-58
5. Berners-Lee, T.; Hendler, J.; Lassila, O. (2001): The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific May (2001)
6. Bloehdorn, S., Petridis, K., Saathoff, C., Simou, N., Tzouvaras, V., Avrithis, Y., Handschuh, S., Kompatsiaris, I., Staab, S., Strintzis, M. G., (2005). "Semantic annotation of images and videos for multimedia analysis". In: *2nd European Semantic Web Conference (ESWC 2005)*. Heraklion, Greece.
7. Cattuto, C., 2006. *Semiotic dynamics in online social communities*. The European Physical Journal C 46, 33–37.
8. Cisco Systems, Reusable Learning Object Strategy: Designing and Developing Learning Objects for Multiple Learning Approaches, (2003)
9. Duygulu, P., Barnard, K., Freitas, de, N., Forsyth, D.: "Object recognition as machine translation: learning a lexicon for a fixed image vocabulary", In Seventh European Conference on Computer Vision (ECCV), Vol. 4, pp. 97-112, 2002.
10. Fan, J., Gao, Y., Luo, H. and Xu, G.: "Automatic image annotation by using conceptsensitive salient objects for image content representation" In *SIGIR '04: Proceedings of the 27th annual international conference on Research and development in information retrieval*, pages 361–368. ACM Press, 2004.
11. Gracia, J. M. (2002), Algebra Lineal tras los buscadores de Internet. <http://www.vc.edu.es/campus/centros/farmacia/deptos-f/depme/profesor/gracia/buscap.pdf>
12. García, R., Celma, O., (2006). "Semantic integration and retrieval of multimedia metadata". In: *5th Knowledge Markup and Semantic Annotation Workshop*. Vol. 185. CEUR Workshop Proceedings, pp. 69–80.
13. Golbeck, J. and Hendler, J.: "Filmtrust: Movie recommendations using trust in web-based social networks". Proceedings of the *IEEE Consumer Communications and Networking Conference*, January 2006.
14. Hare, J. S., Lewis, P. H., Enser, P. G. B. and Sandom, C. J. (2006) Mind the Gap: Another look at the problem of the semantic gap in image retrieval. In Proceedings of Multimedia Content Analysis, Management and Retrieval 2006 SPIE Vol. 6073, pp. 607309-1, San Jose, California, USA. Chang, E. Y., Hanjalic, A. and Sebe, N., Eds.
15. Hunter, J., (2001). "Adding multimedia to the semantic web - building an MPEG-7 ontology". In: *International Semantic Web Working Symposium, SWWS*.

16. Jeon, J., Lavrenko, V., Manmatha, R.: "Automatic Image Annotation and Retrieval using Cross-Media Relevance Models", 26th Annual Int. ACM SIGIR Conference, Toronto, Canada, 2003.
17. Kumar C. A., Srinivas S.: "Latent semantic indexing using eigenvalue analysis for efficient information retrieval", Int. J. Appl. Math. Comput. Sci., 2006, Vol. 16, No. 4, 551-558
18. Labra, J. E., Ordóñez P., Cueva, J. M. (2006): "WESONET: Applying Semantic Web Technologies and Collaborative Tagging to Multimedia Web Information Systems", Computers in Human Behavior, Pendiente de publicación.
19. Linden, G., Smith, B., York, J., 2003. Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing 4 (1).
20. Marlow, C., Naaman, M., Davis, M., Boyd, D., (2006): "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, ToRead", *Proceedings of Hypertext 2006*, New York: ACM Press.
21. Metzler, D., Manmatha, R. (2004): "An Inference Network Approach to Image Retrieval", CIVR04(42-50).
<http://www.visionbib.com/bibliography/applicat800.html#TT56696>
22. Moenne-Loccoz, N., Janvier, B., Marchand-Maillet, S., Ric Bruno, (2005). An integrating framework for the management of video collections. In: First Workshop on Machine Learning and Multimodal Interaction. Vol. MLMI '04, Lecture Notes in Computer Science 3361. Martigny, Switzerland.
23. Monay F., Gatica-Perez D., "On Image Auto-Annotation with Latent Space Models", Proc. ACM Int. Conf. on Multimedia (ACM MM), Berkeley, 2003.
<http://citeseer.ist.psu.edu/monay03image.html>
24. Moral, del, M. E. , Cernea, D. A. "Diseñando Objetos de Aprendizaje como facilitadores de la construcción del conocimiento". II Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos Reutilizables (SPDECE) Barcelona, 19, 20 y 21 de Octubre de 2005
25. Moral, del, M. E. , Cernea, D. A. "Wikis, Folksonomías y Webquests: trabajo colaborativo a través de Objetos de Aprendizaje". III Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos Reutilizables (SPDECE) Oviedo, 2006.
26. Owen, M., Grant, L.; Sayers, S.; Facer, K. (2006): "Social software and learning". En: http://www.futurelab.org.uk/research/opening_education.htm. [Consultado en Abril de 2007]
27. Pan, J. Y., Yang, H. J., Duygulu, P., and Faloutsos, C.: "Automatic Image Captioning." In *Proceedings of the 2004 IEEE International Conference on Multimedia and Expo (ICME 2004)*, 2004.
28. Pecenovic, Z. (1997). *Image retrieval using latent semantic indexing*. Final year graduate thesis, AudioVisual Communications Lab, Ecole Polytechnique Federale de Lausanne, Switzerland, June 1997. <http://citeseer.ist.psu.edu/pecenovic97image.html>
29. Shardanand U., Maes P. (1995): "Social information filtering: algorithms for automating "word of mouth". In: *Conference on Human Factors in Computing Systems archive Proceedings of the SIGCHI conference on Human factors in computing systems table of contents*, Denver, Colorado, United States, Pages: 210 - 217 , ISBN:0-201-84705-1
30. Stamou, G., Kollias, S., (2005). *Multimedia Content and Semantic Web: Standards, Methods and Tools*. John Wiley and Sons.
31. Vallez, M. y Pedraza-Jimenez, R.: *El Procesamiento del Lenguaje Natural en la Recuperación de Información Textual y áreas afines* [on line]. "Hipertext.net", núm. 5, 2007. <<http://www.hipertext.net>> [Consulta: 12/05/2007]. ISSN 1695-5498
32. Wenger, E. (2001): *Comunidades de práctica. Aprendizaje, significado e identidad*. Barcelona: Paidós.

33. Xu, Z., Fu, Y., Mao, J., Su, D., May 2006. "Towards the semantic web: Collaborative tag suggestions". In: *Proceedings of workshop on collaborative Web tagging*. Edinburgh, Scotland.