

## Generación de Tesoros basado en Media Wiki

Ana Flores Cuadrado<sup>1</sup>, Eduardo Villoslada de la Torre<sup>1,3</sup>, y Alberto Peláez Gutiérrez<sup>2</sup>

<sup>1</sup> Telefónica Investigación y Desarrollo

Parque Tecnológico de Boecillo, Parcela 120, 47151. Boecillo. Valladolid, España

{anafc, evdlt}@tid.es

<sup>2</sup> Telefónica Soluciones

Parque Tecnológico de Boecillo, Parcela 120, 47151. Boecillo. Valladolid, España

{alberto javier.pelaez gutierrez}@telefonica.es

<sup>3</sup> Universidad de Valladolid, Edificio TIT.

Campus "Miguel Delibes" s/n, 47011. Valladolid, España

**Resumen.** La Wikipedia alberga a través de su estructura, de forma implícita, un conjunto de relaciones conceptuales muy detalladas y evolución dinámica. Con el objetivo de aprovechar esa información para mejorar los procesos de búsqueda en los sistemas de gestión del conocimiento, se ha definido un algoritmo que permita materializar la estructura conceptual de la Wikipedia y sus relaciones para construir un tesoro.

**Palabras clave:** Tesoros, SKOS, MediaWiki, Wikipedia, screen scrapper

### 1. Introducción

Los sistemas de gestión del conocimiento de una empresa tienen como objetivo compartir y difundir el conocimiento sobre tecnologías y procedimientos entre sus trabajadores. Sin embargo, el rendimiento que se obtienen de estos sistemas es mínimo. A menudo, estos sistemas son utilizados como simples repositorios de documentos, donde los usuarios difícilmente localizan la información usando buscadores basados en el uso palabras claves [23] o búsquedas por metadatos. Aunque el uso las tecnologías semánticas y en especial el uso de tesoros [27], puede mejorar la precisión de las consultas (recall), existen dificultades para su aplicación en el ámbito de las tecnologías de la información, debido a que no existe ningún vocabulario específico para este ámbito y que aunque los vocabularios más generales como *Wordnet* [21] incluyen términos de este dominio, su actualización no es lo suficiente dinámica para soportar la rápida aparición de nuevos conceptos.

Con la aparición del concepto de Web 2.0 [22], Internet se ha constituido como un vehículo de participación ciudadana que permite la creación y difusión del conocimiento de forma dinámica. Una de las herramientas Web 2.0 más populares es la Wikipedia [30], donde los usuarios de forma voluntaria se encargan de introducir información sobre los temas más novedosos, en distintos idiomas. En la Wikipedia, cada página o artículo representa un concepto, y la tecnología que soporta la Wikipedia permite relacionar de forma implícita los distintos conceptos, a través de la estructura de sus páginas.

El objetivo de nuestro trabajo es materializar la estructura conceptual de la Wikipedia, y de sus relaciones y construir con él un tesoro que pueda utilizarse en los sistemas de gestión del conocimiento para mejorar los procesos de búsqueda, por una parte, utilizando las relaciones definidas en el tesoro para expandir las consultas [15,14] y así mejorar la precisión en las consultas, y por otra utilizar este tesoro como parte de un mapa semántico que a modo de asistente vaya guiando a los usuarios en los procesos de búsqueda, y permitiéndoles navegar a través de los distintas relaciones entre conceptos con el objetivo de refinar las búsquedas [14].

En la Sección 2 se describen los principales aspectos que son la base tecnológica durante el proceso de construcción del tesoro. En la Sección 3 se presenta el proceso de mapeo de la estructura de la Wikipedia en base al análisis previo de ésta. En esta sección también se presenta la arquitectura de la aplicación así como el algoritmo utilizado en el proceso de mapeo. La Sección 4 presenta los trabajos relacionados y por último en la Sección 5 se concluye con las ventajas y limitaciones del trabajo presentado y las líneas trabajo futuras.

## 2. Base Tecnológica

El objetivo de esta sección es describir los aspectos tecnológicos que son necesarios conocer para comprender el análisis realizado de la Wikipedia y la tecnología que sustenta la representación de tesauros.

