



# ANALES DE DOCUMENTACIÓN

Anales de Documentación  
ISSN: 1575-2437  
revandoc@um.es  
Universidad de Murcia  
España

Martínez González, M. Mercedes; Alvite Díez, M.ª Luisa  
Propuesta metodológica de evaluación de gestores de tesauros compatibles con la web semántica  
Anales de Documentación, vol. 17, núm. 1, enero-junio, 2014, pp. 38-55  
Universidad de Murcia  
Espinardo, España

Disponible en: <http://www.redalyc.org/articulo.oa?id=63530415004>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

# PROPUESTA METODOLÓGICA DE EVALUACIÓN DE GESTORES DE TESAUROS COMPATIBLES CON LA WEB SEMÁNTICA

**M. Mercedes Martínez González\***

Departamento de Informática. Edificio T.I.T. Universidad de Valladolid.

**M.ª Luisa Alvite Díez\*\***

Área de Biblioteconomía y Documentación. Facultad de Filosofía y Letras. Universidad de León.

**Resumen:** El desarrollo de la web semántica, entendida como un espacio compartido para el intercambio de datos estructurados, ha supuesto un interés renovado por los lenguajes controlados tradicionales. Este estudio presenta una metodología evaluativa que posibilita el análisis sistemático de herramientas para la gestión de tesauros mediante la aplicación de un conjunto de criterios que tienen en cuenta el propósito, las funcionalidades de la aplicación, el mantenimiento de la integridad o coherencia del vocabulario, así como otras cuestiones cruciales referidas a la interoperabilidad, con especial atención a la compatibilidad con los estándares de la web semántica (RDF/SKOS), a la integración de estas herramientas en otros sistemas de información, la interoperabilidad de la información contenida o la habilidad para combinar tesauros procedentes de diferentes fuentes.

**Palabras clave:** Herramientas para la gestión de tesauros; integrabilidad del software; interoperabilidad; metodología de evaluación; SKOS; tesauros; web semántica.

**Title:** METHODOLOGICAL PROPOSAL EVALUATION OF THESAURI MANAGEMENT TOOLS COMPATIBLE WITH THE SEMANTIC WEB.

**Abstract:** The development of the Semantic web, understood as a shared space for exchange of structured data, has led to a renewed interest in traditional controlled vocabularies. This study presents an evaluation methodology that enables a systematic analysis of thesauri management tools applying a set of criteria that takes into account the purpose, functionalities of the application, integrity and consistency of vocabulary and other crucial issues related to interoperability, with particular attention to compatibility with Semantic web standards (RDF/SKOS). Integrability of these tools in other information systems, information interoperability, merging of thesauri between different information systems, or the ability to combine thesauri from different sources have also part of the criteria proposed.

**Keywords:** Evaluation methodology; interoperability; thesauri; thesauri management tools; software integrability; semantic web; SKOS.

## 1. INTRODUCCIÓN.

Los lenguajes documentales han demostrado su eficacia durante décadas. En concreto, los tesauros se impusieron a principios de los años sesenta del pasado siglo impulsando la idea del concepto como núcleo organizador del vocabulario controlado. El desarrollo de la web semántica, entendida como un espacio compartido para el intercambio de datos estructurados, ha supuesto un interés renovado por los tradicionales lenguajes controlados. Ahora bien, esta web semántica requiere que los datos sean legibles y comprensibles por los agentes software, para lo cual son necesarios estándares que permitan la representación de información en formatos interoperables.

En esta línea, en agosto de 2011 se publicó la norma ISO 25964-1:2011 *Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval*. En marzo de 2013 se ha editado la segunda parte -- *Part 2: Interoperability with other vocabularies*. Por otro lado, el Consorcio web (W3C) proporciona un modelo, SKOS (Simple Knowledge Organization Systems) para expresar la estructura básica y el contenido de tesauros, clasificaciones, listas de encabezamiento de materia, taxonomías o folksonomías. SKOS es un "modelo de datos diseñado para compartir sistemas de organización de conocimiento en la web" (W3C, 2009).

Entendemos lenguaje documental como un sistema artificial de signos normalizados que facilita la representación formalizada del contenido de los documentos para permitir la recuperación de información solicitada por los usuarios (Gil Urdiciaín, 2004). Tiene que ser único y estar dotado de una sencilla organización. Subrayamos en este punto la imbricación indisoluble entre recuperación de información y organización o representación del conocimiento.

\* mercedes@infor.uva.es

\*\* luisa.alvite@unileon.es

Recibido: 20-11-2013; 2ª versión: 30-01-2014; aceptado: 17-03-2014.

MARTÍNEZ GONZÁLEZ, M.M. y ALVITE DÍEZ, M.L. Propuesta metodológica de evaluación de gestores de tesauros compatibles con la web semántica. *Anales de Documentación*, 2014, vol. 17, nº 1.

Disponible en: <http://dx.doi.org/10.6018/analesdoc.17.1.186271>

Si bien existen diversos criterios de sistematización de los lenguajes documentales, la literatura suele recoger como lenguajes controlados principalmente; las clasificaciones, taxonomías, las listas de encabezamientos de materia y los tesauros. Todos ellos presentan un vocabulario previamente elaborado y admiten un limitado número de modificaciones en el momento de su utilización.

Deteniéndonos en los tesauros, la Norma ISO 2788-1986 equivalente a la española UNE 50106-1990 entiende el tesauro desde dos perspectivas, según su función, la norma define tesauro como un instrumento de control terminológico que traduce a un lenguaje documental el lenguaje natural empleado en los documentos y por los usuarios. Atendiendo a su estructura, el tesauro es un vocabulario controlado y dinámico de términos relacionados semántica y jerárquicamente, que se aplica a un campo específico de conocimiento. Un tesauro es un lenguaje postcoordinado de estructura combinatoria formado por un conjunto de términos preferidos (descriptores) y no preferidos (no descriptores). Los términos se vinculan entre sí para establecer relaciones de equivalencia, jerarquía y asociación.

Con la llegada de SKOS ha ganado relevancia la visión de tesauros orientados a conceptos frente a la de tesauros basados en términos. Los tesauros basados en términos organizan el vocabulario de un dominio, esto es, el vocabulario consiste en un conjunto de términos relacionados semántica y jerárquicamente. Como indican Pástor Sánchez, Martínez Méndez y Rodríguez Muñoz (2009), los tesauros basados en conceptos contienen conceptos (ideas abstractas), a los que se asocian etiquetas preferentes y no preferentes. En este caso, las relaciones se establecen entre los conceptos.

Los esfuerzos del W3C en SKOS (Miles y Bechhofer, 2009), un estándar que proporciona un modo de representar sistemas de organización del conocimiento mediante RDF (Resource Description Framework), han llevado, incluso antes de su estabilización, a varias iniciativas para representar tesauros con RDF (Lacasta *et al.*, 2007; Faro *et al.*, 2008; Polo *et al.*, 2008; Smedt, 2009) o tentativas para gestionar el uso de tesauros SKOS (Gerbe y Kerhervé, 2010; Koller, 2009; Mayr *et al.*, 2010).

SKOS y RDF constituyen los dos estándares de la web semántica que habrán de permitir a los tesauros ocupar con éxito un espacio en la ansiada interoperabilidad de la información, tan relevante para los sistemas de información web.

Asimismo, la representación de herramientas terminológicas en modelos vinculados a la web semántica guarda una estrecha relación con lo que se viene denominando Linked Data, una iniciativa que pretende facilitar el acceso a los datos, su publicación y reutilización a través de la web y la construcción de aplicaciones en torno a los datos así expuestos. Señala Pastor Sánchez (2011, p. 113) como “Linked Open Data conforma la infraestructura de un objetivo más ambicioso: la interconexión de contenidos, datos adaptados para el desarrollo de procesos inteligentes y dispositivos de todo tipo”.

Este marco de actuaciones nos lleva a fijarnos en las herramientas que permiten en la actualidad la gestión de tesauros y proponer una metodología para su evaluación en la que se tenga en cuenta de modo específico su adaptación a la web semántica.

Si bien tradicionalmente los tesauros se elaboraron manualmente o con editores de texto, las actuales herramientas para la gestión de tesauros permiten crear, editar y utilizar el tesauro para tareas heterogéneas, como anotar semánticamente contenidos (Hildebrand *et al.*, 2009). Sin embargo, no todas estas aplicaciones se han ideado con el mismo propósito, ni permiten las mismas operaciones, ni soportan aspectos relacionados con la integrabilidad e interoperabilidad en niveles semejantes.

En este trabajo se realiza un análisis sistemático de herramientas para la gestión de tesauros mediante la aplicación de un conjunto de criterios y que conforma una propuesta metodológica que tiene en cuenta el propósito y la funcionalidad de la aplicación, así como otras cuestiones cruciales referidas a la integración de estas herramientas en otros sistemas de información o a la interoperabilidad de la información contenida, la capacidad de intercambio de tesauros entre diferentes sistemas de información o la habilidad para combinar tesauros de diferentes fuentes.

Consideramos que la evolución actual de las herramientas terminológicas para la organización del conocimiento ha de entenderse dentro del escenario de la web semántica, un marco que aglutina un conjunto de tecnologías con un grado de abstracción cada vez mayor y en cuya evolución, como apunta Pastor Sánchez (2011, p. 42) “son fundamentales las constantes sinergias entre cada uno de los niveles, producto de la aplicación y evaluación de las diferentes tecnologías”.

## 2. OBJETIVOS Y METODOLOGÍA.

El objetivo principal de este estudio se dirige a proponer criterios evaluativos que permitan valorar sistemáticamente aplicaciones informáticas ideadas para la gestión de tesauros. Los indicadores atenderán a las funcionalidades exigibles a estas herramientas y prestarán una atención especial a su adecuación a los presupuestos de la web semántica. Asimismo, pretendemos indagar en el estado actual de desarrollo del software de gestión de tesauros y su adaptación a los nuevos estándares surgidos en el contexto de la web semántica.

Para la formalización de los indicadores planteados se han tenido en cuenta trabajos con propuestas evaluativas en campos cercanos, Moya Martínez y Gil Leiva (2001), García Castro y Gómez Pérez (2005, 2007, 2009), estos últimos relativos a *frameworks* de la web semántica, estudios de carácter generalista cuya pretensión consiste en fijar directrices para proponer marcos de evaluación en la web semántica, y que nos han servido de guía para la propuesta metodológica objeto de este artículo, propuesta específicamente dirigida a la evaluación de gestores de tesauros. Nuestro planteamiento surge ante la carencia de una metodología de evaluación de herramientas de tesauros actualizada a tenor de los últimos avances de la web semántica en la representación y tratamiento de estos vocabularios. Cubre, por tanto, un espacio que, hasta donde conocemos, no estaba cubierto.

