

Conversión normalizada (SKOS) de sistemas de organización del conocimiento interoperables en la web

Gonzalo Mochón Bezares*
Eva María Méndez Rodríguez*
Gema Bueno de la Fuente**

Artículo recibido:
29 de enero de 2019

Artículo aceptado:
9 de abril de 2019

Artículo de investigación

RESUMEN

Este estudio examina de forma exhaustiva la literatura científica dedicada a los procesos de *skosificación* de vocabularios y sistemas de organización del conocimiento. Se analizan en profundidad 49 trabajos que describen y detallan la transformación de un total de 59 vocabularios controlados convencionales o SOC (Sistemas de Organización del Conocimiento) a Simple Knowledge Organization System (SKOS). Se identifican los puntos clave para hacer el análisis de metodologías de transformación de vocabularios en SKOS para la web y se comparan los estudios para determinar las aproximaciones y parámetros más recomendables para llevar a cabo estos procesos de conversión de vocabularios, cada vez más frecuentes y necesarios

* Departamento Biblioteconomía y Documentación, Universidad Carlos III de Madrid, España
gmochonb@gmail.com emendez@bib.uc3m.es

** Departamento de Ciencias de la Documentación e Historia de la Ciencia, Universidad de Zaragoza, España
gbueno@unizar.es

en la web semántica y en entornos de *linked data* (LD). Los resultados señalan que la mayor parte de SOC transformados son tesauros, que los formatos mayoritarios son de texto o registros bibliográficos, que el objetivo más común al cambiar a SKOS es la mejora de la interoperabilidad de los vocabularios, y que los procesos de conversión pueden agruparse mediante tres formas: *scripts* realizados en distintos lenguajes, transformaciones XSL y lenguajes de mapeo. Se concluye que SKOS es considerado por los autores como una buena opción para mejorar la interoperabilidad de vocabularios controlados.

Palabras clave: Conversión de Vocabularios; Revisión Bibliográfica; Sistemas de Organización del Conocimiento; SKOS

Normalized conversion (SKOS) of interoperable knowledge organization systems on the Web

Gonzalo Mochón Bezares, Eva María Méndez Rodríguez and Gema Bueno de la Fuente

ABSTRACT

This paper examines exhaustively the scientific literature about skosification processes of vocabularies and knowledge organization systems (KOS). Forty nine documents that describe and detail transformation of 59 conventional controlled vocabularies or KOS to Simple Knowledge Organization System (SKOS) were analyzed. The key points for the analysis of vocabulary transformation methodologies in SKOS for the Web were identified, and studies were compared to determine the most recommendable approximations and parameters, for conversion processes of vocabularies were each time more frequent and necessary in semantic web and linked data (LD) environments. Results indicated that the majority of transformed SOC's are thesauri, that most formats were text or bibliographic records, that the most common objective in the change to SKOS was improvement of interoperability of the vocabularies and that the conversion processes could be grouped by three forms: *scripts* made in different languages; XSL transformations, and mapping languages. It was concluded that SKOS is considered by the authors as a good option for improving the interoperability of controlled vocabularies.

Keywords: Vocabulary conversion, Bibliographic Review, Knowledge Organization Systems, SKOS

INTRODUCCIÓN

Los Sistemas de Organización del Conocimiento (SOC), entendidos de un modo genérico como aquellas clases de esquemas diseñados con el fin de organizar la información y la gestión del conocimiento, y también como las herramientas que puedan ofrecer términos y conceptos en la labor de recuperación de la información (Hodge, 2000: 3), han sido empleados durante décadas en las tareas llevadas a cabo por el personal y los usuarios de los distintos tipos de unidades de información. El diseño de SOC ha sido y es objeto de estudio dentro de la bibliotecología y las ciencias de la información.

Las categorías de SOC incluyen desde vocabularios en los que apenas se establece una estructuración entre los términos, como ocurre con los glosarios o las terminologías, hasta otros que disponen de una elevada organización jerárquica y una elevada riqueza de relaciones entre sus conceptos, como se puede ver en los tesauros y las ontologías. La evolución que se ha dado en los últimos 15 años en los distintos tipos de sistemas de organización del conocimiento ha sido enorme. Esta necesaria transformación ha satisfecho las nuevas demandas sobre gestión y recuperación que se han planteado en un entorno informativo como la World Wide Web, que cada vez resulta más semántico y complejo.

