

# ARTÍCULOS



## TESAUROS: ESTÁNDARES Y RECOMENDACIONES



**Sonia Sánchez-Cuadrado, María-Jesús Colmenero-Ruiz y José-Antonio Moreiro**



**Sonia Sánchez-Cuadrado** es doctora en documentación por la *Universidad Carlos III de Madrid*. Desde 2002 es profesora en el *Departamento de Informática* de dicha *Universidad* con docencia en documentación e informática sobre temas relacionados con el modelado de tesauros y elaboración de sistemas de organización del conocimiento y recuperación de información.

*Universidad Carlos III de Madrid*  
Avda. Universidad, 30. 28911 Leganés, Madrid  
ssanhec@ie.inf.uc3m.es



**María-Jesús Colmenero-Ruiz** es licenciada en ciencias biológicas por la *Universidad de Alcalá de Henares*. Diplomada, licenciada y doctora en documentación por la *Universidad Carlos III de Madrid*. Trabaja como profesora ayudante en el *Departamento de Biblioteconomía y Documentación*, y forma parte del grupo de investigación *Acrópolis*. Su investigación se centra en tesauros y vocabularios controlados, web semántica y *topic maps*.

*Universidad Carlos III de Madrid*  
C/ Madrid, 126. 28903 Getafe, Madrid  
mcolmene@bib.uc3m.es



**José-Antonio Moreiro** es catedrático de biblioteconomía y documentación de la *Universidad Carlos III de Madrid*. Es autor de trabajos sobre técnicas de análisis de contenido documental y cuestiones conceptuales de la documentación. Ha sido director del *Departamento de Biblioteconomía y Documentación* y decano de la *Facultad de Humanidades, Comunicación y Documentación* de la *Universidad Carlos III de Madrid*.

*Universidad Carlos III de Madrid*  
C/ Madrid, 126. 28903 Getafe, Madrid  
jamore@bib.uc3m.es

### Resumen

Los tesauros se han adaptado al entorno tecnológico y constituyen una pieza esencial como medio de organización del conocimiento para la web semántica. Tras un largo período sin abordar la actualización de estándares y recomendaciones de los tesauros, en los últimos cinco años se han realizado revisiones en el seno de los organismos normalizadores. Un trabajo de consenso internacional ha culminado con la publicación de la primera parte de la norma *ISO* para tesauros. Aporta un modelo de datos y pone el acento en la interoperabilidad. Se recoge lo más destacado de las diferentes directrices para tesauros que se han sucedido desde los años ochenta hasta el estándar *ISO 25964*. Se reflexiona sobre cuáles han sido las características más destacadas de esta evolución y se exploran los usos más novedosos de los sistemas de organización del conocimiento en la Web.

### Palabras clave

Tesauros, Estándares, Directrices, Recomendaciones, Pautas, Vocabularios controlados, Sistemas de organización del conocimiento.

**Title: Thesauri: Standards and guidelines**

### Abstract

Thesauri and other knowledge organization systems have been adapted to the technological environment. Currently, this represents an essential aspect of knowledge organization on the semantic web. After a long period in which thesaurus guidelines and standards were not updated, some have been revised by the governing organisations within the past five years. Consensus among international experts has resulted in publication of the first part of an *ISO* standard for thesauri, which provides a data model and emphasizes the interoperability process. This paper compiles the most relevant guidelines for thesauri published between the '80s and the *ISO 25964* document. We reflect on the most relevant characteristics in this evolution and explore the most recent applications of knowledge organization systems on the Web.

Artículo recibido el 10-02-12  
Aceptación definitiva: 16-03-12

## Keywords

Thesauri, Standards, Guidelines, Recommendations, Controlled vocabulary, Knowledge organization systems.

**Sánchez-Cuadrado, Sonia; Colmenero-Ruiz, María-Jesús; Moreiro, José-Antonio.** "Tesauros: estándares y recomendaciones". *El profesional de la información*, 2012, mayo-junio, v. 21, n. 3, pp. 229-235.  
<http://dx.doi.org/10.3145/epi.2012.may.02>

## Introducción

En este estudio se presenta una revisión de las recomendaciones y estándares sobre sistemas de organización de conocimiento, también conocidos como *KOS* por la denominación inglesa *knowledge organization systems*, con atención preferente a los tesauros. Los *KOS* son esquemas para organizar información y gestionar conocimiento. Dentro del concepto se incluyen diversos tipos de listas de términos (diccionarios, glosarios, listas de autoridades y de nombres geográficos o *gazetters*), taxonomías, encabezamientos de materia, tesauros, redes semánticas y ontologías (Hodge, 2000; Zeng; Chan, 2004). La diferencia entre unos y otros reside en el tipo de estructuración y la información que pueden representar (Sánchez-Cuadrado *et al.*, 2007). Las tecnologías de la información, y en concreto internet, han motivado la aparición de diferentes iniciativas para adecuarlas a este nuevo medio digital (Pérez-Agüera, 2004). Inicialmente, las tendencias se orientaban a representar *KOS* sólo para indexar y recuperar información y después han evolucionado hacia la idea de la web semántica para compartir conocimiento.