Los parámetros en los que se focaliza el análisis son los siguientes:

- Propósito de la aplicación
- Requisitos del sistema
- Funcionalidades
- Estructuras conceptuales y terminológicas soportadas
- Gestión de la integridad
- Interoperabilidad de la información
- Interoperabilidad e integrabilidad del software

Para testear la propuesta evaluativa se examinan un conjunto de aplicaciones detalladas en el Apartado 4. Pretendemos compartir una metodología de evaluación y para plantearla se han empleado seis herramientas de diversa orientación, sin pretender en ningún caso confrontar o contrastar productos.

Para la ejecución de los tests de evaluación se han utilizado los tesauros EuroVoc -Tesauro multilingüe de la Unión Europea-, UKAT -UK Archival Thesaurus-, NAL -National Agricultural Library thesaurus- y un microtesauro elaborado *ad hoc* a partir de EuroVoc.

## 3. PROPUESTA METODOLÓGICA DE EVALUACIÓN.

La construcción y mantenimiento manual de los tesauros o con el limitado auxilio de los procesadores de texto fue con el tiempo reemplazada por aplicaciones específicas ideadas para la elaboración y gestión de estos vocabularios. No obstante, no todas las herramientas se han creado con el mismo propósito, ni permiten las mismas acciones ni favorecen la integrabilidad e interoperabilidad en el mismo grado.

Nos proponemos establecer un conjunto de criterios que posibiliten la evaluación comparativa de estas aplicaciones para lo cual se han tenido en cuenta prioritariamente los trabajos de Moya Martínez y Gil Leiva (2001), investigación centrada en el análisis de distintos softwares de gestión de tesauros, y los estudios de García Castro y Gómez Pérez (2005, 2007, 2009) focalizados en la evaluación comparativa de distintas aplicaciones relacionadas con tecnologías de la web semántica. Se han revisado igualmente otros trabajos como el de Severino (2007) resultado del análisis de encuestas realizadas a usuarios de tesauros de diversas instituciones y empresas o el de Willpower Information (2004) en el que se valoran funcionalidades de distintos softwares como la edición, la gestión de las relaciones entre los términos, la importación, la visualización del tesauro, entre otras.

Planteamos en este estudio la inclusión novedosa entre los indicadores evaluativos de aquellos aspectos relacionados con la integrabilidad y la interoperabilidad de la información gestionada, cuestiones no contempladas en la literatura previamente. Tal como se ha indicado en la sección anterior, hasta donde conocemos, no existe una propuesta específica para la evaluación de gestores de tesauros, cuyos criterios estén actualizados en concordancia con la evolución actual de la web semántica y su impacto en la representación y manipulación de tesauros. Asimismo, dada la trascendencia de SKOS como estándar de la web semántica y como estándar preferido para la representación de tesauros, se tendrá en cuenta si la Recomendación establecida por el W3C en agosto de 2009 (Miles y Bechhofer,

2009), o sus borradores previos (Miles y Bechhofer, 2008a,b) son tenidos en cuenta en las aplicaciones de gestión de tesauros.

Este trabajo, como se ha apuntado, pretende centrarse en la propuesta metodológica de evaluación, incluyendo cuestiones de interoperabilidad, con especial atención a la compatibilidad con los estándares de la web semántica (RDF/SKOS), e integrabilidad del software. Sin embargo, no se pueden olvidar aspectos trascendentales como las funcionalidades de la aplicación, determinantes para usuarios finales, como editores de tesauros o creadores a la hora de seleccionar una de estas herramientas. Asimismo, se ha de tener en cuenta que un mayor número de estructuras de tesauros soportadas por la herramienta supone una mayor garantía de compatibilidad con otros tesauros. La interoperabilidad sintáctica está directamente vinculada con los formatos compatibles. Además, la capacidad de soportar vocabularios normalizados, como SKOS, acerca la herramienta a la interoperabilidad semántica en lo que respecta a la información, aunque lograr la interoperabilidad semántica completa pueda ser un problema más complejo.

Como subraya Pastor Sánchez (2013), la nueva norma ISO 25964 está mucho más cercana a SKOS que las anteriores, en las que se ofrecía un enfoque léxico frente al nuevo modelo de datos que presenta la ISO de 2011, que facilita el desarrollo de aplicaciones informáticas y su aplicación en procesos de recuperación de información. Para favorecer la interoperabilidad entre ISO y SKOS en 2012 el grupo de trabajo ISO TC46/SC9/WG8 e ISAAC ha publicado un documento que establece las relaciones entre ambos modelos.

Por otra parte, los desarrolladores de aplicaciones de tesauros presentan un perfil de usuario particular en relación a los componentes que incluyen en las mismas. Generalmente, suelen preferir reutilizar paquetes de software o servicios de terceros. La integración del software se consigue por medio de paquetes de software que pueden ser utilizados por otras herramientas (paquetes *.jar* para aplicaciones Java que pueden ser compiladas con la aplicación empleada, widgets y otros que se puedan integrar en páginas web, servicios que intercambian mensajes XML, etc.) o el uso de API (Application Programming Interface). Cuando una API que proporciona la interoperabilidad necesaria para integrar este software en otras aplicaciones se encuentra disponible, ésta debería estar bien documentada, con un sistema de interfaces bien definido y documentado que se pueda reutilizar (Bloch, 2006). Por otro lado, en el contexto de internet, los servicios web se han establecido como una técnica apreciada para la reutilización de funcionalidad externa. Los servicios web impulsan la interoperabilidad a través del intercambio de mensajes XML. En nuestra propuesta se comprueba la presencia de ambas posibilidades, API y servicios web.

Con las acotaciones señaladas, los criterios evaluativos propuestos son los siguientes:

1. Propósito de la aplicación.

Atiende este criterio a la finalidad para la que se ha diseñado la herramienta. Esto va a condicionar las funcionalidades de la aplicación. La mayoría de las herramientas tratan de facilitar la tarea de los usuarios de tesauros (que buscan en los tesauros, pero no editan los mismos) y editores de tesauros (que crean y editan tesauros).

2. Requisitos del sistema.

Los requisitos del sistema condicionan las posibilidades de utilización del software. Este parámetro evaluativo hace referencia especialmente a los usuarios que no cuentan con aptitudes especiales para la instalación y gestión de software sofisticado o para aquellos usuarios que simplemente quieren minimizar el software instalado en sus equipos.

3. Funcionalidades.

Cada herramienta dispone de sus propias funcionalidades. Se han seleccionado un conjunto de las mismas por ser comunes a la mayoría de las herramientas examinadas o por ser consideradas significativas en análisis previos sobre funcionalidades exigibles a las herramientas de gestión de tesauros (Severino, 2007; Moya Martínez y Gil Leiva, 2001). En este trabajo, persiguiendo la mayor concisión posible, se han considerado como especialmente relevantes las siguientes:

- Creación y edición de tesauros. Este ítem examina las tareas siguientes:

- Crear y eliminar tesauros
- Crear, borrar y editar elementos de los tesauros (términos/ conceptos, relaciones, dominios, sub-dominios)
- Crear, borrar y editar colecciones<sup>1</sup>

- Revisión de los elementos de los tesauros, es decir, la posibilidad de marcar términos o conceptos como candidatos a ser incluidos, eliminados, modificados, etc.
- Búsqueda y recuperación en los tesauros. En esta categoría se incluye la revisión de los aspectos siguientes:
  - Recuperación de metadatos de los tesauros, términos cabeceras (top concepts), colecciones y componentes de las colecciones
  - Búsqueda de etiquetas (recuperación de conceptos cuyas etiquetas contengan una cadena de búsqueda dada)
  - Búsquedas que exploten las relaciones semánticas, es decir, que recuperen términos y etiquetas de un concepto así como información sobre sus posibles conceptos específicos y conceptos relacionados
  - Búsqueda avanzada: combinación de términos de búsqueda con operadores lógicos (AND, OR, NOT), búsquedas expandidas (a través de términos relacionados o términos que guarden alguna relación jerárquica), búsquedas con comodines, etc.
- Navegación
- Combinación ( fusión ) de tesauros
- Importación y exportación de tesauros desde y hacia ficheros diversos
- Provisión de ficheros *log* que incorporen las modificaciones efectuadas en los tesauros o la prestación de otros registros o informes procedentes de este tipo de ficheros
- Gestión y colaboración de usuarios. Posibilidad de apoyar a los usuarios que colaboran en la edición/creación de tesauros

#### 4. Estructuras conceptuales y terminológicas soportadas.

Los tesauros proporcionan una serie compleja de estructuras para la organización del vocabulario que los conforma. No obstante, no todas ellas tienen la misma trascendencia ni aparecen en todos los tesauros. Las estructuras más cruciales que consideramos ha de gestionar un software de gestión de tesauros son:

- Relaciones semánticas: Términos preferentes (Descriptores) / Términos no preferentes (No descriptores). Cuando se trata de un tesauro orientado a conceptos, encontramos Conceptos, que tienen asociada una etiqueta preferente y etiquetas alternativas. Las relaciones jerárquicas se establecerán en este último caso entre Conceptos en lugar de Términos preferentes.
- Relaciones de equivalencia: UF, UP (Usado Por), colocado delante del término no preferente y USE (Use) colocado delante del descriptor para guiar al usuario hacia el término preferente. En los tesauros basados en conceptos estas relaciones están embebidas en el rol atribuido a cada etiqueta asociada a un concepto.
- Relaciones jerárquicas:
  - Que expresan la subordinación entre conceptos: BT, TG (Término Genérico) y NT, TE (Término específico). Términos cabecera o Top Terms (Top Concepts en tesauros basados en conceptos) / Términos no cabecera o No Top Terms.
  - Dominios / subdominios. Aunque no es común a todos los tesauros, muchos de ellos cuentan con múltiples niveles de dominio, esto es, presentan clasificaciones jerárquicas en varios niveles, subdominios que se integran en otros dominios. Puede servir de ejemplo EuroVoc, que cuenta con una estructura de subdominios o microtesauros.
  - Asimismo, aunque no resulta aconsejable, hay tesauros que cuentan con conceptos que pueden pertenecer a varios campos temáticos a la vez, esto es, pueden subordinarse a varios términos genéricos. Este fenómeno se denomina polijerarquía.
- Relaciones asociativas, entre dos términos o conceptos que no pertenecen a la misma estructura jerárquica aunque sean semántica o contextualmente similares. RT, TR (Término relacionado).
- Notas que aclaran el significado o el uso de un término: notas de alcance, notas históricas, notas del editor, notas de uso, etc.