Una ventaja de la adaptación de los SOC al entorno de la web es la mejora que ello supone para el intercambio de datos entre distintos sistemas de información mediante protocolos de comunicación. El aumento en la capacidad para combinar los datos propios con los existentes en otros recursos, conocido como interoperabilidad semántica, ha resultado un avance considerable para muchos tipos de SOC, especialmente tesauros y ontologías. La conversión de SOC a la web no debe limitarse a un mero traslado del contenido a un formato electrónico que sólo pueda ser interpretado por personas, sino que también debe hacerse comprensible para las computadoras, con el fin de que éstas puedan manejar mejor los recursos. Con la adaptación de los vocabularios controlados a los estándares de la web semántica no sólo se ponen éstos en la web, sino que también se convierten en herramientas útiles para la web (Méndez y Greenberg, 2012: 240).

La expresión de SOC ya existentes en la web semántica puede realizarse por diferentes vías, que varían en función del formato empleado para estructurar los contenidos. En uno de los tipos de SOC más empleados, los tesauros, la mejor opción de adaptación a la web semántica es el formato Simple Knowledge Organization System (SKOS). Una ventaja de dicho formato frente a otras opciones en XML, como Zthes, es el empleo del modelo de intercambio de datos RDF. Si se consideran los modelos de vocabularios en RDF,

SKOS destaca sobre los demás por ser un estándar reconocido por el Consorcio W3C y porque su implementación resulta menos compleja y costosa que el formato OWL, más propio de las ontologías (Pastor-Sánchez, Martínez Méndez y Rodríguez-Muñoz, 2009).

Además de ser el más apropiado para la representación de contenidos de tesauros, el modelo de datos SKOS puede representar a otros tipos de SOC semiformales como taxonomías, sistemas de clasificación o listas de palabras clave (World Wide Web Consortium, 2009a y 2009b). El formato SKOS supone un primer paso para lograr la consecución de la interoperabilidad semántica entre los vocabularios, y se ha ido convirtiendo con el tiempo en un estándar *de facto* para la publicación de tesauros en la web semántica (Saorín Pérez y Pastor Sánchez, 2012: 234; Pastor Sánchez, 2013: 189). Este modelo permite la representación de las estructuras de muchos tipos de SOC y ofrece la posibilidad de mapeo de los conceptos hacia otros vocabularios, por lo que ha sido muy utilizado como formato para la representación de vocabularios en la web. Como ejemplos de la expansión del uso del formato SKOS entre los SOC podemos ver el catálogo Basel Register of Thesauri, Ontologies & Classifications (BARTOC, <https://bartoc.org/>) y en la versión antigua de la plataforma Datahub (<https://old.datahub.io/>), que en febrero de 2019 recogían un total de 410 vocabularios y 233 conjuntos de datos disponibles en SKOS, respectivamente.

La adaptación de SOC ya existentes, conocida también por los autores como reingeniería de vocabularios, a un formato adaptado a la web semántica es un proceso necesario para evitar la pérdida de los vocabularios controlados que se han realizado sobre muchas áreas del conocimiento. El cambio de formato a SKOS en los SOC debe ser contemplado como parte de un proceso evolutivo de adaptación de sus estructuras y concepciones semánticas al escenario de una web que se basa en los principios de datos enlazados (*Linked Data*).

Entre los estudios sobre reingeniería de vocabularios destacan los trabajos de Leroi y Holland (2010) y Villazón Terrazas (2011). El primero es un informe del proyecto *Athena* sobre la construcción de un tesauro extraído de diversas fuentes terminológicas con el fin de asegurar la interoperabilidad semántica dentro de la red Europea. En dicho informe se recogen los pasos necesarios para transformación de vocabularios y se ofrecen pautas para la *skosificación*, entendida como el proceso de conversión o transformación de un vocabulario controlado al formato SKOS.