### 2.1. Tecnología Media Wiki

Se entiende por Wiki una herramienta web colaborativa, accesible mediante un navegador, cuyo contenido puede ser visualizado y editado directamente por cualquier usuario con permisos sin necesidad de tener unos conocimientos especializados. La primera Wiki fue creada en 1995 por Ward Cunningham [28]. Posteriormente en 2001 se usó este modelo para crear *Wikipedia* utilizando un software específico para realizar wikis: *MediaWiki* [18]. Dentro de *Wikipedia* se distinguen cinco tipos de elementos primarios [11,4]:

- **Páginas de Artículos:** texto que describe un concepto; es la entidad básica de MediaWiki.
- **Páginas de Re-dirección:** con el fin de asegurarse de que existe un único artículo por concepto, MediaWiki provee re-direcciones de otros artículos hacia el principal. De esta es la forma se maneja diferentes formas de un mismo término (por ejemplo: ortografías múltiples, seudónimos, mayúsculas/minúsculas, abreviaturas, sinónimos) [24].
- **Enlaces a artículos:** dirigen al usuario de un artículo a otro. Las etiquetas de los enlaces no tienen que ser iguales al título del artículo enlazado, pueden usarse sinónimos, términos relacionados y variaciones lingüísticas.
- **Categorías:** MediaWiki también ofrece la posibilidad a los autores de incluir categorías a modo de metadatos para la clasificación de artículos. Estas categorías pueden ser utilizadas para encontrar artículos relacionados.
- **Páginas de Desambiguación:** Páginas que contienen un conjunto de enlaces a artículos con los diferentes sentidos que puede tener el término buscado, es decir, describen múltiples conceptos asociados con un mismo término.

### 2.2. SKOS

Un *tesoro* es un conjunto de conceptos o términos de un vocabulario controlado organizados en una estructura jerárquica, donde cada término puede tener una o más relaciones *padre-hijo* con otros términos [27]. En los tesauros es posible representar otro tipo de relaciones como las *relaciones de equivalencia* o de sinonimia, y las *relaciones asociativas o related* [12].

*Simple Knowledge Organization System (SKOS)*[7], es un vocabulario desarrollado por el *World Wide Web Consortium (W3C)*, y que actualmente es el modelo ideal para representar cualquier tipo de vocabulario controlado, y especialmente para representar tesauros en *Resource Description Framework (RDF)* [16]. A diferencia de *Web Ontology Language (OWL)* [2,1], que permite expresar estructuras conceptuales complejas donde el tipo de relaciones

que se establecen entre los elementos depende del dominio, *SKOS* permite la representación de vocabularios más complejos, incluyendo ontologías [5] *OWL* [25], mediante esquemas sencillos de conceptos. En *SKOS*, los elementos únicamente se pueden relacionar por relaciones jerárquicas, asociativas o de equivalencia, lo que facilita su tratamiento, cuando se quiere utilizar la navegación entre las relaciones de los elementos como ayuda o asistencia a los usuarios en las búsquedas.