Además, merece la pena apuntar que SKOS tiene en cuenta ciertas construcciones que podrían estar soportadas por las aplicaciones para la gestión de tesauros:

- Conceptos (*skos:Concept*). Con el estándar SKOS los conceptos se definen como un recurso de la clase “*skos:Concept*” y se identifican con un URI. Este enfoque considera los vocabularios compuestos por conceptos en lugar de términos, conceptos que están representados por etiquetas. Las etiquetas pueden representar términos preferentes (*skos:prefLabel*), alternativos (*skos:altLabel*) u ocultos (*skos:hiddenLabel*).
- Concept Schemes (*skos:ConceptScheme*). Estructura que recopila y organiza un conjunto de conceptos. Puede ser utilizado para representar tesauros y dominios.
- Colecciones (*skos:Collection*). Las colecciones son grupos de conceptos, que no necesitan estar vinculados por relaciones jerárquicas (TG o TE) (Isaac y Summers, 2009). Los miembros de una colección pueden proceder de diferentes *ConceptSchemes*, lo que significa que una colección puede agrupar conceptos de diferentes tesauros.

### 5. Gestión de la integridad.

El mantenimiento de la integridad de los tesauros es considerada unánimemente una funcionalidad primordial entre aquellas que han de presentar las herramientas de gestión de tesauros (Moya Martínez y Gil Leiva, 2001; Pastor Sánchez, 2009). Estas aplicaciones han de gestionar la coherencia semántica de las relaciones creadas en el tesauro, impidiendo automáticamente la introducción de posibles relaciones incoherentes que romperían la consistencia del vocabulario. Además, es deseable que una herramienta que se ocupa de la gestión automática de la integridad sea capaz de incorporar automáticamente las relaciones que se puedan derivar de las ya incluidas en el tesauro.

Las herramientas de verificación automática de la integridad suponen una inestimable ayuda para la adecuada elaboración y mantenimiento de estos vocabularios evitando errores de consistencia y garantizando la calidad del tesauro. Como sucede con las ontologías (Noy *et al.*, 2001), esta funcionalidad es especialmente relevante a medida que el tamaño del tesauro crece. Igualmente, cuando la herramienta de tesauros es invocada desde otras aplicaciones, al igual que ocurre con los software de bases de datos, es preferible que sea el propio gestor (de tesauros en este caso) quien verifique la integridad.

En esta propuesta, la gestión de la integridad del tesauro se traduce en la capacidad de:

- Registrar automáticamente las relaciones inversas cuando un término es incluido en el tesauro (por ejemplo, si *A* es NT de *B*, entonces *B* es BT de *A*)
- Eliminar automáticamente las relaciones en las que participa un término cuando este se elimina del vocabulario
- Impedir la introducción de relaciones inconsistentes, por ejemplo, prohibir la inclusión de un término simultáneamente como término genérico y como término específico de otro dato

### 6. Interoperabilidad de la información: formatos compatibles.

La capacidad de importación y/o exportación de tesauros representados con diferentes formatos es una cuestión fundamental para lograr la interoperabilidad. Por otra parte, la capacidad para permitir las representaciones en SKOS resulta crucial para atender los objetivos perseguidos por la web semántica. Tendremos en cuenta en este parámetro tanto la capacidad de importar y exportar tesauros como los formatos soportados por las distintas aplicaciones. Tendremos en consideración en el examen la posibilidad de texto plano así como los formatos relacionados con la representación de datos en la web semántica: RDF/SKOS y sus variantes sintácticas, RDF/XML, N3 Turtle, JSON, etc.

### 7. Interoperabilidad e integrabilidad del software: provisión de paquetes de software o servicios.

El software puede ser reutilizable por medio de APIs que ofrecen un conjunto de librerías, provistas de interfaces que describen las definiciones de clase y su comportamiento (en el caso de los lenguajes orientados a objetos).

Para esta propuesta evaluativa se considera una cuestión adicional, específica del contexto de esta evaluación, si la API es independiente de la representación de tesauros o no. Algunos de los programas de gestión de tesauros son aplicaciones orientadas a SKOS, lo que significa que la API asociada utiliza un vocabulario SKOS (gestionan *ConceptSchemes* en lugar de tesauros, dominios o microtesauros, manejan etiquetas *Preferred / NonPreferred* en lugar de Descriptores / No Descriptores, etc.) En estos casos, el usuario de la API debe conocer el vocabulario de SKOS y las construcciones que se pueden utilizar para representar cada concepto en los tesauros.

Por su parte, un servicio web es un "sistema de software diseñado para apoyar la interoperabilidad e interacción entre equipos en una red" (Haas y Brown, 2004). Se trata, por tanto, de tecnología que utiliza un conjunto de

protocolos y estándares que sirven para intercambiar datos entre aplicaciones. Permiten que distintas aplicaciones de software, desarrolladas en lenguajes de programación diferentes, y ejecutadas sobre cualquier plataforma, puedan utilizar los servicios web para intercambiar datos, esto es, se convierten en aplicaciones web. La interoperabilidad se consigue mediante la adopción de estándares abiertos.

Atendiendo a la documentación del Tesauro GEMET, que incluye una API para acceder a los tesauros a través de un servicio web, la primera API basada en servicios web para tesauros fue la de la Biblioteca Nacional de Agricultura de Estados Unidos<sup>2</sup>. En el año 2001 apareció la especificación SKOS RDF y con ella la API SKOS<sup>3</sup>. Sin embargo, fue abandonada aproximadamente en 2004. Las posibles razones de este abandono, a tenor de la documentación de GEMET, podrían relacionarse con la propia naturaleza de la API, muy específica para SOAP (Simple Object Access Protocol) y con su carácter monolingüe.

## **4. APPLICACIÓN DE LOS CRITERIOS EVALUATIVOS.**

### **4.1 Selección de herramientas.**

Para validar la propuesta metodológica planteada en este estudio se procedió a una recopilación de herramientas para la gestión de tesauros que cumplieran con ciertos requisitos.

En primer lugar debían ser aplicaciones libres o contar con una versión de evaluación. Esto coloca al lector interesado en situación de reproducir los experimentos, sea cual sea su capacidad de adquirir software con licencia. Creemos, además, que las herramientas de libre distribución tienen una implantación suficiente para justificar centrarnos en ellas sin mermar la validez o interés de esta propuesta por ello. En segundo lugar debían contar con versiones posteriores a la aparición de SKOS, esto es, posteriores al año 2004 y candidatas posibles, por tanto, a soportar el estándar. En tercer lugar, debían ser herramientas creadas para la gestión de tesauros o con cierta experiencia documentada sobre el empleo de las mismas para este fin. Así se excluyeron de nuestros tests aplicaciones cuyo propósito general es la edición de ontologías (Protégé, Top Braid Suite o SWOOP), herramientas orientadas a la gestión de terminología (Wintertree, Thesaurus Engine), sistemas de gestión de bases de datos que soportan tecnologías semánticas o cualquier otra herramienta susceptible de ser empleada para el desarrollo de aplicaciones de tesauros, pero que no ha sido específicamente diseñada para este propósito.

Las herramientas seleccionadas para los tests fueron localizadas en sedes web que mantienen directorios de software para la gestión de tesauros, tecnologías semánticas y aplicaciones para SKOS<sup>4,5,6,7,8</sup>.

Finalmente, las aplicaciones objeto de análisis para testear la metodología propuesta fueron:

- MultiTes (2005)
- One-2-One (2009). Esta herramienta reemplaza a TermTree
- PoolParty 2.7 (2010)
- SKOSEd 1.0: editor de SKOS para Protégé-OWL 4.0 (enero de 2009)
- TemaTres 1.2 (2011)
- ThManager 2.0 (2007)

Además, se ha examinado la API GEMET, un servicio web que a pesar de no contar con una herramienta asociada, creemos que tiene interés por soportar SKOS y ajustarse a los criterios de interoperabilidad perseguidos por las tecnologías semánticas.

Existen, no obstante otras herramientas que no han sido incluidas, como ISGAT<sup>9</sup> o Skosify<sup>10</sup>, y que cumplen los criterios indicados, pero que aún no estaban accesibles en los repositorios visitados cuando se inició este trabajo. Sin embargo, dado que el foco de esta propuesta es la metodología de evaluación, se decidió mantener acotado el conjunto de herramientas sobre el que se probaría.

### **4.2 Tesauros utilizados.**

Para llevar a cabo las pruebas se emplearon cuatro tesauros:

- EuroVoc<sup>11</sup>, versión del Tesauro Eurovoc representada con SKOS. Esta representación de EuroVoc se obtuvo transformando los ficheros de EuroVoc en XML facilitados por la Oficina de Publicaciones de la Unión Europea (Alvite *et al.*, 2010). Se trata de un tesauro multilingüe y con polijerarquía. Este tesauro fue empleado para comprobar la importación en formato SKOS.

- UKAT Thesaurus<sup>12</sup>, representación en SKOS/RDF. Se trata de un tesauro disponible públicamente para su descarga en formato texto y en ficheros rtf cuyo contenido es una representación que sigue SKOS-Core 1.0 RDF schema. Se trata de un tesauro con microtesauros. La descarga está disponible en inglés.
- NAL -National Agricultural Library thesaurus<sup>13</sup>. Representación en SKOS/RDF y listado alfabético en ASCII. Tesauro disponible en español e inglés. Se permite la descarga en formato pdf, ASCII y en formatos XML, MARC y SKOS-Core. En las pruebas fue utilizado el formato SKOS-Core.
- Tesauro ad-hoc, un mini-tesauro construido a partir de Eurovoc con 70 términos, 42 descriptores, 24 no descriptores y 2 notas de alcance. Presenta relaciones jerárquicas de dos niveles (TG1, TG2 y TE1, TE2), relaciones de equivalencia (USE y UF) y relaciones asociativas (17 TR). Presenta 8 campos temáticos y 15 subdominios o microtesauros. Para comprobar la gestión de la polijerarquía en los tests se introdujo un término que pertenece a tres microtesauros.

#### 4.3 Resultados de la evaluación.

##### 4.3.1 Descripción y funcionalidades de las herramientas analizadas.

A continuación se presenta una síntesis de los resultados de la aplicación de los criterios evaluativos propuestos en el epígrafe anterior.

MultiTes Pro<sup>14</sup> es una herramienta comercial distribuida por multites.com. La empresa ofrece tres productos en su sede web: MultiTes Pro, MultiTes WDK y MultiTes EDK. La versión de evaluación de MultiTes Pro está disponible para su descarga en la web tras cumplimentar un formulario de registro.