## Normas y recomendaciones: evolución

La evolución de las recomendaciones de los tesauros y los sistemas de organización del conocimiento en general gira en torno a algunas nociones fundamentales como: vocabulario controlado, término, concepto, descriptores, relaciones, interoperabilidad, *mapping*, etc. Cuando en las normas se habla de "términos" se refiere a palabras o frases cortas que representan un concepto unívoco, de tal manera que si se trata de una palabra homónima, entonces se deben utilizar mecanismos de desambiguación, es decir, que precisen el significado del término. Un "descriptor" designa un término aprobado para ser usado en indización. Asumiendo fundamentos comunes con la terminología, los conceptos de un tesauro son términos especializados de un campo determinado en los que se especifican las relaciones conceptuales y se asegura que un término representa inequívocamente un concepto (Amtz; Picht, 1995). Los conceptos se precisan como una unidad de pensamiento que puede ser definida o descrita. Sin duda la idea más novedosa desde que surgieron los tesauros es la de interoperabili-

dad tomada del campo de las TIC: Se denomina interoperabilidad a la propiedad de un sistema cuyas interfaces pueden trabajar con otros sistemas sin restricciones de acceso o implementación.

Aunque de forma reiterativa se plantea la cuestión de si los tesauros siguen teniendo vigencia (Rowley, 1994; Sveinonius, 1986; Sánchez-Cuadrado *et al.*, 2007), lo cierto es que continúan apareciendo ejemplos e iniciativas a favor de los vocabularios controlados. Recientemente *Google*, *Microsoft* y *Yahoo!* se unieron para ofrecer una colección de vocabularios compartidos: *Schema.org*. Los webmasters pueden usar el vocabulario definido en este esquema para definir y marcar sus páginas. El principal beneficio para los tres buscadores es compartir los mismos mecanismos de entendimiento.

<http://schema.org>

Lancaster (2002) planteó la evolución de los vocabularios controlados desde sus inicios hasta mediados de los ochenta, mostrando la influencia de unas recomendaciones en otras. Dextre-Clarke (2008) revisó los últimos 50 años de la organización del conocimiento desde su perspectiva personal. La llegada de los ordenadores personales, y más tarde la necesidad de interconectar diferentes vocabularios controlados marcó un punto de inflexión importante. Otros estudios comparan enfoques entre diferentes propuestas como el estándar *ISO 2788* y la recomendación *SKOS (W3C, 2008)*. Se puede consultar una visión de algunas normas y los principales protocolos o estándar de intercambio de datos en Dextre-Clarke (2010).

La figura 1 muestra la evolución de las normas y recomendaciones sobre tesauros. Se ha generalizado a las normas o recomendaciones de los sistemas de organización del cono-

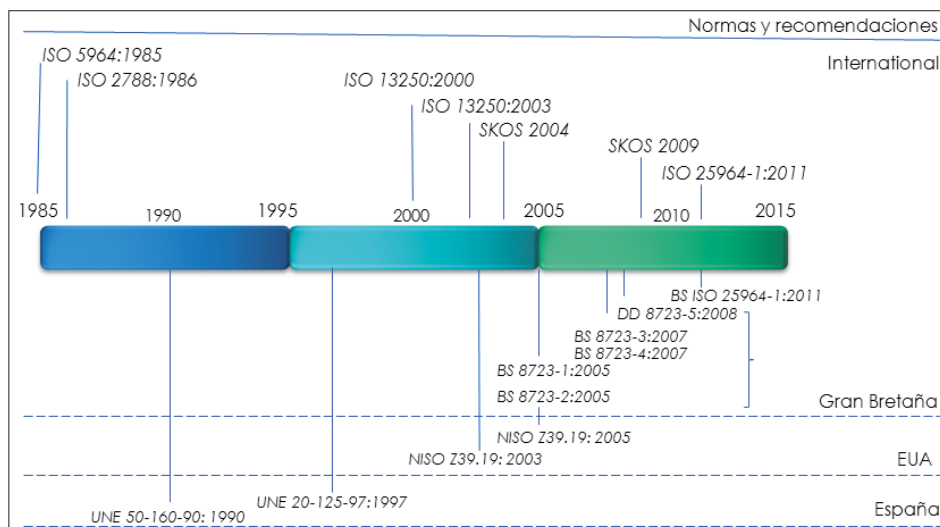


Figura 1. Evolución de normas y recomendaciones sobre tesauros y sistemas de organización de conocimiento

cimiento de instituciones de estándares de algunos países como Gran Bretaña o Estados Unidos por la innegable influencia y peso que han tenido en la última norma internacional publicada. El punto de inicio se ha marcado con la primera norma, que hasta la aparición de esta última estaba vigente y como punto de continuación para el esquema desarrollado por **Lancaster** (2002). En la parte superior se presentan las normas aparecidas a nivel internacional, mientras que en la parte inferior se han recogido algunas normas representativas aparecidas en los últimos años a nivel local.