Dentro de la documentación del proyecto *Athena* se evalúa si SKOS es el mejor formato para representar el contenido de la terminología a representar. En dicha revisión se destacan los siguientes puntos para evaluar las terminologías (Leroi y Holland, 2010: 39-45):

- evaluación de la representación de conceptos, sus relaciones semánticas y grado de interoperabilidad;
- identificación de conceptos mediante URI usando un sistema de identificación permanente como Permanent URL (PURL);
- definición de etiquetas únicas para la expresión de conceptos;
- evitar duplicación de información empleando propiedades inversas y simétricas;
- detallar la precisión en las relaciones semánticas entre conceptos, y
- asegurar la documentación de los conceptos y la terminología.

En el trabajo de Villazón Terrazas (2011) se tratan los procesos de reingeniería de recursos no ontológicos a ontologías, ordenándolos por los recursos sobre los que se realiza la transformación (esquemas de clasificación, folksonomías, tesauros y léxicos) y por los métodos utilizados en la implementación (bases de datos, XML, archivos de texto plano y hojas de cálculo). Este autor señala que los procesos de reingeniería deben tratar todos los tipos de recursos no ontológicos, tomando en consideración toda la información sobre los mismos y teniendo en cuenta sus datos internos (Villazón Terrazas, 2011: 39).

ALCANCE Y OBJETIVOS

A pesar de la importancia de las técnicas de *skosificación* de vocabularios y que cada vez es más frecuente convertir o transformar en SKOS vocabularios, hay pocos estudios que permitan comprobar las metodologías y las formas de adaptación de los SOC al formato SKOS, especialmente en escenarios que demandan una mayor interoperabilidad entre los sistemas de información.

Así pues, el objetivo general que se plantea es realizar una revisión exhaustiva de trabajos que recojan el proceso de transformación de uno o más vocabularios al formato SKOS (*skosificación*). Se busca comprobar cuáles son los métodos empleados para la conversión de los elementos de vocabularios controlados al formato SKOS, al tiempo que se analizan las circunstancias y razones expuestas por los autores para decidirse a *skosificar* sus vocabularios y adoptar su codificación a la web semántica.

Además, se identifican objetivos específicos que han servido para guiar los criterios en la evaluación de los procesos de *skosificación* de SOC. Dichos objetivos específicos son los siguientes:

- Analizar técnicas y métodos más frecuentes al *skosificar* vocabularios y SOC.
- Comparar la expresión del mapeo entre los elementos de los SOC y sus equivalentes en el formato SKOS.
- Comprobar la incidencia en la conversión según cada tipo de SOC.

METODOLOGÍA

Para abordar los objetivos y analizar la bibliografía sobre procesos de transformación o *skosificación* de vocabularios, se ha establecido una metodología que permitirá reflejar un estado del arte autorizado y clasificado, atendiendo a las dos fases principales que se detallan a continuación.

Fase 1. Selección, búsqueda e identificación exhaustiva de recursos para el análisis

En primer lugar, se plantea el alcance de la búsqueda, centrándose en aquellos trabajos que detallan la conversión al formato SKOS de cualquier tipo de SOC, bien sea relatando el proceso completo o una parte del mismo. Después se seleccionan nueve bases de datos (cuatro generales y cinco especializadas) en las cuales realizar una serie de consultas sistemáticas.¹ De dichas bases de datos solamente una está en español: *ÍNDICES* (fusión de las antiguas bases de datos ICYT, ISOC e IME).

A continuación se detalla la estrategia de búsqueda a utilizar en las bases de datos seleccionadas mediante las palabras clave SKOS, o bien el desarrollo de dichas siglas (*Simple Knowledge Organization System*), en unión con la denominación de los principales tipos de SOC, tanto en español (*vocabularios, taxonomías, ontologías, tesauros, encabezamientos de materias, esquemas de clasificación, listas de autoridades o listas de palabras clave*) como sus equivalentes en inglés (*vocabularies, taxonomies, ontologies, thesauri, subject headings, classification schemes, authority files o keyword lists*), dependiendo de la lengua de interrogación que se debe emplear en cada base de datos (*Tabla 1*).