Se trata de un editor de tesauros, una aplicación dirigida a la construcción y mantenimiento de este tipo de lenguajes. Este software requiere Windows 2000, XP o superior. La aplicación necesita un mínimo de 3MB de espacio en disco. Asimismo, estiman unos 15MB de ocupación de disco para 50.000 términos.

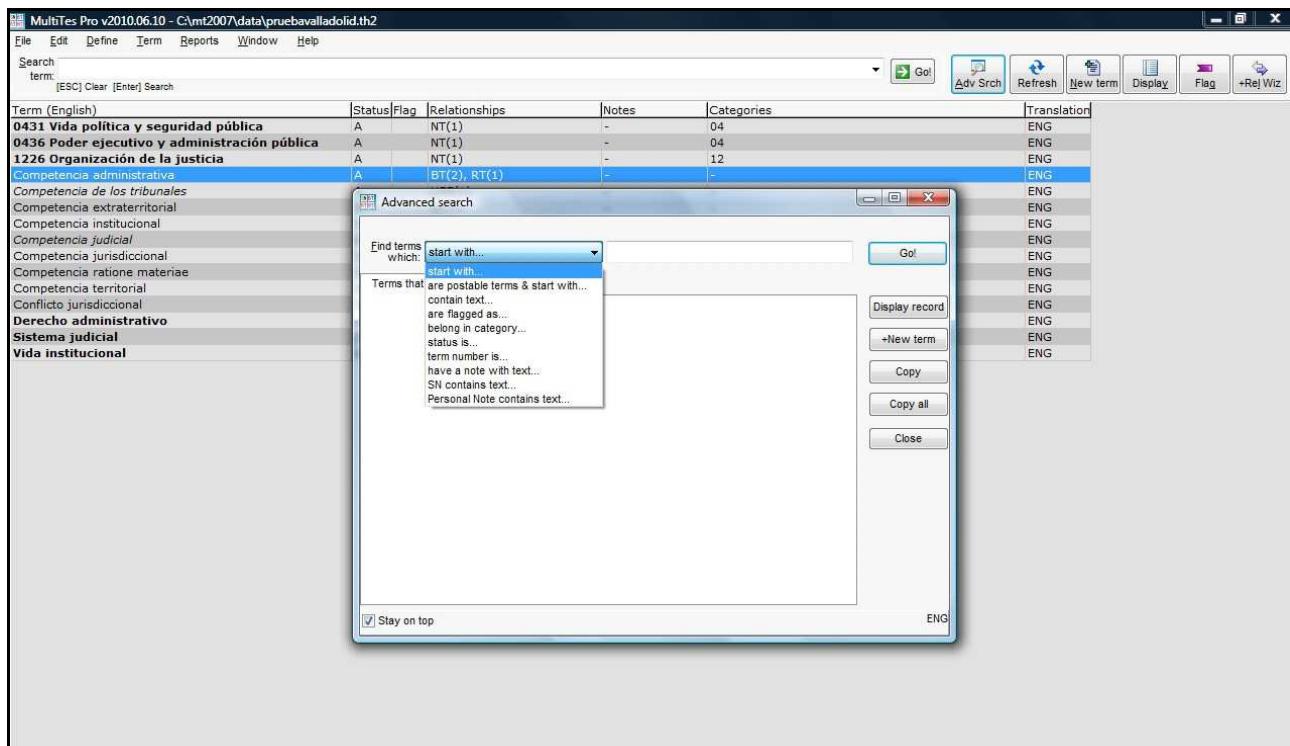


Figura 1. MultiTes Pro.

One-2-One<sup>15</sup> es el sustituto de Term Tree, herramienta lanzada en 1999. Cuenta con la posibilidad de descarga de una licencia de evaluación. Se trata de una herramienta para la gestión de tesauros entendidos de modo tradicional. El software permite la edición de tesauros y clasificaciones, si bien parece más dirigido a la gestión de estas últimas. El sistema requiere la instalación de Access y MS SQL Server.

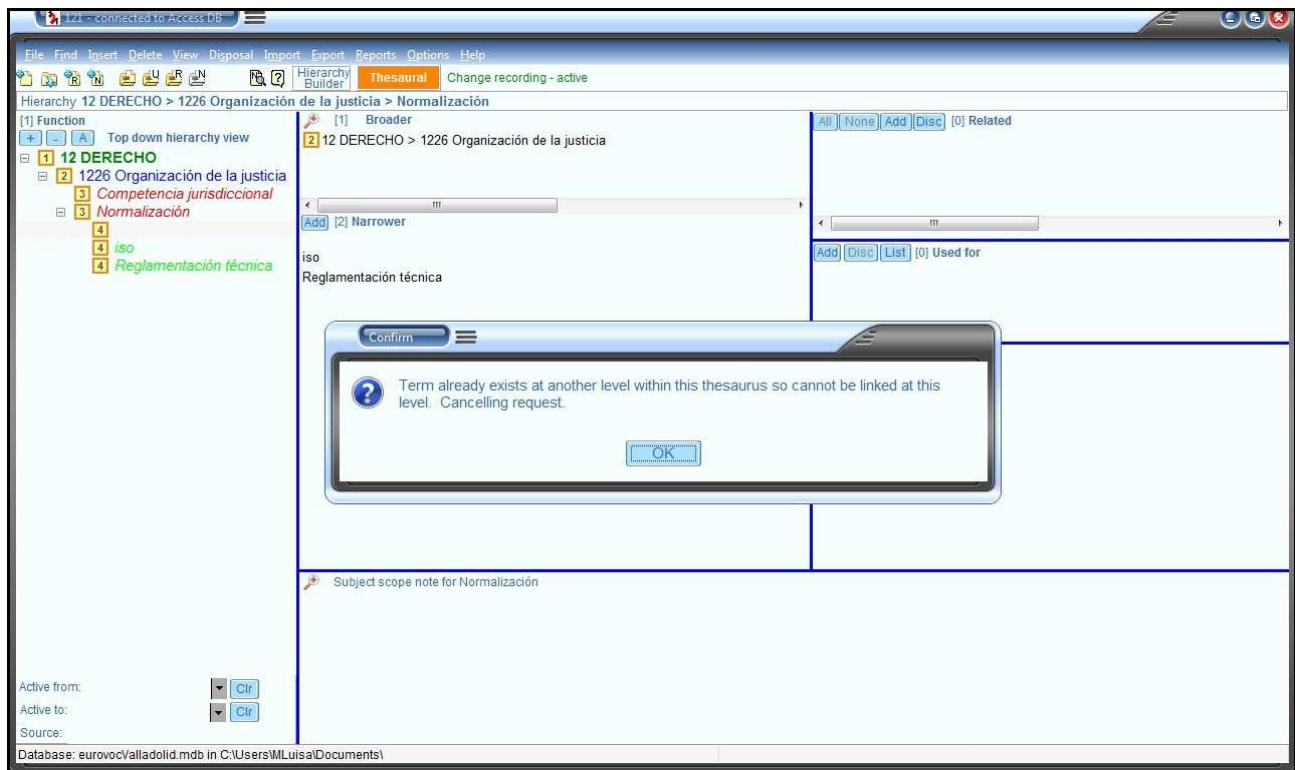


Figura 2. One-2-One.

PoolParty<sup>16</sup> es un sistema de gestión de tesauros dirigido a la web semántica (Schandl y Blumauer, 2010). Se trata de un software comercial que cuenta con una versión de evaluación en la que algunas de las utilidades del programa no están incluidas. PoolParty utiliza SKOS y trabaja con las especificaciones de este estándar: Concepts, ConceptSchemes, etc. Cuenta con una interfaz de usuario sencilla y permite crear y mantener tesauros multilingües.

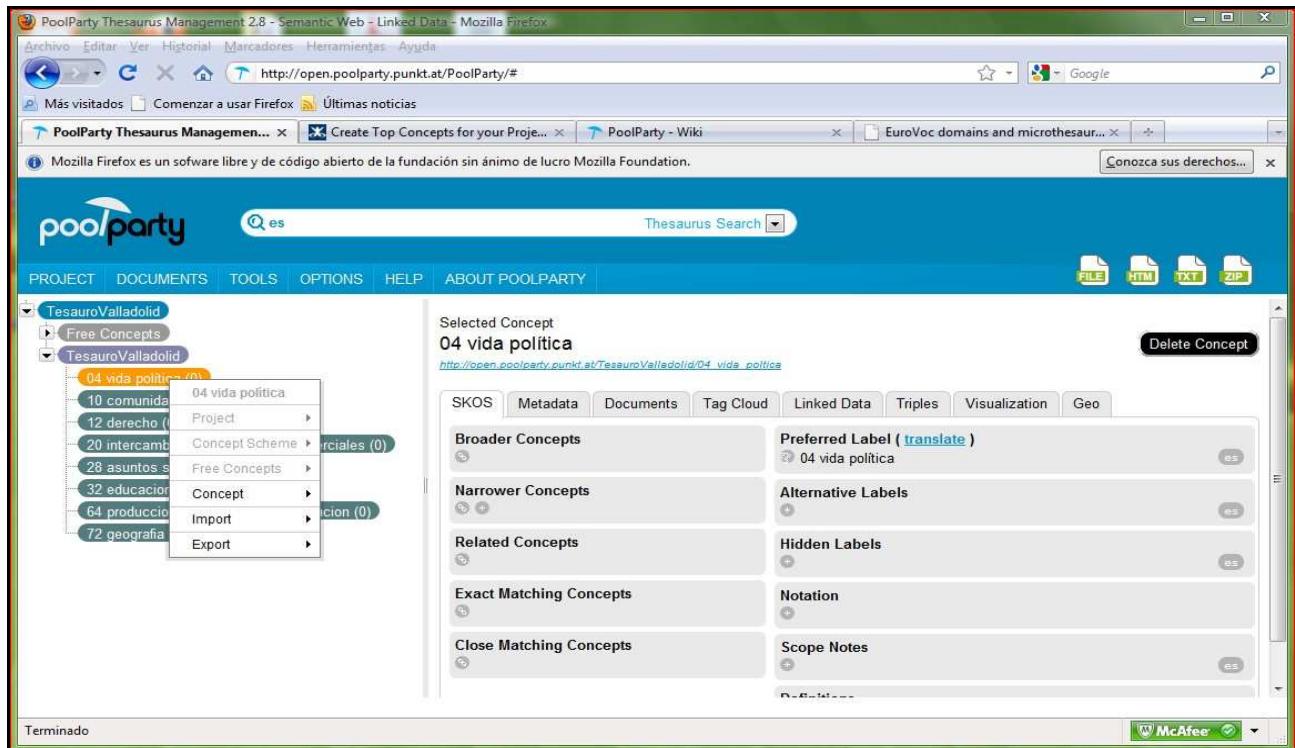


Figura 3. PoolParty.