#### **ISO 5964:1985**

Denominada *Guidelines for the establishment and development of multilingual thesauri*, la norma *ISO 5964* de 1985 insta los principales procedimientos para el tratamiento de varias lenguas en un mismo recurso de conocimiento, identificando los principales problemas en cuanto a la equivalencia semántica. Esta norma fue adoptada por el equivalente estándar nacional en diferentes países. En España se tradujo en la *UNE 50-160-90:1990*.

Sus directrices definen una lengua fuente y una serie de lenguas objetivo. Los conceptos de la lengua fuente se traducen a la lengua objetivo mediante reglas que dependen de la existencia de equivalencias exactas, inexactas, parciales, compuestas o no equivalencia.

#### **ISO 2788:1986**

La norma *ISO 2788* para tesauros monolingües (*Guidelines for the establishment and development of monolingual thesauri*), inicialmente publicada en 1974, fue revisada y publicada en 1986. Esta norma a nivel internacional se ha mantenido durante 25 años, hasta la aparición de la *ISO 25964-1* en 2011. Cada organización nacional publicó su versión traducida y adaptada a cada país. Estas normas locales han tenido posteriores revisiones al año 1986 como la británica y la americana.

#### **ISO 13250:2000/2003**

El estándar *ISO 13250* define los *topic maps (TM)*. *Topic maps* es un metamodelo para categorizar, fusionar y navegar mediante relaciones semánticas entre documentos y secciones de los mismos a partir de su contenido. Están especificados en lenguaje xml, con una sintaxis que se denomina xtm (**Pepper; Moore**, 2001). En el estándar xtm se declaran los elementos y los conceptos claves (*topic, association, occurrence, published subject* y *scope*). Un *TM* tiene como objetivo la indización de contenido en conjuntos de documentos, y permite realizar indizaciones múltiples para la misma colección de documentos web. El estándar *TM* también se suele incluir como un formato de intercambio electrónico de tesauros, aunque no está específicamente definido para ellos, siendo sólo una de sus aplicaciones.

#### **SKOS 2004/2009**

*Simple knowledge organization systems (SKOS)* es una recomendación del *W3C* que tiene su origen en la iniciativa *SWAD-Europe*. Propone un modelo de datos para definir e implementar estructuras básicas de contenido como organización del conocimiento semi-formal que permita un entorno compartido basado en un modelo simplificado y que haga posible su reutilización e interoperabilidad. La primera

publicación tuvo lugar en 2004 siendo actualizada y publicada en 2009. El lenguaje de formalización es *rdf*. Su potencia reside en que puede ser combinado y extendido con lenguajes formales de representación del conocimiento como *owl* y que admite la incorporación de otros vocabularios de metadatos (*Dublin core, FOAF*). Se dirige a conceptos que pueden ser representados por un *uri (universal resource identifier)*, permitiendo representar *KOS* multilingües mediante la combinación de propiedades. La especialización del modelo *SKOS* con la extensión *XL (SKOS-XL)* permite el uso de propiedades de *rdf (rdfs:subPropertyOf* y *rdfs:subClassOf)* para crear subtipos de relaciones interpretables por un motor de inferencia adecuado para la web semántica. *SKOS* incluye la especificación básica del modelo de representación (*SKOS core*), recursos para el alineamiento e interoperabilidad de los conceptos (*SKOS mapping*), mecanismos de ampliación de propiedades de representación de conocimiento (*SKOS extensions*) y servicios para la consulta web de *KOS (SKOS API)*.

En la literatura se refleja una amplia aceptación para formalizar los vocabularios controlados con *SKOS (Agrovoc, Eurovoc, LCSH)*. Como modelo ligero de datos, *SKOS* ha demostrado su utilidad para los procesos de interoperabilidad a nivel terminológico (**Sah et al.**, 2007), y para el conocimiento colaborativo en la generación social de esquemas de clasificación (**Sah; Wade**, 2010).