1 *Academic Search Ultimate, ACM Digital Library, ÍNDICES, Library and Information Science Abstracts, Library, Information Science & Technology Abstracts, ProQuest Library Science Database, Proquest Research Library, Scopus y Web of Science.*

SKOS OR "Simple Knowledge Organization System" AND vocabularies
SKOS OR "Simple Knowledge Organization System" AND taxonomies
SKOS OR "Simple Knowledge Organization System" AND ontologies
SKOS OR "Simple Knowledge Organization System" AND thesauri
SKOS OR "Simple Knowledge Organization System" AND subject headings

Tabla 1. Ejemplos de sentencias de búsqueda utilizadas en las BBDD seleccionadas

Finalmente, la búsqueda se completa con un repaso a diferentes sumarios y resúmenes presentes en otras fuentes, como las bases de datos sobre tesis doctorales españolas TESEO (<https://www.educacion.gob.es/teseo/irGestionarConsulta.do>) y de todo el mundo OATD (<https://oatd.org/>), diferentes páginas sobre actas de congresos e informes de grupos de trabajo o incluso sitios web con información sobre aplicaciones de software recogidas en plataformas de desarrollo colaborativo como Github (<https://github.com/>). De entre esas fuentes destacan las siguientes: el sitio web de la Networked Knowledge Organization Systems (NKOS, <http://nkos.slis.kent.edu/>), los seminarios de UDC Consortium (<http://seminar.udcc.org/>) y la European Semantic Web Conference (ESWC, <https://eswc-conferences.org/events>). También se ha incluido información bibliográfica de sitios web de los proyectos sobre interoperabilidad de vocabularios controlados de la University of South Wales (Facet, Seneschal, Star, Stellar), los sitios relacionados con el proyecto Europeana (Athena Plus) y de grupos de estudio como el Library Linked Data Incubator Group del World Wide Web Consortium (W3C).

Como el objetivo principal de este trabajo es hacer una revisión lo más exhaustiva posible, los criterios establecidos para la selección de documentos no han sido muy restrictivos, por lo que no se han establecido restricciones temporales. Asimismo, se considera relevante para el estudio cualquier tipo de documento de texto en el que se trate el proceso de conversión de uno o varios SOC a SKOS, analizando incluso breves comunicaciones a congresos. En cuanto a los idiomas de los documentos analizados, se consideraron inglés, español, francés, alemán e italiano, y se rechazaron los documentos redactados en otras lenguas. También se han excluido del estudio aquellos casos o documentos en los que no se identifica de forma clara el nombre del vocabulario que se pretende trasladar a SKOS.

Con todo, finalmente se han identificado un total de 49 documentos en los que se trata el proceso de conversión de SOC a SKOS, cuyas referencias se recogen en la bibliografía. No se trata de una muestra, sino de una selección exhaustiva de documentos científicos que han tratado la *skosificación* de SOC, los cuales se someten a análisis y tratan de acuerdo a lo que se expone en la fase siguiente.

Fase 2. Establecimiento de criterios y análisis de los documentos seleccionados

En el análisis de los trabajos seleccionados se identifican los siguientes aspectos o elementos que permiten establecer conclusiones sobre metodologías y procesos más utilizados para la *skosificación* de SOC:

A. Características básicas de cada vocabulario *skosificado* ofrecidas en los documentos tratados:

1. Título completo del SOC.
2. Categoría del SOC que se describe. Si es nombrado con un término genérico (por ejemplo, vocabulario) se le asignará una categoría de acuerdo a la estructura que presente su contenido.
3. Materia o materias sobre las que trate su contenido.
4. Definición del objetivo para el que se construyó el SOC. La finalidad original para la que se creó el vocabulario controlado puede que sea distinta a la que se busca con el cambio a SKOS.
5. Descripción física de los SOC, en la que se atenderá al tamaño de los vocabularios, los elementos de su estructura y las relaciones que se establecen entre sus conceptos.
6. Formato del fichero fuente en el que se presenta el contenido del vocabulario, previo a su *skosificación*.