SKOSED<sup>17</sup> es un plugin de Protégé que permite crear y editar vocabularios controlados representados mediante SKOS. Temates<sup>18</sup> es una aplicación web desarrollada en PHP, con licencia GPL, que permite la gestión y la consulta de tesauros y taxonomías. Por su parte, ThManager es una herramienta de código abierto que permite crear y visualizar vocabularios SKOS.

En lo que se refiere a los requisitos del sistema, la versión de evaluación de PoolParty es una aplicación web accesible que requiere Firefox. La interfaz despliega la información en dos áreas, que se muestran en la Figura 3: el árbol jerárquico de los conceptos y la vista de detalle de los mismos. En esta última opción, muestra distintas pestañas para elegir entre la visualización SKOS, el examen de metadatos, edición de linked data, vista de triplets RDF, visualización gráfica o una vista de datos Geo.

SKOSED es un plugin del editor de ontologías OWL, Protégé<sup>19</sup>. Esta aplicación permite crear y editar vocabularios representados en SKOS. Se trata de un software claramente orientado a SKOS, frente a herramientas tradicionales dirigidas a la edición de tesauros o incluso herramientas para tesauros que usan SKOS, como PoolParty. Introduce ciertas variaciones con respecto a las notaciones comunes en estos lenguajes normalizados, así aparece la relación “Narrower Transitive” y “Broader Transitive”. En cierto sentido, el usuario puede sentirse desbordado por la aplicación si no está familiarizado con las funcionalidades de Protégé o con editores de ontologías.

SKOSED es una extensión de Protégé que presenta un conjunto de menús y vistas útiles. Una de las utilidades significativas es la opción de descargar la ontología SKOS, como se observa en la Figura 4. La finalidad o propósito de esta aplicación coincide con la de Protégé, esto es, la edición de ontologías, si bien en el caso de SKOSED esta edición se restringe a ontologías SKOS. En cuanto a los requisitos del sistema, el primordial es la instalación de Protégé 4, una aplicación disponible para Windows, Unix y sistemas Mac.

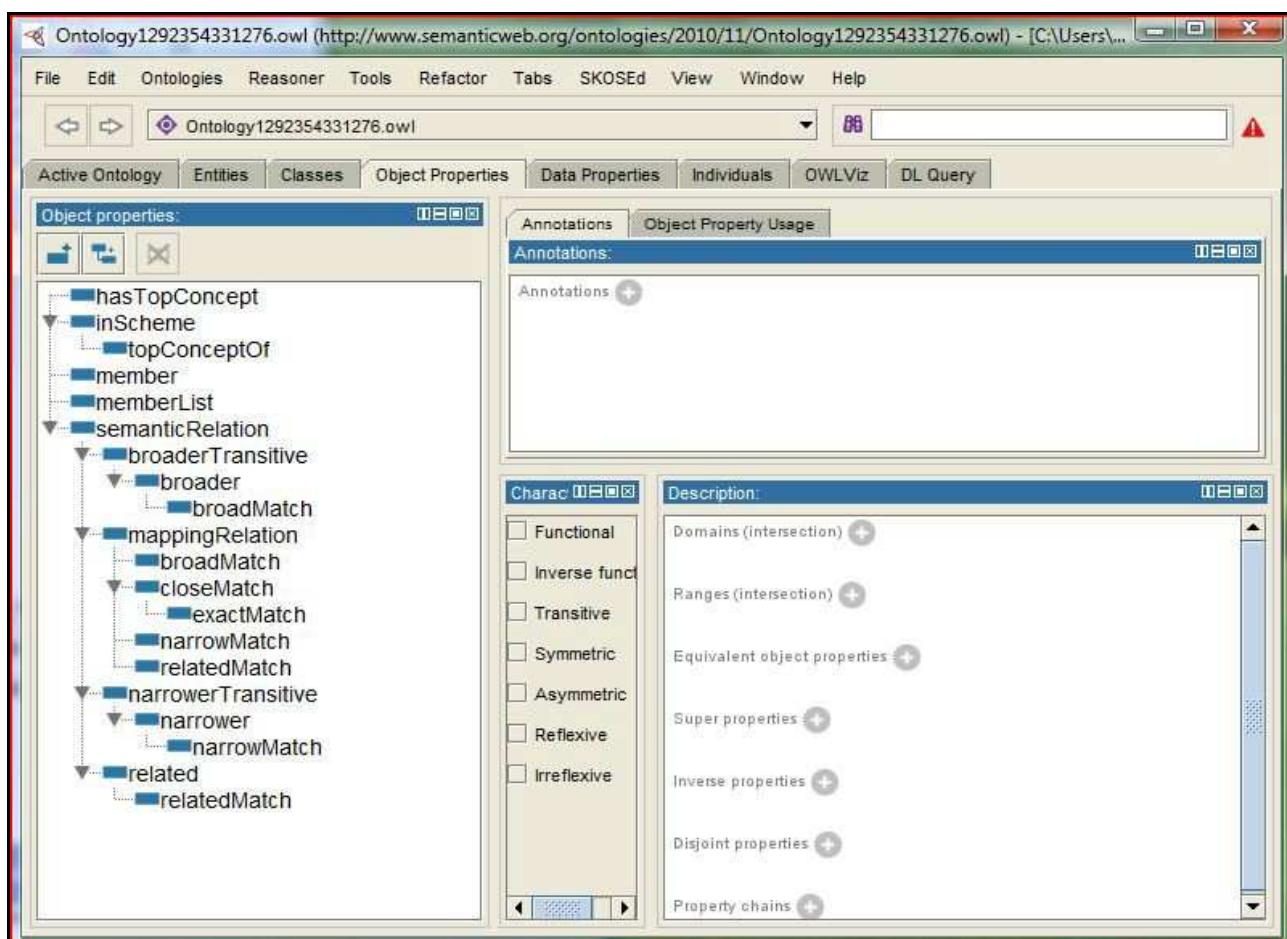


Figura 4. SKOSED.

TemaTres<sup>20</sup> hace posible la edición de vocabularios controlados por medio de una sencilla interfaz web. Cuenta con una edición web libre que puede ser usada para navegar y hacer búsquedas en un conjunto dado de vocabularios. Para crear y editar un tesauro propio dispone de un kit de software, descargable desde la sede web sourceforge.net, que requiere ser instalado en un servidor web local.

Se trata de una herramienta dirigida a desarrolladores tradicionales de tesauros, como puede apreciarse en la Figura 5. Así, en la aplicación los tesauros se componen de términos que pueden ser propuestos para su inclusión y ser aceptados o rechazados por los responsables finales del vocabulario. Además, TemaTres se asienta fundamentalmente en funcionalidades comunes de navegación con las que los usuarios no expertos en aplicaciones avanzadas o sofisticadas están perfectamente familiarizados. Cuenta con extensiones de búsqueda de términos en Google, Google Académico, Wikipedia, etc. La sencillez de la interfaz y la terminología empleada en la herramienta son funcionalidades que, sin duda, se adecuan a los desarrolladores tradicionales de tesauros.

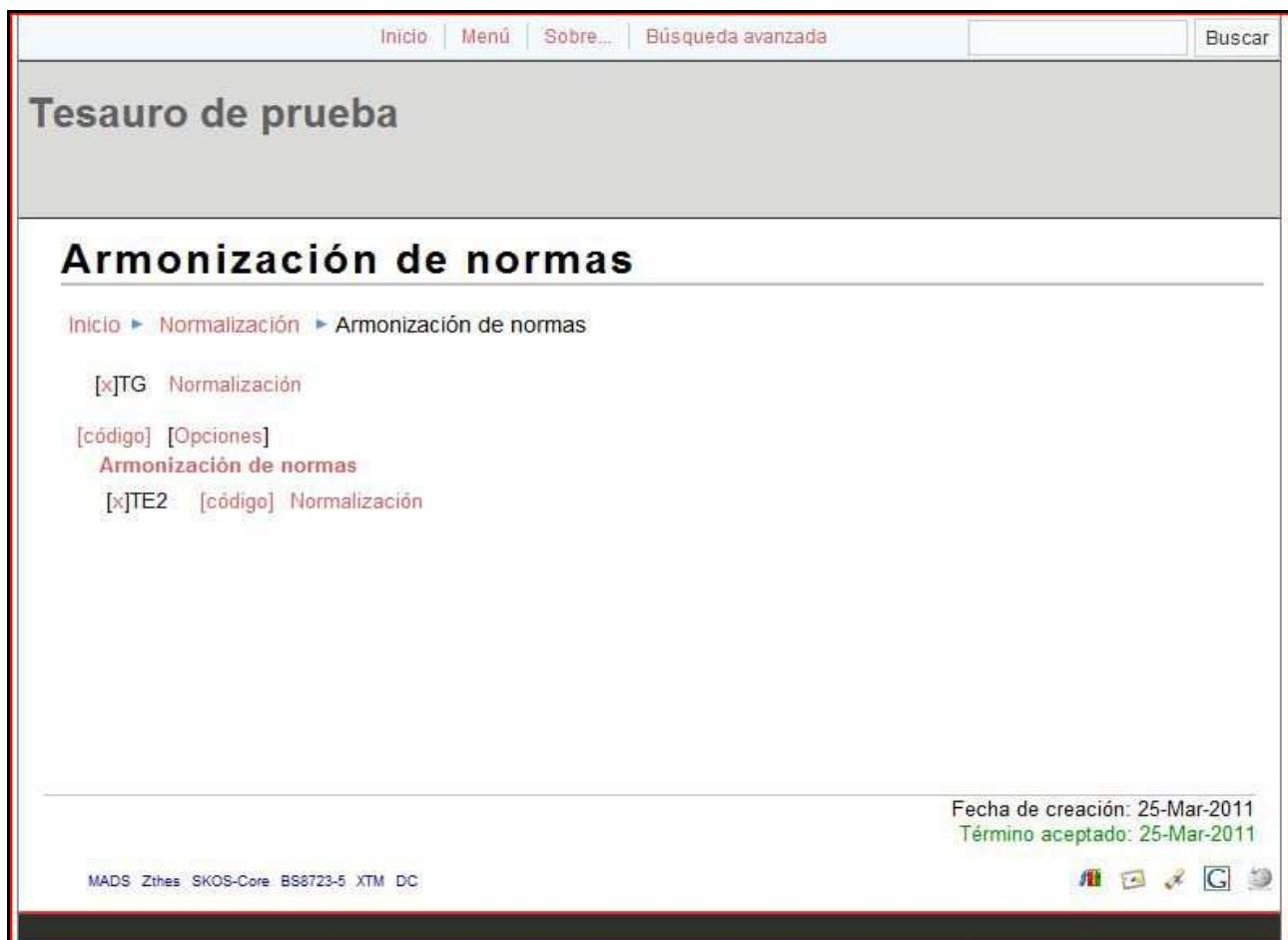


Figura 5. TemaTres.