#### **ISO 25964-1:2011**

La norma *Thesauri and interoperability with other vocabularies* está siendo actualizada. La primera parte denominada *Thesauri for information retrieval (ISO 25964-1:2011)* fue publicada en 2011. La segunda parte está prevista para 2012 y está orientada a guiar la interoperabilidad con otros vocabularios. Se basa en el estándar *BS8723 Structured vocabularies for information retrieval* que a su vez se inspiró en la propuesta *ANSI*. La nueva actualización de la norma *ISO 25964-1:2011* a diferencia de la británica está constituida en Ods partes.

SKOS propone un modelo de datos para definir e implementar estructuras básicas de contenido que permita un entorno compartido y que haga posible su reutilización e interoperabilidad

Es probablemente la renovación más profunda que ha experimentado el estándar *ISO* de los tesauros, ya demandada por los profesionales desde tiempo atrás. Fusiona los estándares anteriores para tesauros monolingües y multilingües. También propone un formato de intercambio basado en un esquema xml. Su objetivo es proporcionar recomendaciones para el desarrollo y mantenimiento de los tesauros. Se observa un aumento en el uso de facetas para la organización del esquema de conocimiento. Está dirigida a diferentes tipos de recursos de información, con independencia del medio utilizado (texto, multimedia, bases de conocimiento, etc.), que permitan su uso en aplicaciones de recuperación

de información. En esta versión *ISO* se incorporan técnicas de procesamiento textual que flexibilicen el tratamiento del vocabulario controlado y las técnicas de recuperación de información que ya estaban siendo aplicadas en software de tesauros y motores de búsqueda (Marrero *et al.*, 2010). Por ejemplo se tiene en cuenta la casuística de combinaciones de palabras posibles y en formatos diferentes teniendo en cuenta que una computadora debe recibir instrucciones para identificar variantes ortográficas de un mismo término.

### **ANSI/NISO Z39.19:2005**

El *American National Standards Institute* realizó una profunda revisión del estándar de los tesauros bajo la denominación *Guidelines for the construction, format and management of monolingual controlled vocabularies*. Proponía una guía para la construcción de tesauros, incidiendo en los aspectos de formato y gestión de vocabularios monolingües. Las recomendaciones de la norma *NISO Z39.19:2005* incluían criterios para el mantenimiento de los sistemas de organización del conocimiento mediante gestores de tesauros automatizados. Esta norma estaba orientada a cubrir la representación de contenido de diferentes *KOS* como las listas de sinónimos, taxonomías, tesauros y otros tipos de vocabularios controlados.

La propuesta de esta *NISO (National Information Standards Organization)* entiende que los vocabularios controlados suelen ser utilizados para describir el contenido por asignación de términos para representar metadatos asociados al contenido de los objetos (*NISO Z39.19:2005:12*). El estándar *NISO Z39.19:2005* está pensado para ser aplicado a tesauros monolingües (*NISO Z39.19:2005*, apartado 2.4). Es decir, a diferencia de los estándares *ISO* aportados hasta ese momento, esta iniciativa incorporaba ya su aplicación en entornos informatizados, como demuestra el tratamiento de esquemas de metadatos como *Dublin Core*.

### **IFLA 2009**

La *IFLA* publicó su informe sobre las directrices para tesauros multilingües en 2009 con el objetivo de completar la norma *ISO 5964:1985*. Complementaba también a otras directrices dirigidas a los encabezamientos de materias y al estándar *NISO Z39.19:2005* (*IFLA*, 2009). La principal aportación de este informe es sobre el tratamiento de construir tesauros de forma no-simétrica y la vinculación entre diferentes vocabularios controlados. Por tesauros no-simétricos entienden un tipo de tesauros multilingüe donde el número de descriptores de cada lengua no es necesariamente el mismo y tampoco el modo en que los descriptores se relacionan en las diferentes lenguas.

### **BS8723 (2005-2008) y BS ISO 25964-1:2011**

El estándar *BS8723 (Structured vocabularies for information retrieval)* se publicó durante los años 2005-2008. Está constituido por cinco partes dónde la última, conocida como *DD8723-5*, está dedicada a protocolos y formatos de intercambios para la interoperabilidad. Está dirigido a vocabularios estructurados en general, caracterizando y comparando los diferentes tipos de *KOS* como esquemas de clasificación, taxonomías, esquemas de encabezamiento de materias, te-

sauros, listas de autoridades y ontologías. El software pasa a tener una función esencial en: la gestión del vocabulario, los recursos de formalización necesarios para los mapeos entre vocabularios, y los protocolos y los formatos de intercambios. El enfoque de organización de conocimiento se basa en la representación de conceptos.

Las distintas partes de esta norma británica han quedado anuladas tras la publicación de la *BS ISO 25964:2011* con el título *Information and documentation. Thesauri and interoperability with other vocabularies. Thesauri for information retrieval*, idéntica a la *ISO* del 2011 y que supone la traslación local de la norma internacional, con su plena adhesión.