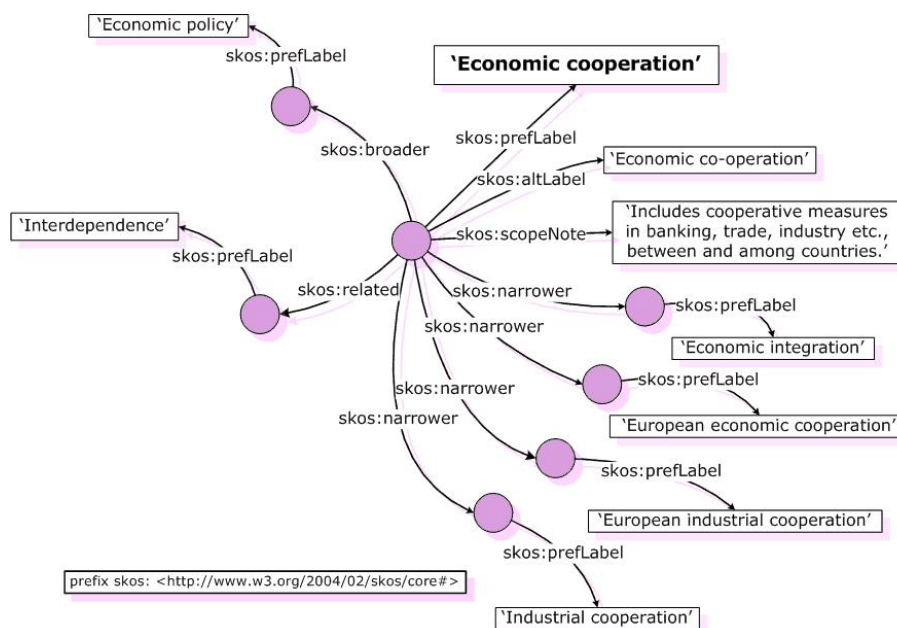


Fig. 1. Representación de un Tesauro usando SKOS

En *SKOS* (figura 1), cada concepto (*skos:Concept*) se identifica por una *Uniform Resource Identifiers (URI)* y representa una colección de términos sinónimos. Uno de los elementos de esta colección es el descriptor del concepto (*skos:prefLabel*) y se utiliza como identificador único de cada concepto. El resto de los sinónimos se representan como términos alternativos (*skos:altLabel*) asociados al descriptor, pudiendo existir un número indefinido de ellos. En el caso de los tesauros multi-idiomáticos, además se puede establecer relaciones entre un mismo concepto en distintos idiomas, mediante el atributo *lang* de la etiqueta (*skos:prefLabel*). Aunque sólo es posible definir una etiqueta preferida por idioma y tantas etiquetas alternativas como se desee en los diferentes idiomas. Análogamente a los tesauros, las relaciones semánticas que se pueden establecer en *SKOS* son relaciones jerárquicas, donde el término padre (*skos:broader*) suele representar el término más amplio o genérico y el término hijo representa al más concreto o específico (*skos:narrower*). Las relaciones de equivalencia o entre sinónimos se representan con las etiquetas (*skos:altLabel*) y (*skos:prefLabel*). Y finalmente se pueden establecer relaciones asociativas o *related* entre términos del tesauro que no están dentro de la misma jerarquía, pero que tienen cierta relación entre ellos mediante la etiqueta (*skos:related*).

### 3. Proceso de Construcción del Tesauro

En esta sección se presenta el proceso de mapeo de la estructura de la Wikipedia y de construcción del tesauro, así como la arquitectura de la aplicación utilizada para efectuar el proceso de mapeo.

#### 3.1. Análisis de la información de la Wikipedia

La forma de realizar la extracción de la información de Media Wiki y de realizar la transformación entre sus elementos y los conceptos de *SKOS* (sección 2.2) hay que establecerla basándose en los contenidos de la *MediaWiki* (sección 2.1). Esta transformación no es única, sino que pueden darse distintas interpretaciones y pueden generarse distintos algoritmos en función de los objetivos del tesauro a generar.

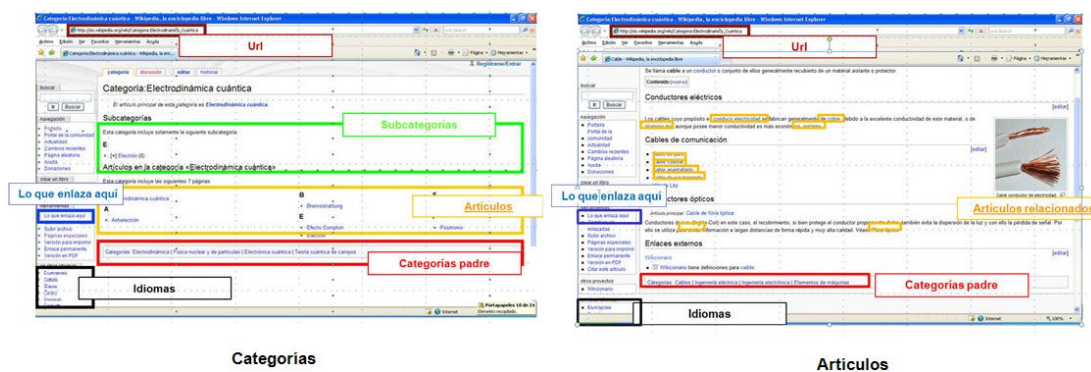


Fig. 2. Esquemas de las páginas de categorías ya artículos.