El propósito de esta herramienta se dirige a facilitar la edición de vocabularios controlados, incluyendo tesauros, taxonomías y listas de encabezamientos de materia, etc. Disponible para Windows, Mac OS y Linux (Gonzales Aguilar; Ramírez Posada y Ferreyra, 2012). En cuanto a los requisitos del sistema, la edición completa de TemaTres requiere la instalación de un servidor web como Apache, base de datos MySQL y PHP.

ThManager<sup>21</sup>, en la Figura 6, es software libre, distribuido bajo licencia GNU (Lesser General Public License). Esta herramienta gestiona tesauros representados en SKOS, sin embargo, su interfaz emplea el vocabulario clásico empleado en los tesauros, esto es, un vocabulario compuesto por términos. Cabe destacar la particularidad de emplear los vocablos “hijo” y “padre” para referirse a las relaciones expresadas en el tesauro como término específico y término genérico. La aplicación presenta un detallado manual de usuario.

Esta herramienta actúa como un editor de tesauros que permite igualmente representar los mismos en SKOS para su importación, edición y gestión. La aplicación requiere que la máquina virtual de Java esté instalada en el equipo.

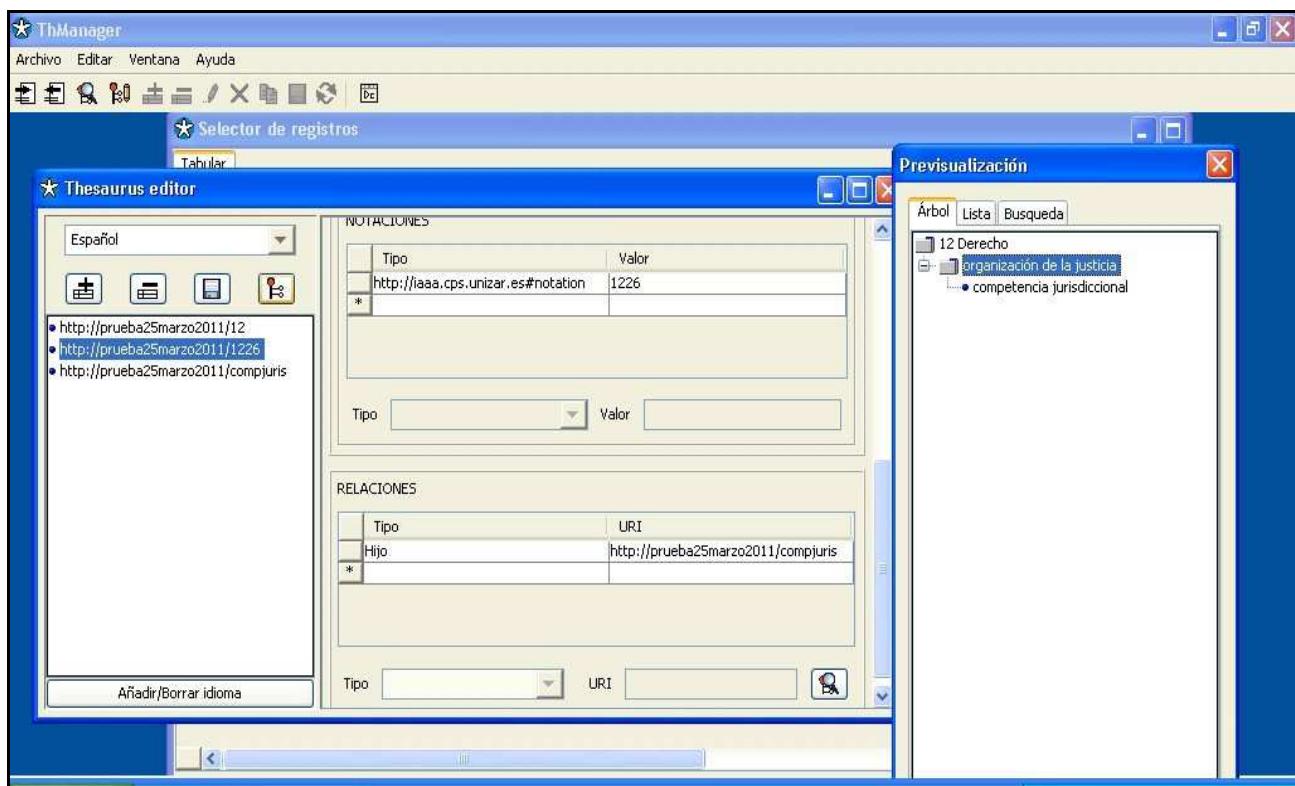


Figura 6. ThManager.

Finalmente, la API de GEMET, descrita en GEMET-webserviceAPI, ofrece alrededor de doce métodos que permiten la recuperación de términos cabecera (getTopmostConcepts), de las diferentes traducciones de un concepto (getAllTranslationsForConcept), de conceptos que coinciden con una expresión dada (getConceptsMatchingRegexByThesaurus), etc. Esta API basada en la propuesta API web service de SKOS introduce ciertas modificaciones para resolver cuestiones como el multilingüismo y la posibilidad de emplearla sobre interfaces ReST y XML-RPC.

La propuesta de API GEMET ReST-ful, originariamente basada en la API de la National Agricultural Library Thesaurus<sup>22</sup> es la propuesta REST para la API GEMET. Para cada recurso proporciona posibles representaciones y los parámetros correspondientes que pueden ser utilizados con la operación HTTP GET para recuperar esa representación. El acceso a los recursos GEMET (thesaurus collection, thesauri, concepts, themes, groups, INSPIRE terms y relationships) se realiza a través de operaciones HTTP GET. La documentación incluye la descripción de las clases y el método de las interfaces junto con ejemplos de uso.

La Tabla I resume las funcionalidades de las herramientas examinadas.

	<b>MultiTes</b>	<b>One-2-One</b>	<b>PoolParty</b>	<b>SKOSED</b>	<b>TemaTres</b>	<b>ThManager</b>
<b>Creación y edición</b>	Sí	Sí	Sí (Metadatos Dublin Core)	Sí (Metadatos Dublin Core)	Sí (Metadatos Dublin Core)	Sí (Metadatos Dublin Core)
<b>Búsqueda y recuperación</b>	Búsqueda de cadenas de texto en etiquetas, categorías, estado, tipos y notas	Búsqueda de cadenas de texto en términos y notas	Búsqueda de cadenas de texto en las etiquetas Preferred, Alternative y Hidden	Búsqueda a través de Protégé DL Queries	Búsqueda de cadenas de texto en los descriptores	Búsqueda de cadenas de texto en los descriptores
<b>Navegación</b>	A través de las relaciones	A través de las relaciones, jerarquía y listas	A través de las relaciones y jerarquía	A través de las relaciones y en las vistas SKOS	A través de las relaciones y listas	A través de las relaciones, jerarquía y listas
<b>Fusión de tesauros</b>	No	Sí	Sí (desde ficheros SKOS)	Sí (desde Protégé ontology import)	No	No
<b>Importación y exportación</b>	Ambas	Ambas	Ambas	Ambas	Ambas	Ambas
<b>Ficheros log</b>	No	Sí	No	No	No	No
<b>Colaboración de usuarios</b>	Sí (en redes LAN)	No	No	No	Sí	No

Tabla I. Funcionalidades.

Las funcionalidades básicas exigibles están presentes en todas las aplicaciones. Además, se da soporte en muchos de los casos a Dublin Core. Hay algunas diferencias en las búsquedas, aunque la mayoría de las herramientas permiten buscar cadenas de texto en distintas formas. No obstante, mientras algunas aplicaciones, tales como ThManager y TemaTres, solo permiten la búsqueda de descriptores o términos preferentes, otras herramientas expanden la búsqueda a las etiquetas alternativas, notas e incluso etiquetas ocultas.

En general, no es posible combinar varios tesauros. La fusión de tesauros, que podría tener utilidad en algunos casos, se soporta en One-2-One y en las herramientas que incorporan tecnologías propias de la web semántica. Resulta un aspecto congruente con las características de estos softwares, dado que la fusión de ontologías es un tema importante para los desarrolladores de ontologías en el contexto de la web semántica (Euzenat y Shvaiko, 2007).

Las utilidades de importación y exportación de tesauros están implementadas en las aplicaciones, si bien se observan variaciones importantes con respecto a la garantía de interoperabilidad de la información, como comentaremos en las reflexiones finales. Los ficheros *log* no parecen revestir importancia para los desarrolladores de estas herramientas; con la excepción de One-2-One esta funcionalidad está ausente en todas las aplicaciones. Por último, sólo TemaTres soporta la edición colaborativa de tesauros.

#### 4.3.2 Estructuras conceptuales y terminológicas soportadas.

Las construcciones de tesauros compatibles con cada herramienta se muestran en la Tabla II. La primera fila se refiere al vocabulario básico que presenta la aplicación para la gestión del vocabulario: la estructura clásica de Tesauro y Términos del tesauro o bien *Concept* y *ConceptScheme* de SKOS. Esta clara diferenciación es indicativa de la

orientación de la herramienta, enfocada a SKOS, o si ha sido concebida desde una aproximación a la construcción clásica de tesauros, con independencia de la representación de estos vocabularios.

	<b>MultiTes</b>	<b>One-2-One</b>	<b>PoolParty</b>	<b>SKOSED</b>	<b>TemaTres</b>	<b>ThManager</b>
<b>Estructuras básicas</b>	Tesauros, Términos	Tesauros, Términos	Concept-Scheme, Concept	Concept-Scheme, Concept	Tesauros, Términos	Tesauros, Términos
<b>Términos preferentes / Términos no preferentes</b>	Sí	Sí	Sí	Sí	Sí	Sí
<b>Relaciones de equivalencia</b>	Sí	Sí	Sí	Sí	Sí	Sí
<b>Relaciones jerárquicas</b>	Sí	Sí	Sí	Sí	Sí	Sí
<b>Relaciones asociativas</b>	Sí	Sí	Sí	Sí	Sí	Sí
<b>Notas</b>	Alcance, Privadas	Alcance	Notas SKOS	Notas SKOS	Alcance, Históricas, Catálogo, Privadas	Alcance, Definición

**Tabla II. Estructuras conceptuales y terminológicas soportadas.**

Aunque SKOS se confirma como el estándar indiscutible para la representación de tesauros, es preciso conocer cómo representar las estructuras de los tesauros en las construcciones establecidas por SKOS, lo que puede resultar una limitación para determinados usuarios.