### **Implicaciones de la evolución de la norma de los tesauros**

Las revisiones más influyentes en la última *ISO* aprobada han sido la americana *NISO Z3919:2005* y la británica *BS5723-1-5:2005-2008*. Un aspecto a destacar es la modificación del nombre de la norma. La *NISO* generalizó la norma para los “vocabularios controlados”. En el caso del estándar británico se establecen diferencias entre tesauros y lo que no son tesauros como vocabularios controlados para la recuperación de información. Sin embargo, la *ISO* siempre ha mantenido en el título y en el objetivo a los tesauros.

“ Una de las diferencias más significativas la encontramos en el paso de estructuración terminológica a una dimensión de representación conceptual dónde las equivalencias se crean entre conceptos y no entre palabras ”

La norma *ISO 25964:2011* surge como revisión de las normas *ISO 2788* y *5964*, que después de 25 años ya no responden a las prácticas actuales de procesos de indización y recuperación de información, ni cubren la demanda de aplicaciones de red que quieren utilizar recursos heterogéneos. Por tanto, por una parte estamos hablando de un cambio por obsolescencia del formato (del papel a soporte electrónico), y por otro de un cambio o incremento de funcionalidad. Pero también hablamos de una adaptación a nuevos perfiles de usuarios. La unión de recursos distintos implica el desarrollo de mecanismos de *mapping* o alineamiento que permitan la interoperabilidad. Una de las diferencias más significativas la encontramos en el paso de estructuración terminológica a una dimensión de representación conceptual dónde las equivalencias se crean entre conceptos y no entre palabras. Si nos remitimos a las definiciones puede no tener sentido si consideramos que los términos representan conceptos. Los conceptos a su vez se definen con palabras y un término es una palabra de un lenguaje de especialidad. La explicación reside en si la finalidad de definir y representar se hace al modo de las bases de datos léxicas o las ontologías lingüísticas como *WordNet*, o bien desde un punto de vista formal más próximo al modelo matemático. Este hecho repercutirá en la definición de restricciones en los tipos de relaciones

como ocurre en otros esquemas de organización de conocimiento como las ontologías.

Los modelos de datos del estándar *ISO 25964-1:2011* y la recomendación de *SKOS* (2009) del *W3C* están basados en la representación de conceptos. A nivel conceptual debería ser fácil la conversión entre ellos, puesto que tratan de representar lo mismo. Ambos grupos han mostrado interés en sus trabajos recíprocos por la sensibilización existente en la profesión frente a las barreras que puedan crearse con el uso de distintos estándares, en lugar de producirse una suma de esfuerzos. Los trabajos del estándar *ISO* justifican su iniciativa con la inclusión de muchos países y la consideración de iniciativas paralelas. Según la *ISO* los formatos para la interoperabilidad existentes (*MARC*, *SKOS*, *Zthes*, esquema XML de la *DD8723-5* británica y la *ISO 16642*, el *Terminological markup framework (TMF)*) no son adecuados para expresar todos los aspectos de un tesauro según se definen en el estándar (*ISO 25964-1:2011*). **Dextre-Clarke** (2010) asegura que la comunidad *SKOS* está considerando hacer extensiones a la propia recomendación para mejorar el alineamiento con el estándar *ISO*.

**Dextre-Clarke** (2009) señala diferencias entre los dos modelos de intercambio. En primer lugar, *SKOS* no establece directrices para la construcción de un sistema de organización de conocimiento, solamente proporciona un lenguaje de formalización para su publicación en la Web. En segundo lugar, defiende que el modelo de datos de la norma *ISO* incluye todas las características de un tesauro, mientras que *SKOS* pretende servir a varios tipos de vocabularios diferentes, por lo que no prevé mecanismos para algunas de las características del tesauro. Sin embargo, una de las limitaciones que presenta la norma *ISO 25964-1:2011* es que no es aplicable a bases de datos o software que pueda ser utilizado directamente en las aplicaciones de búsqueda o indexación, limitándose a sugerir recomendaciones para la gestión de diccionarios de sinónimos.

La nueva norma de los tesauros incide en la importancia del *mapping* entre recursos asumiendo que es un proceso complejo. Una de las dificultades del proceso reside en la rareza de encontrar auténtica equivalencia exacta. A esto debemos sumar que muchos de los vocabularios están pobremente contruidos con objetivos distintos, con diferentes niveles de granularidad en la jerarquía, etc. Sin duda estas peculiaridades dificultan el mapeo e intensifican su labor, viéndose agudizado por los escasos recursos que se dedican normalmente. El metatesauro de medicina *UMLS (Unified medical language system)*, de la *US National Library of Medicine*, que integra 100 tesauros de medicina y cinco millones de conceptos, es uno de esos ejemplos donde se ha invertido gran esfuerzo para mapear estructuras de conocimiento diferentes en el entorno digital.