El algoritmo aquí propuesto trata de ofrecer una visión más general e independiente del dominio del tesauro. Para generar este algoritmo se ha analizado la estructura de las páginas de MediaWiki para determinar su relación con cada elemento del vocabulario *SKOS*. Para ello, en la figura 2 se muestran las distintas partes de las páginas de categorías y de artículos de donde se extrae la información que almacenará el fichero *SKOS*. La conclusión obtenida del análisis es la siguiente:

- **Identificador del concepto** (*skos:Concept*): Se obtiene de la *Uniform Resource Locator (URL)* de la página de la wikipedia.
- El descriptor o **elemento identificador** del concepto (*skos:prefLabel*): Se obtiene directamente de la parte final *URL* de los enlaces.
- El conjunto de **sinónimos o Alternatives** (*skos:altLabel*): Serán las redirecciones a la página actual y se obtiene de los enlaces cuyo id de la página *mw-whatlinkshere-list* (Lo que enlaza aquí).
- La **relaciones jerárquicas** (padre-hijo) o *Broader/Narrower*: Cada artículo o categoría pertenece a una o varias categorías, que son asignadas como *skos:broader*. Se identifican como aquellos enlaces cuyo id es *mw-normal-catlinks*. La asignación de *skos:narrower* no se realiza de forma explícita, sino que se diseña el software para que cuando se asigna un término *broader*, el tesauro incluya automáticamente la relación *narrower* correspondiente.
- La **relaciones asociativas** (*skos:related*): La estructura de las páginas de categorías y de artículos es diferente, por lo hay que diferenciar:

- para los artículos se toman como *related* los enlaces del contenido del artículo a otros artículos, que vendrán identificados con `column-content` en el `id`.
- para las categorías se toman como *related* aquellos padres que estén relacionados dentro de un mismo concepto; es decir, si un concepto pertenece a varias categorías se asume que hay cierta relación entre ellas.
- **Idiomas:** La página cargada establece un idioma para el término de *preflabel*, pero es posible extraer los términos en otros idiomas a través de los enlaces cuyo `id` es `p-lang`.

### 3.2. Arquitectura

Para la extracción de la información de la Wikipedia la arquitectura propuesta (figura 3) está formada por un *spider*, un *parser*, un módulo encargado del procesamiento de la información y otro del almacenamiento, en este caso a fichero utilizando **Jena** [8]. El *spider*, es el software que recorre las páginas a partir de sus enlaces, tomando como punto de partida la establecida por el usuario; se ha utilizado **JSpider** [10] y para controlar su comportamiento y filtrar las páginas a parsear se ha desarrollado un *plugin* específico. El *parser*, permite extraer la información de las páginas a partir de una etiqueta *HTML* recuperando y procesando su contenido; se ha utilizado **Jericho HTML** [9].

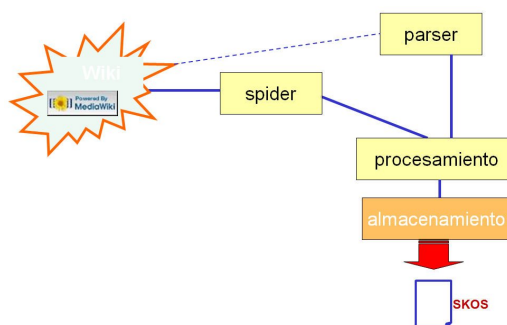


Fig. 3. Arquitectura de generación del fichero *SKOS*.

El algoritmo para generar el fichero *SKOS* (figura 4) se divide en dos fases:

En la primera etapa el *spider* construye la estructura del tesoro: a partir de una página inicial, según se recorre la Wiki se consigue recuperar todos los conceptos (*URI* y *preflabel*), así como sus relaciones (*broader*, *narrower* y *related*). El tamaño de la Wiki, en concreto, en el caso de Wikipedia, hace que el espacio y el tiempo de extracción de los conceptos aumente exponencialmente, por lo que interesa definir estrategias de parada para indicar al *spider* el momento de finalizar su labor de una forma ordenada. Se han planteado las posibles opciones:

- **Parada por número de conceptos:** cuando el *spider* alcance el número de conceptos configurado pasará a la siguiente fase.
- **Parada por niveles:** cuando el *spider* da el número de saltos configurado pasará a la siguiente fase.

La segunda fase recorre el tesoro generado completando la información de cada uno de los conceptos: *alternatives* y *preflabel* en distintos idiomas; así mismo para asegurar una estructura adecuada del tesoro se asigna *broader* a todos aquellos conceptos que no los tengan y se establecen como *Top Concepts* del tesoro aquellos que no tienen *broader*.