Como se observa en la Tabla II, todos los softwares aquí examinados soportan las construcciones básicas necesarias para la gestión de un tesauro. Se aprecian, no obstante, diferencias significativas en el rango de las posibles notas incluidas en el vocabulario. Solamente SKOSED permite el uso de *Collections* SKOS, un aspecto significativo de lo limitado de este tipo de construcción para los creadores y usuarios de tesauros.

#### **4.3.3 Integridad, interoperabilidad e integrabilidad.**

En la Tabla III se muestran de modo sumario los resultados obtenidos respecto a la integridad o consistencia de la información, la interoperabilidad y los problemas de integrabilidad.

	<b>MultiTes</b>	<b>One-2-One</b>	<b>PoolParty</b>	<b>SKOSED</b>	<b>TemaTres</b>	<b>ThManager</b>
<b>Relaciones inversas automáticas</b>	Sí	Sí	Sí	Sí	Sí	No
<b>Borrado automático de relaciones</b>	Sí	Parcial	Sí	Sí	Sí	No
<b>Impedir relaciones jerárquicas inconsistentes</b>	Sí	Sí	Sí	Depende del razonador seleccionado	No	No
<b>Importación de tesauros</b>	Txt	Trim Ascii, eDocs, etc.	SKOS (Turtle, N3, N-triples)	SKOS	Ficheros Tab	SKOS
<b>Exportación de tesauros</b>	Txt, XML, SKOS	Excel, CSV, XML, SKOS	SKOS (Turtle, N3, N-triples)	SKOS	Txt, SKOS Core, Zthes, MADS	SKOS
<b>Integrabilidad</b>	Servicios web	No	Servicios web	API SKOS	Servicios web	No

**Tabla III. Integridad, interoperabilidad e integrabilidad.**

Hay diferencias notables en la gestión de la integridad. MultiTes y PoolParty superaron todos los tests realizados. En el extremo opuesto se situaría ThManager, quedando en una posición intermedia One-2-One, SKOSED y TemaTres, que superaron algunas de las pruebas, pero no todas. En algún caso, creemos que la inconsistencia puede derivar del modo en que el software gestiona la jerarquía en el vocabulario. Por su parte, en el caso concreto de SKOSED los errores de integridad parecen tener su origen en el hecho de que los conceptos del tesauro no son clases de una ontología, sino instancias de las clases de la ontología.

Las herramientas que utilizan tecnologías semánticas, tales como PoolParty o SKOSED, soportan un abanico amplio de formatos propios de la web semántica -RDF, N3, Turtle, etc. - para la importación y exportación de tesauros. Esto es lógico, debido a que estos formatos están asociados de modo habitual a RDF. Parece recomendable que, atendiendo al grado de implementación de SKOS como estándar, las herramientas de gestión de tesauros garanticen este formato para la importación y exportación del vocabulario.

Una API general y estándar para tesauros parece ser un ideal que no se intuye cercano. De hecho, el abandono de la API SKOS no parece ser un buen indicio sobre una posible API disponible en un tiempo razonable. En la actualidad, parece que cada gran tesauro o herramienta de gestión de tesauros desarrolla su propia API para su propio acceso. En este sentido, cabe destacar la API GEMET, exhaustiva y muy bien documentada, que ofrece ideas muy útiles sobre lo que una API de tesauro debe contener, o su implementación en servicios web.

La API de SKOS (Jupp *et al.*, 2009) es un esfuerzo interesante en materia de gestión de SKOS. Es una aplicación de código abierto bajo licencia LGPL. El fichero *.jar* está disponible para su descarga y su diseño se dirige a manipular Sistemas de Organización del Conocimiento representados con SKOS. En cuanto a la aplicación de TemaTres, basada en el enfoque de servicio web, presenta como gran ventaja la simplicidad que anima al usuario a emplear esta herramienta de gestión de tesauros como servicio web.

## 5. REFLEXIONES FINALES.

Se ha propuesto una metodología para llevar a cabo una evaluación de herramientas de gestión de tesauros que permite un análisis sistemático de estas aplicaciones en orden a conocer tanto su cumplimiento con respecto a las funcionalidades clásicas exigidas a estas herramientas, como su grado de adaptación a las exigencias derivadas de la adaptación a la web semántica.

La metodología propuesta en este trabajo toma como referencia estudios previos de evaluación de gestores de tesauros así como propuestas genéricas para la evaluación de herramientas vinculadas a la web semántica. No obstante, esta propuesta viene a cubrir un vacío en cuanto a la existencia de una metodología específicamente diseñada para evaluar herramientas de tesauros incluyendo entre sus criterios aquellos capaces de reflejar los últimos avances, motivados por el auge de la web semántica, que han afectado a la representación de estos vocabularios controlados. Así pues, esta propuesta aúna dos valores. Primero, estar diseñada particularmente para herramientas de tesauros, lo cual entendemos resulta útil a los usuarios de estas herramientas que buscan una guía para su selección y que, de otro modo, se verían obligados a adaptarla a partir de propuestas genéricas. Segundo, incluir los criterios que reflejan la evolución y estado más actual de la representación de tesauros, vinculada a la evolución de los estándares surgidos en el marco de la web semántica.

El análisis realizado para aplicar los criterios evaluativos propuestos nos ha permitido observar una clara distinción entre aplicaciones tradicionales de gestión de tesauros y software orientado hacia tecnologías semánticas. Dentro de este segundo grupo de herramientas se distinguen aquellas que como PoolParty aportan sobre SKOS las restricciones características de los tesauros frente a otras como SKOSED, genérica para KOS representados con SKOS, que se limita a lo que señala la ontología SKOS sin atender a las normas tradicionales de construcción de tesauros. En nuestra opinión, en un futuro cercano, se decidirá qué tendencia resulta predominante: aplicaciones dirigidas a tesauros o aplicaciones orientadas a SKOS.

Herramientas como PoolParty o SKOSED, apoyan a una gama más amplia de formatos de la web semántica -RDF, Turtle, N3, etc., para la importación y exportación de tesauros. Igualmente son estos programas los que permiten la fusión de tesauros.

Los servicios web con libre acceso a la documentación (GEMET, TemaTres) cubren las utilidades de consulta. No obstante, el conjunto de funcionalidades relacionadas con la creación y edición de tesauros no parece del todo resuelto. En cuanto a las herramientas propietarias, tales como PoolParty, no han podido ser examinadas por completo dado que

el acceso a la documentación está restringido a los clientes del producto. Los resultados del examen de las aplicaciones nos llevan a abogar por una API general y estándar para la gestión de tesauros.

En nuestra opinión, en un futuro próximo, veremos una evolución que aclarará si la preferencia de este tipo de software se decanta hacia APIs e interfaces orientadas a SKOS o APIs y herramientas específicas para tesauros.

## NOTAS

<sup>1</sup> Se utiliza el término colección con el significado que tiene en la Recomendación de SKOS.

<sup>2</sup> <<http://agclass.nal.usda.gov/agt/agt.shtml>>.

<sup>3</sup> <<http://www.w3.org/2001/sw/Europe/reports/thes/skosapi.html>>.

<sup>4</sup> Tools for managing taxonomies (or Thesauri, or Ontologies). <<https://www.earley.com/blog/tools-managing-taxonomies-or-thesauri-or-ontologies>>.

<sup>5</sup> Software for building and editing thesauri. <<http://www.willpowerinfo.co.uk/thessoft.htm>>.

<sup>6</sup> W3C Tools - SKOS wiki. <<http://www.w3.org/2004/02/skos/wiki/Tools>>.

<sup>7</sup> W3C SKOS community wiki. <<http://www.w3.org/2001/sw/wiki/SKOS>>.

<sup>8</sup> SKOS page in Wikipedia. <[http://en.wikipedia.org/wiki/-Simple\\_Knowledge\\_Organization\\_System#Tools](http://en.wikipedia.org/wiki/-Simple_Knowledge_Organization_System#Tools)>.

<sup>9</sup> <<http://www.um.es/igat>>.

<sup>10</sup> <<http://code.google.com/p/skosify/>>.

<sup>11</sup> <<http://eurovoc.europa.eu/>>.

<sup>12</sup> <<http://www.ukat.org.uk/thesaurus/>>.

<sup>13</sup> <<http://agclass.nal.usda.gov/>>.

<sup>14</sup> <<http://www.multites.com/>>.

<sup>15</sup> <<http://www.acs121.com/one2one.html>>.

<sup>16</sup> <<http://www.poolparty.biz/>>.

<sup>17</sup> <<https://code.google.com/p/skoseditor/>>.

<sup>18</sup> <<http://www.r020.com.ar/tematres/>>.

<sup>19</sup> <<http://protege.stanford.edu/>>.

<sup>20</sup> <<http://www.r020.com.ar/tematres/>>.

<sup>21</sup> <<http://thmanager.sourceforge.net/>>.

<sup>22</sup> <<http://agclass.nal.usda.gov/agt/agt.shtml>>.

## 6. BIBLIOGRAFÍA.

ALVITE DÍEZ, M.L.; PÉREZ LEÓN, B.; MARTÍNEZ GONZÁLEZ, M.M. y VICENTE BLANCO, D.J. Propuesta de representación del tesauro Eurovoc en SKOS para su integración en sistemas de información jurídica. *Scire: Representación y Organización del Conocimiento*, 2010, vol. 16, nº 2, p. 47-51.

BLOCH, J.J. How to design a good API and why it matters. En: TARR, P.L.; COOK, W.R. (eds.) *OOPSLA Companion*. New York: ACM Press, 2006, p. 506-507.

EUZENAT, J. y SHVAIKO, P. *Ontology matching*. Berlín: Springer, 2007.