<http://www.nlm.nih.gov/research/umls>

## Sistemas de organización de conocimiento en la Web

El *mapping* persigue entendimiento común, y cuando hablamos de intercambio de datos se convierte en interoperabilidad. El *mapping* de recursos diversos implica enlazar

datos o información. Una forma de crear tesauros también es posible mediante el enlace de conceptos de diferentes vocabularios controlados. Este método se caracterizaría por una gestión colaborativa del tesauro y la posibilidad de crear de forma flexible tesauros descentralizados. Esto significa que el sistema de organización del conocimiento puede ser creado y mantenido de forma externa, pero conectar con conceptos equivalentes. Esta idea está relacionada con los principios de *linked data*.

El grupo de *linked data W3C Library Linked Data Incubator Group* trabaja enlazando datos en rdf. Utiliza URIs para identificar los conceptos y ofrece información sobre los recursos usando rdf. Esto le permite incluir enlaces a otras URIs. De este modo consigue interconectar datos diferentes y a su vez reutilizar información en lugar de duplicarla.

Los aspectos básicos del diseño son los siguientes:

- que el dato esté en la Web;
- que sea un dato interpretable por las máquinas;
- que no esté en formato propietario;
- que se use el estándar rdf;
- que sean enlazados mediante rdf.

Estos principios apuntan directamente a la reutilización de conocimiento. El proyecto *DBpedia* es un ejemplo de este planteamiento. Trabaja para extraer datos de la *Wikipedia* y ponerlos a disposición web de forma estructurada, mostrando los resultados de búsqueda con *Zitgist*. Otros ejemplos de resultados de *linked data* significativos son *Geonames* y *The Data Hub*.

<http://zitgist.com>

El modo de unir diferentes vocabularios puede hacerse con mapeos uno a uno de los esquemas, pero también mediante una ontología de alto nivel o una ontología de referencia. Un ejemplo es el proyecto *Umbel (Upper mapping and binding exchange layer)* diseñado para ayudar a los contenidos a interactuar. Se establecen vínculos entre la ontología conceptual de referencia y los vocabularios controlados o de dominio y conjuntos de datos. Incluye vocabularios de metadatos y ontologías de amplia difusión como *SKOS*, *WordNet*, *RSS*, *FOAF*, *Sioc*, *PIM*, *GeoNames*, y *OpenCyc*, entre otras.

<http://rdfs.org/sioc/spec>

Ontologías y vocabularios de metadatos se vinculan mediante diferentes etiquetas para unir conceptos idénticos, ayudando así a la fusión y coexistencia de diferentes conjuntos de etiquetas. Es decir, el objetivo no es transformar los recursos, sino crear un mapa de su significado como en *TagCommons* o *Folcsonomías*. Otro proyecto relacionado es *Semse*, un sistema que permite la representación y recuperación conceptual de esquemas de metadatos mediante su mapeo contra una ontología de alto nivel (**Palcios**, 2010).

Los tesauros también han experimentado un fuerte impulso con el desarrollo de las ontologías. La ingeniería ontológica y las iniciativas para elaborar la web semántica se han centrado en los mismos fundamentos de estructuración y organización del conocimiento, pero con la aplicación de lenguajes de formalización como rdf y owl (**Daconta et al.**, 2003, p. 157).

2005 fue un año clave para un planteamiento interoperable. La *OMG (Organization Management Group)* publicó *Ontology definition metamodel*, que integra seis metamodelos entre los que se encuentra *UML*. Los tesauros, entendidos como un tipo de ontología ligera (**Daconta et al.**, 2003) o sistemas de representación del conocimiento (**Sánchez-Cuadrado et al.**, 2007), pueden ser definidos según el metamodelo de la *OMG*. *UML* no se creó pensando expresamente en los tesauros, pero pueden ser modelados (**Génova et al.**, 2011) como podemos observar en los modelos de tesauros publicados por las normas de 2011.

## Conclusiones

El análisis de las normas de tesauros y *KOS* en general permite determinar que son un recurso caracterizado por su contenido, su estructura y su función. En su proceso evolutivo estos aspectos han cambiado. La nueva concepción de sistema de organización de conocimiento deja a un lado las diferenciaciones entre mono y multilingüismo. Requieren la representación formal del contenido, lo que implica un lenguaje de formalización más lógico. Este cambio de representación condiciona sin duda la estructura relacional. La función de los *KOS* se ha ampliado en la última década con las TIC, convirtiéndose la interoperabilidad, la vinculación entre *KOS* y la reutilización en factores esenciales.