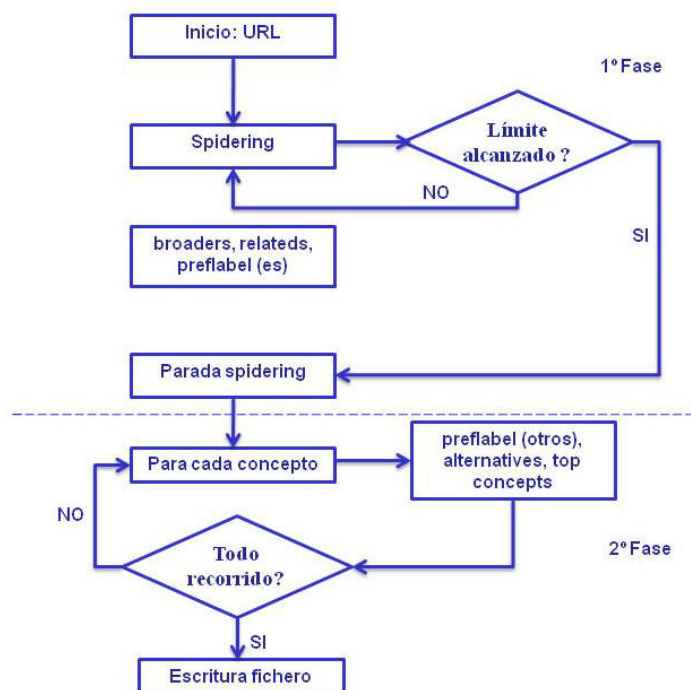


Fig. 4. Algoritmo de generación del fichero SKOS.

### 3.3. Experimentación

Usando la arquitectura anterior se han realizado diversas pruebas de generación de tesauros, tanto sobre Wikipedia, como desde otras Wikis. En la tabla 1 se recogen los datos de la generación del tesauro a partir de la categoría Microcontroladores de Wikipedia, configurado para recuperar 20 idiomas por concepto.

Descripción	Número conceptos (1ª Vuelta/2ª Vuelta)	Tiempo (1ª Vuelta/2ª Vuelta)	Memoria (1ª Vuelta/2ª Vuelta)
3 Niveles, sin límite de conceptos (Microcontroladores)	7549 / 10767	20 min. / 3h 56min.	29MB / 635MB
1000 Conceptos, sin niveles (Microcontroladores)	1000 / 1754	1 min. 20s. / 35 min.	17MB / 126MB

Tabla 1. Ejecución del algoritmo sobre Wikipedia

Analizando los datos de las ejecuciones, se observa que, independientemente de la estrategia de finalización, existe una diferencia entre los conceptos recuperados en la primera vuelta y en la segunda debido a que en la segunda parte del algoritmo, se crean todos los *broaders* de cada uno de los conceptos obtenidos en la primera vuelta, que aún no existan. En cuanto al tiempo de procesamiento, la segunda parte del procesamiento es bastante más

alta que la primera porque en la segunda parte se entra de nuevo en dos páginas para cada concepto. Para el concepto de partida el resultado obtenido en cualquiera de los casos es el de la figura 5.

```
<rdf:Description rdf:about="http://es.wikipedia.org/wiki/Categor%C3%A9a:Microcontroladores">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core\#Concept"/>
  <skos:prefLabel xml:lang="fi">Mikrokontrollerit </skos:prefLabel>
  <skos:prefLabel xml:lang="en">Microcontrollers </skos:prefLabel>
  <skos:prefLabel xml:lang="pl">Mikrokontrolery </skos:prefLabel>
  <skos:prefLabel xml:lang="es">Microcontroladores </skos:prefLabel>
  <skos:prefLabel xml:lang="fr">Microcontrôleur </skos:prefLabel>
  <skos:prefLabel xml:lang="da">Mikrokontrollere </skos:prefLabel>
  ...
  <skos:narrower rdf:resource="http://es.wikipedia.org/wiki/PIC16F84"/>
  <skos:narrower rdf:resource="http://es.wikipedia.org/wiki/AMD_29000"/>
  <skos:narrower rdf:resource="http://es.wikipedia.org/wiki/Logochip"/>
  <skos:narrower rdf:resource="http://es.wikipedia.org/wiki/Freescale_68HC08"/>
  <skos:narrower rdf:resource="http://es.wikipedia.org/wiki/MPLAB"/>
  <skos:narrower rdf:resource="http://es.wikipedia.org/wiki/Serie_AT89_de_Atmel"/>
  ...
  <skos:related rdf:resource="http://es.wikipedia.org/wiki/Categor%C3%A9a:Microprocesadores"/>
  <skos:related rdf:resource="http://es.wikipedia.org/wiki/Categor%C3%A9a:Electr%C3%B3nica"/>
  <skos:related rdf:resource="http://es.wikipedia.org/wiki/Categor%C3%A9a:Intel"/>
</rdf:Description>
```