FARO, S.; FRANCESCONI, E.; MARINAI, E. y SANDRUCCI, V. *EUROVOC Studies LOT2 D2.3. Report on execution and results of the interoperability tests*. Tech. Rep. 10118 [en línea]. Florencia: Publications Office of the EC, Institute of Legal Information Theory and Techniques ITTIG, 2008. Disponible en: <[http://eurovoc.europa.eu/drupal/sites/all/files/Presentation-D2.3\\_Mapping\\_results.pdf](http://eurovoc.europa.eu/drupal/sites/all/files/Presentation-D2.3_Mapping_results.pdf)> [Consulta: 3 de septiembre de 2013].

GARCÍA CASTRO, R. y GÓMEZ PÉREZ, A. Guidelines for benchmarking the performance of ontology management APIs [en línea]. En: *Proceedings of the 4th International Semantic web Conference (ISWC2005) LNCS 3729*, 2005, p. 277-292. Disponible en: <[http://oa.upm.es/4866/1/ISWC\\_2005.pdf](http://oa.upm.es/4866/1/ISWC_2005.pdf)> [Consulta: 6 de septiembre de 2013].

GARCÍA CASTRO, R.; GÓMEZ PÉREZ, A. y SURE, Y. Benchmarking the RDF(S) interoperability of ontology tools [en línea]. *Proceedings of the 19th International Conference on Software Engineering and Knowledge Engineering (SEKE2007)*. Boston, USA: Knowledge Systems Institute, 2007. Disponible en: <[http://oa.upm.es/4212/1/SEKE\\_2007.pdf](http://oa.upm.es/4212/1/SEKE_2007.pdf)> [Consulta: 6 de septiembre de 2013].

GARCÍA CASTRO, R. y GÓMEZ PÉREZ, A. Benchmarking in the semantic web [en línea]. En: CARDOSO, J.; LYTRAS, M. (eds.) *Semantic web Engineering in the Knowledge Society*. Information Science Reference IGI Global, 2009, p. 341-370. Disponible en: <[http://oa.upm.es/3993/1/Benchmarking\\_in\\_the\\_Semantic\\_web.pdf](http://oa.upm.es/3993/1/Benchmarking_in_the_Semantic_web.pdf)> [Consulta: 6 de septiembre de 2013].

- GEMET-ReST. Some Thoughts on a ReST-ful Interface for the new GEMET API [en línea]. *Taskman project: Software development EEA website systems*. Disponible en: <<http://taskman.eionet.europa.eu/projects/zope/wiki/ReSTInterface>> [Consulta: 9 de septiembre de 2013].
- GEMET-General Multilingual Environmental Thesaurus [en línea]. *Taskman project: Software development EEA website systems*. Disponible en: <<http://taskman.eionet.europa.eu/projects/zope/wiki/GEMET>> [Consulta: 9 de septiembre de 2013].
- GEMET websevice API [en línea]. *Taskman project: Software development EEA website*. Disponible en: <<http://taskman.eionet.europa.eu/projects/zope/wiki/GEMETwebServiceAPI>> [Consulta: 9 de septiembre de 2013].
- GERBE, O. y KERHERVE, B. A Model-Driven Approach to SKOS Implementation. En: *Internet and web Applications and Services (ICIW)*. IEEE Computer Society, 2010, p. 484–488.
- GIL URDICIAIN, B. *Manual de lenguajes documentales*. 2ªed. revisada y ampliada. Gijón: Trea, 2004.
- GONZALES AGUILAR, A.; RAMÍREZ POSADA, M. y FERREYRA, D. TemaTres: software para gestionar tesauros. *El Profesional de la Información*, 2012, vol. 21, nº 3, p. 319-325.
- HAAS, H. y BROWN, A. *Web Services Glossary* [en línea]. W3C Working Group, Note 11 February 2004. Disponible en: <<http://www.w3.org/TR/ws-gloss/>> [Consulta: 9 de septiembre de 2013].
- HILDEBRAND, M.; VAN OSSENBRUGGEN, J.; HARDMAN, L. y JACOBS, G. Supporting subject matter annotation using heterogeneous thesauri: A user study in web data reuse. *International Journal of Human-Computer Studies*, 2009, vol. 67, nº 10, p. 887-902.
- ISAAC, A. y SUMMERS, E. *SKOS Simple Knowledge Organization System Primer* [en línea]. W3C Working Group, Note 18 August 2009. Disponible en: <<http://www.w3.org/TR/skos-primer/>> [Consulta: 9 de septiembre de 2013].
- ISO 25964-1:2011 Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval, 2011.
- ISO 25964-1:2011 Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 2: Interoperability with other vocabularies, 2013.
- ISO TC46/SC9/WG8; ISAAC, A. *Correspondence between ISO25964 and SKOS/SKOS-XL models* [en línea]. NISO, 2012. Disponible en: <[http://www.niso.org/apps/group\\_public/download.php?9627/Correspondence%20ISO25964-SKOSXL-MADS-2012-10-21.pdf](http://www.niso.org/apps/group_public/download.php?9627/Correspondence%20ISO25964-SKOSXL-MADS-2012-10-21.pdf)> [Consulta: 9 de septiembre de 2013].
- JUPP, S.; BECHHOFER, S. y STEVENS, R. A flexible API and editor for SKOS. En: *Proceedings of the 6th European Semantic web Conference on The Semantic web: Research and Applications (ESWC 2009 Heraklion)*. Berlin: Springer-Verlag, 2009, p. 506-520.
- KOLLER, A. SKOS Thesaurus Management based on Linked Data with Poolparty [en línea]. En: *Third Annual European Semantic Technology Conference, 2 y 3 de diciembre de 2009*. Vienna, 2009. Disponible en: <[http://triplify.org/files/challenge\\_2009/poolparty.pdf](http://triplify.org/files/challenge_2009/poolparty.pdf)> [Consulta: 11 de septiembre de 2013].
- LACASTA, J.; NOGUERAS ISO, J.; LÓPEZ PELLICER, F.J.; MURO MEDRANO, P.R. y ZARAZAGA SORIA, F.J. ThManager: An Open Source Tool for creating and visualizing SKOS. *Information Technology and Libraries (ITAL)*, 2007, vol. 26, p. 39-51.
- MAYR, P.; ZAPILKO, B. y SURE, Y. Establishing a multi-thesauri-scenario based on SKOS and cross-concordances [en línea]. En: *Proceedings of the 2010 International Conference on Dublin Core and Metadata Applications*. Pittsburgh, 2010, p. 194-196. Disponible en: <<http://dcpapers.dublincore.org/pubs/article/view/1031>> [Consulta: 11 de septiembre de 2013].
- MILES, A. y BECHHOFER, S. *SKOS Simple Knowledge Organization System Reference. W3C Working Draft 25 January 2008* [en línea]. W3C, 2008a. Disponible en: <<http://www.w3.org/TR/2008/WD-skos-reference-20080125/>> [Consulta: 11 de septiembre de 2013].
- MILES, A. y BECHHOFER, S. *SKOS Simple Knowledge Organization System Reference. W3C Working Draft 9 June 2008* [en línea]. W3C: 2008b. Disponible en: <<http://www.w3.org/TR/2008/WD-skos-reference-20080125/>> [Consulta: 11 de septiembre de 2013].
- MILES, A. y BECHHOFER, S. *SKOS Simple Knowledge Organization System Reference. W3C Recommendation 18 August 2009* [en línea]. W3C, 2009. Disponible en: <<http://www.w3.org/TR/2009/PR-skos-reference-20090818/>> [Consulta: 11 de septiembre de 2013].
- MOYA MARTÍNEZ, G. y GIL LEIVA, I. Evaluación de softwares de gestión de tesauros. *Ciencias de la Información*, 2001, vol. 32, nº 3, p. 3-23.
- NOY, N.F.; SINTEK, M.; DECKER, S.; CRUBEZY, M.; FERGERSON, R.W. y MUSEN, M.A. Creating Semantic web Contents with Protégé-2000. *IEEE Intelligent Systems*, 2001, vol. 16, nº 2, p. 60-71.
- PASTOR SÁNCHEZ, J.A. ISO-THESS: ampliando Skos a partir de la norma de tesauros ISO 25964. *Anuario ThinkEPI*, 2013, vol. 7, p. 189-193.
- PASTOR SÁNCHEZ, J.A. *Tecnologías de la web semántica*. Barcelona: Editorial UOC, 2011.

- PASTOR SÁNCHEZ, J.A. *Diseño de un sistema colaborativo para la creación y gestión de tesauros en Internet basado en SKOS*. Tesis doctoral. Universidad de Murcia, 2009.
- PASTOR SÁNCHEZ, J.A.; MARTÍNEZ MÉNDEZ, F.J. y RODRÍGUEZ MUÑOZ, J.V. Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives [en línea]. *Information Research*, 2009, vol. 14, nº 4. Disponible en: <<http://informationr.net/ir/14-4/paper422.html>> [Consulta: 11 de septiembre de 2013].
- POLO PAREDES, L.; ÁLVAREZ RODRÍGUEZ, J.M. y RUBIERA AZCONA, E. Promoting Government Controlled Vocabularies for the Semantic web: the EUROVOC Thesaurus and the CPV Product Classification System. En: *Semantic Interoperability in the European Digital Library: Proceedings of the First International Workshop, SIEDL 2008, Tenerife, June 2, 2008. Lecture Notes in Computer Science 50212*, 2008, p. 111-122.
- SCHANDL, T. y BLUMAUER, A. Poolparty: Skos thesaurus management utilizing linked data [en línea]. En: Aroyo, L. et al. (eds.). *The Semantic web: Research and Applications, Lecture Notes in Computer Science 6089*, 2010, p. 421-425. Disponible en: <<http://www.poolparty.biz/wp-content/uploads/2013/04/PoolParty-SKOS-Thesaurus-Management-utilizing-Linked-Data.pdf>> [Consulta: 11 de septiembre de 2013].
- SEVERINO, F. *The term development in the thesauri of international organisations*. London: Frank Cass, 2007.
- SMEDT, J.D. SKOS Extensions for the EUROVOC Thesaurus. En: *Third Annual European Semantic Technology Conference (ESTC)*, Vienna, 2009.
- UNE: 50106-1990, Documentación. *Directrices para el establecimiento y desarrollo de tesauros monolingües*. Madrid: AENOR, 1990.
- WILLPOWER INFORMATION. *Software for building and editing thesauri*. 2012. Disponible en: <<http://www.willpowerinfo.co.uk/thessoft.htm>> [Consulta: 9 de septiembre de 2013].
- World Wide Web Consortium (W3C). *SKOS Simple Knowledge Organization System Reference. W3C Recommendation 18 August 2009* [en línea]. W3C, 2009. Disponible en: <<http://www.w3.org/TR/2009/REC-skos-reference-20090818/>> [Consulta: 12 de septiembre de 2013].