Las normas y recomendaciones de la *ISO* se han centrado más en tesauros que en otros tipos de *KOS*, mientras que *SKOS*, planteando una formalización simple ha facilitado el uso y la eficiencia de los vocabularios controlados en la web semántica. Desde la eclosión de la Web, los tesauros han generado un vínculo inseparable con la informática, aunque esto no ha sido recogido por las recomendaciones hasta 2004 por el *W3C* y hasta 2005 por los organismos de estandarización. Este reconocimiento surgió de su necesidad de adaptación a los entornos tecnológicos actuales y pone de manifiesto la importancia y vigencia de los tesauros, así como otros *KOS* para la gestión del conocimiento. Su adaptación técnica supone un gran avance para estos sistemas de organización del conocimiento de gran tradición en nuestro campo.

## Bibliografía

Aenor. *UNE 50106:1990. Directrices para el establecimiento y desarrollo de tesauros monolingües*. Madrid: Aenor, 1990.

Aenor. *UNE 50125:1997. Directrices para el establecimiento y desarrollo de tesauros multilingües*. Madrid: Aenor, 1997.

**Amtz, Reiner; Picht, Heribert**. *Introducción a la terminología*. Madrid: *Fundación Germán Sánchez Ruipérez*, 1995. ISBN: 84 8616 895 3

ANSI/NISO. *Z39.19: 2005 Guidelines for the construction, format and management of monolingual controlled vocabularies*. NISO Press: Bethesda, MD, 2005. ISBN: 1 880124 65 3

[http://www.niso.org/apps/group\\_public/project/details.php?project\\_id=46](http://www.niso.org/apps/group_public/project/details.php?project_id=46)

ISO. *BS 8723-1:2005 Structured vocabularies for information retrieval. Guide. Definitions, symbols and abbrevia-*

*tions*. British Standards Institution: London, 2005. ISBN: 0 580 46798 8

ISO. *BS 8723-2:2005 Structured vocabularies for information retrieval. Guide. Thesauri*. British Standards Institution: London, 2005, 60. ISBN: 0 580 46799 6

ISO. *BS 8723-3:2007 Structured vocabularies for information retrieval. Guide. Vocabularies other than thesauri*. British Standards Institution: London, 2007, 52. ISBN: 978 0 580 63072 9

ISO. *BS 8723-4:2007 Structured Vocabularies for information retrieval. Guide. Interoperability between vocabularies*. British Standards Institution: London, 2007, 62. ISBN: 978 0 580 63073 6

**Daconta, Michael C.; Obrst, Leo J.; Smith, Kevin T.** *The semantic web. A guide to the future of XML, web services, and knowledge management*. Indianapolis: Wiley, 2003. ISBN: 978 0471432579

DD 8723-5:2008 *Structured vocabularies for information retrieval. Guide. Formats and protocols for data exchange*. British Standards Institution: London, 2008, 60. ISBN: 978 0 580 53864 3

**Dextre-Clarke, Stella G.** "International standard ISO 25964". En: *The 9<sup>th</sup> European networked knowledge organization systems (NKOS). Workshop at the 13 ECDL Conf, Corfu, Greece, Oct. 2009*.

**Dextre-Clarke, Stella G.** "Knowledge organization systems standards". *Encyclopedia of library and information sciences*, Third Edition. Edition by Taylor & Francis, 2010, pp. 3164-3175. ISBN: 0 8493 9711 1

**Dextre-Clarke, Stella G.** "The last 50 years of knowledge organization: a journey through my personal archives". *Journal of information science*, 2008, v. 34, n. 4, pp. 427-437. <http://dx.doi.org/10.1177/0165551508089225>

**Dextre-Clarke, Stella G.** "Thesaurus standards on a converging track". *Legal information management*, 2010, v. 10, n. 1, pp. 43-45. <http://dx.doi.org/10.1017/S1472669610000265>

**Génova-Fuster, Gonzalo; Llorens-Morillo, Juan; Fuentes-Torres, José-Miguel; Morato-Lara, Jorge; Martínez-Fernández, Paloma.** "Las jerarquías conceptuales en UML comparando la norma ISO 2788 con el metamodelo de UML". *Técnica administrativa*, 2011, v. 10, n. 45. <http://www.cyta.com.ar/ta1001/v10n1a1.htm>

**Hodge, Gail.** *Systems of knowledge organization for digital libraries. Beyond traditional authority files*. Washington, DC: The Council on Library and Information Resources, 2000. <http://www.clir.org/pubs/reports/pub91/contents.html>

IFLA. *Guidelines for multilingual thesauri*. (IFLA professional reports: 115). IFLA, 2009. ISBN: 978 90 77897 35 5 <http://www.ifap.ru/library/book411.pdf>

ISO. *ISO 25964-1:2011. Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval, 2011*.

ISO. *ISO 2788:1986. Guidelines for the establishment and*

*development of monolingual thesauri*. 2<sup>nd</sup> ed. Geneva: International Organization for Standardization, 1986.