Fig. 5. Resultado de la experimentación

Puesto que el objetivo no concierne únicamente a Wikipedia, sino a la tecnología MediaWiki se han realizado pruebas con otras Wikis que utilizan dicha tecnología, como Mardripedia [13], Qwiki [26] o WikiMusicGuide [29]. En estos casos lo que se ha comprobado es hay ciertos parámetros de la Wiki que hay contemplar en el algoritmo de extracción, como el contexto del servidor o las palabras reservadas para categorías y desambiguación.

### 3.4. Aplicación

La aplicación más inmediata es utilizar el tesauro generado como índice rápido para el manejo de una Wiki, de tal forma que se presente el mapa de conceptos de forma explícita y se pueda navegar por él, hasta llegar a los conceptos deseados, en vez de tener que realizar la búsqueda y navegación de forma explícita a través de la Wiki; una vez identificados los conceptos de interés se podría acceder a ellos directamente a través del enlace. Esta navegación por el mapa de conceptos es aún más útil cuando la Wiki es multiidioma, ya que el tesauro mostrará esa información y es posible ver que dependiendo del idioma las relaciones son más completas, en función de los completos que sean los contenidos en cada idioma.

Al margen de esto, puesto que el tesauro formaliza la relación entre conceptos, puede servir de fuente para realizar un asistente de búsquedas que aproveche la relación entre los conceptos para afinar las búsquedas en otros contextos ya que visualizando los conceptos *related* extraídos, se puede hacer una idea rápida sobre cierto tema y ver que otros conceptos están implicados en lo que se buscando, para así ubicarse mejor en aquello que se pretende buscar.

Un punto de vista complementario es no sólo utilizar el tesauro para afinar los criterios de búsqueda, sino para aplicarlo en la formación y priorización de los resultados obtenidos en la búsqueda, de forma que los primeros resultados serían los que están más ligados

semánticamente al concepto buscado y todos los que le rodean. O en este mismo sentido, no sólo hacer búsquedas por los *preflabel*, sino también por los correspondientes *alternatives*.

#### 4. Trabajos Relacionados

Existen variadas iniciativas con propuestas para estructurar los contenidos Wikipedia, como recoge [3] o [11], pero en estos casos la estructura obtenida no es un tesoro y si bien muestran posibilidades de información a extraer no son comparables a esta propuesta. Existe un trabajo para generar tesauros a partir de Wikipedia en [4], que tiene como objetivo realizar un mapeo de palabras a su concepto correspondiente, a partir de una base previa de palabras (nombre propios y comunes), utilizando Wikipedia como fuente de información de los conceptos, para de esta forma facilitar búsquedas y lograr una mayor precisión. Y es en base a este objetivo como se define el algoritmo de extracción y generación del tesoro. En este caso los títulos de los artículos son utilizados para definir los *preflabel*, las páginas de redirección generan los *altlabel* y se examinan todos los enlaces de los artículos para establecer relaciones entre conceptos (*related*), dentro de los cuales diferencian dos tipos de relaciones: las débiles, en las cuales desde un concepto se referencia a otro, pero no a la inversa, y las fuertes en las cuales la referencia es bidireccional. La cuestión es que *related* expresa una relación simétrica, por lo que la decisión es si incluir las relaciones débiles o solo las fuertes. Aunque hay ideas interesantes, como el tratamiento de *related*, el algoritmo propuesto está muy asociado al objetivo planteado y no puede ser considerado dentro de una visión más general, como la aquí propuesta; por ejemplo, no plantean una estrategia para establecer relaciones *broader* ni *narrower*, y al utilizar únicamente el inglés, tampoco hacen tratamiento de los idiomas, con lo que se pierde conocimiento existente en Wikipedia.

#### 5. Conclusiones y Líneas Futuras

En este artículo se ha presentado un algoritmo que materializa la estructura conceptual de la Wikipedia, y de sus relaciones, en un tesoro que permitirá fácilmente mejorar las búsquedas en los sistemas de gestión del conocimiento aplicando sus propiedades de sinonimia, traducción, generalización y particularización conceptual.