ISO. *ISO 5964:1985. Guidelines for the establishment and development of multilingual thesauri*. Geneva: International Organization for Standardization, 1985.

ISO. *ISO/IEC 13250:2003. Information technology. SGML Applications. Topic maps*. Geneva: International Organization for Standardization, 2003.

**Lancaster, Frederick W.** *El control de vocabulario en la recuperación de la información*. Valencia: Universitat de València, 2002. ISBN: 978 84 3705 444 5

**Marrero, Mónica; Sánchez-Cuadrado, Sonia; Urbano, Julián; Morato, Jorge; Moreira, José-Antonio.** "Sistemas de recuperación de información adaptados al dominio biomédico". *El profesional de la información*, 2010, mayo-junio, v. 19, n. 3, pp. 246-254.

<http://dx.doi.org/10.3145/epi.2010.may.04>

**Palacios-Madrid, Vicente.** *Sistema de recuperación conceptual mediante niveles semánticos en la representación de esquemas de metadatos*. Tesis doctoral. Universidad Carlos III de Madrid, 2010.

<http://e-archivo.uc3m.es/handle/10016/9332>

**Pepper, Steve; Moore, Graham.** "XML topic maps (XTM) 1.0. Specification". *TopicMaps.Org*, 2001.

<http://www.topicmaps.org/xtm/1.0>

**Pérez-Agüera, José-Ramón.** "Automatización de tesauros y su utilización en la web semántica". *Textos universitaris de biblioteconomia i documentació*, 2004, dic., n. 13.

<http://www.ub.edu/bid/13perez2.htm>

**Rowley, Jennifer.** "The controlled versus natural languages debate revisited: a perspective on information retrieval practice and research". *Journal of information science*, 1994, April, v. 20, n. 2, pp. 108-118.

<http://dx.doi.org/10.1177/016555159402000204>

**Sah, Melike; Hall, Wendy; Gibbins, Nicholas M.; De Roure, David C.** "Sempport: a personalized semantic portal". En: *Procs of the 18<sup>th</sup> conf on hypertext and hypermedia*. ACM, New York, 2007, pp. 31-32.

<http://eprints.soton.ac.uk/264441>

**Sah, Melike; Wade, Vicent.** "Automatic metadata extraction from multilingual enterprise content". En: *Procs of the 19<sup>th</sup> ACM intl conf on information and knowledge management*. New York: ACM, 2010, p. 1665-1668.

<http://www.cngl.ie/drupal/sites/default/files/papers2/Automatic%20Metadata%20Extraction.pdf>

**Sánchez-Cuadrado, Sonia; Morato-Lara, Jorge; Palacios-Madrid, Vicente; Llorens-Morillo, Juan; Moreira, José-Antonio.** "De repente, ¿todos hablamos de ontologías?". *El profesional de la información*, 2007, nov.-dic., v. 16, n. 6, pp. 562-568.

<http://www.elprofesionaldelainformacion.com/contenidos/2007/noviembre/03.pdf>

<http://dx.doi.org/10.3145/epi.2007.nov.03>

**Svenonius, Elaine.** "Unanswered questions in the design of controlled vocabularies". *Journal of the American Society for Information Science*, 1986, Sept., v. 37, n. 5, pp. 331-340.

[http://polaris.gseis.ucla.edu/gleazer/462\\_readings/Svenonius\\_1986.pdf](http://polaris.gseis.ucla.edu/gleazer/462_readings/Svenonius_1986.pdf)

W3C. *SKOS Simple Knowledge Organization Systems Primer*. W3C Working Group, 2008.

<http://www.w3.org/TR/2008/WD-skos-primer-20080221>

**Zeng, Marcia L.; Chan, Lois M.** "Trends and issues in establishing interoperability among knowledge organization systems". *Journal of the American Society for Information Science and Technology*, 2004, v. 55, n. 5, pp. 377-395.

<http://dx.doi.org/10.1002/asi.10387>

## Próximos temas centrales

Julio 2012	Comunicación digital
Septiembre 2012	Información y derecho
Noviembre 2012	Bibliotecas académicas
Enero 2013	Soportes digitales
Marzo 2013	Educación y biblioteca
Mayo 2013	Bibliotecas y documentación de museos
Julio 2013	Economía de la información
Septiembre 2013	Gestión de contenidos

Los interesados por favor consulten detenidamente las Normas para autores:

<http://www.elprofesionaldelainformacion.com/autores.html>

y luego envíen sus artículos a través del gestor de manuscritos OJS de la plataforma del Recyt:

<http://recyt.fecyt.es/index.php/EPI/index>

Copyright of El Profesional de la Información is the property of EPI SCP and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.