Aunque existe distintas aproximaciones similares[17,20,6,19], el algoritmo aquí propuesto trata de ofrecer una visión más general e independiente del dominio del tesoro, permitiendo generar de forma sencilla tesauros multi-idioma en torno a un concepto dado a partir de la estructura de la Wikipedia. Adicionalmente, al ser un algoritmo basado en Media Wiki, puede ser aplicado a cualquier Wiki realizada con dicha tecnología, permitiendo la creación de un mapa conceptual completo y navegable de la información de las Wikis empresariales.

El trabajo aquí presentado, únicamente se basa en la generación de tesauros, pero no considera la actualización de los mismos, ya que ante cambios sustanciales del mapa de conceptos es posible volver a generar el tesoro; no obstante como líneas futuras se plantea la posibilidad de contemplar cambios incrementales en el mapa de conceptos o tesoro.

#### Referencias

1. G. Antoniou and F. vanHarmelen. *A Semantic Web Primer*. MIT Press, Cambridge, 2004.
2. S. Bechhofer, F. van Harmelen, J. Hendler, I. Horrocks, D.L. McGuinness, P.F. Patel-Schneider, L.A. Steinnd, and F.W. Olin. Owl web ontology language reference, February 2004.
3. dbpedia. <http://dbpedia.org>.
4. A. Gregorowicz and M A. Kramer. Mining a large-scale term-concept network from wikipedia. *Mitre Technical Report*, 6(1).

5. T. R. Gruber. Towards Principles for the Design of Ontologies Used for Knowledge Sharing. In N. Guarino and R. Poli, editors, *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Deventer, The Netherlands, 1993. Kluwer Academic Publishers.
6. Bachlechner D. Hepp, M. and K. Siorpaes. Harvesting wiki consensus - using wikipedia entries as ontology elements. *First Workshop on Semantic Wikis*, 2006.
7. A. Isaac and E. Summers. Skos simple knowledge organization system primer, February 2008.
8. jena. <http://jena.org/>.
9. jericho. <http://jericho.org/>.
10. jspider. <http://jspider.org>.
11. S. Kruger. Acquiring the semantic relationships of links between wikipedia articles. research proposal. <http://www.coloradostoutenburg.com/cs58920v1.doc>.
12. M.J.Lamarca. Tesauros. <http://www.hipertexto.info/documentos/tesauros.htm>.
13. Madripedia. <http://www.madripedia.es/>.
14. E. Mäkelä. Survey of semantic search research. In *Proceedings of the Seminar on Knowledge Management on the Semantic Web*, 2005.
15. C. Mangold. A survey and classification of semantic search approaches. *Int. J. Metadata Semant. Ontologies*, 2(1):23–34, 2007.
16. F. Manola and E. Miller. Rdf primer, February 2004.
17. Legg C. Milne D. Medelyan, O. and I.H Witten. Mining meaning from wikipedia. Hamilton, New Zealand: University of Waikato, Department of Computer Science. Computer Science Working Papers, November 2008.
18. MediaWiki. <http://es.wikipedia.org/wiki/MediaWiki>.
19. D Milne and Nichols D. M. Witten, I.H. Extracting corpus specific knowledge bases from wikipedia. Technical Report Working Paper 03/2007, University of Waikato, 2007.
20. Medelyan O. Milne, D. and I.H. Witten. Mining domain-specific thesauri from wikipedia: A case study. In *WI '06: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*, pages 442–448, Washington, DC, USA, 2006. IEEE Computer Society.
21. D.I Moldovan and R. Mihalcea. Using wordnet and lexical operators to improve internet searches. *IEEE Internet Computing*, 4(1):34–43, 2000.
22. T. O'Reilly. What is web 2.0—design patterns and business models for the next generation of software, September 2005. <http://www.oreillynet.com/lpt/a/6228>.
23. T.B. Passin. *Explorer's Guide to the Semantic Web*. Manning Publications Co, 2004.
24. Redirecciones. <http://es.wikipedia.org/wiki/Wikipedia:Redirecciones>.
25. Simple Knowledge Organization System. <http://en.wikipedia.org/wiki/SKOS>.
26. Stanford University. Qwiki. <http://qwiki.stanford.edu/>.
27. Pidcock W. What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model?, January 2003. <http://www.metamodel.com/article.php?story=20030115211223271>.
28. Wiki. <http://es.wikipedia.org/wiki/Wiki>.
29. WikiMusicGuide. <http://www.wikimusicguide.com>.
30. Retrieved Wikipedia: Wiki, September 2008. <http://en.wikipedia.org/wiki/Wiki>.