B. Razones que han llevado a la conversión de los SOC al formato SKOS.

C. Explicación de los instrumentos y procedimientos/metodologías empleados en la operación para la conversión de los SOC:

1. Identificación de las herramientas de software y los lenguajes usados en cada proyecto como vía de la transformación a SKOS.
2. Revisión de las equivalencias que se establecen entre los elementos de los SOC y el formato SKOS. Estas equivalencias deben estar claramente expresadas en la documentación manejada.
3. Se tendrán en cuenta los problemas encontrados al realizar el mapeo a SKOS y las soluciones que se hayan aplicado en los distintos casos para resolverlos.

RESULTADOS

En los 49 documentos que han sido tratados en este estudio, la exposición de la información que realizan los autores resulta más homogénea en la parte que se ha dedicado a las características básicas de los vocabularios

(estructura, tamaño y relaciones entre términos), mientras que en las otras secciones estudiadas (causas de *skosificación* y explicación de instrumentos y procedimientos empleados) la cantidad de información varía considerablemente de un documento a otro, siendo en algunos aspectos escasos o nulos los datos ofrecidos por los autores.

1. AGROVOC Thesaurus	27. Habitat Types
2. Amsterdam Museum Thesaurus	28. HASSET Thesaurus
3. Archaeological Sciences Thesaurus	29. Iconclass
4. Aria Catalogue	30. iDAI.gazetteer
5. Australian Public Affairs Information Service	31. Integrated Public Sector Vocabulary
6. Bibliopolis Thesaurus	32. International Press Telecommunication Council News Code
7. Bibliotheca Iuris Antiqui Thesaurus	33. International Standard for Country Codes ISO 3166
8. Building Materials Thesaurus	34. International Union for Conservation of Nature Classification
9. Camps	35. Library of Congress Subject Headings
10. Charda	36. Lista de encabezamientos de materias para bibliotecas públicas
11. Chinese Agricultural Thesaurus	37. Mathematics Subject Classification
12. Common Procurement Vocabulary	38. MDA Object Type Thesaurus
13. Dottrina Giurudica Classificazione materie giuridiche	39. Medical Subject Headings
14. e.Jewish.info Thesaurus	40. Thesaurus of Monument Types
15. EHRI Thesaurus	41. Nature 2000 AI Taxonomy
16. English Heritage Aircraft Type Thesaurus	42. Nuovo Soggetario Thesaurus
17. Environmental Application Reference Thesaurus	43. Organic Edunet Ontology
18. European Commission Library Thesaurus	44. Ra Thesaurus
19. European Nature Information System Species	45. Roman Thesaurus
20. European Training Thesaurus	46. Russian Library Bibliographic Classification
21. EUROVOC Thesaurus	47. Standard Thesaurus Wirtschaft
22. Evidence Thesaurus	48. TAXREF
23. Federal Register Thesaurus for Indexing Terms	49. Tesoro del Habitat del CEHAP
24. Gemeenschappelijke Thesaurus Audiovisuele Archieven (GEMET)	50. Thesaurus for the Social Sciences
25. General Management Environmental Thesaurus	51. Thesaurus of Greek Terms
26. Ghettos	52. Thésaurus pour description et indexation des archives locales

53. Thesaurus VRT	57. Universal Decimal Classification
54. Trilingual Cultural Thesaurus	58. Yleinen suomalainen asiasanasto
55. UNESCO Nomenclature	59. Zenon thesaurus
56. UNESCO Thesaurus	

Tabla 2. Títulos de los SOC *skosificados*

Los 49 documentos analizados presentan una distribución temporal desigual que va de 2004 hasta 2017, tal y como se puede apreciar en la *Figura 1*. En esos 14 años se han dado dos periodos con una mayor actividad de publicación en los que se recogen casi las tres cuartas partes de la producción total. En el cuatrienio que va desde 2006 hasta 2009 se publicaron 20 trabajos, y en el bienio 2012-2013 se editaron otros 16. Desde 2014 el número anual de publicaciones ha caído de forma considerable hasta 2017, en el que se publicó un único documento sobre conversión de vocabularios a SKOS.

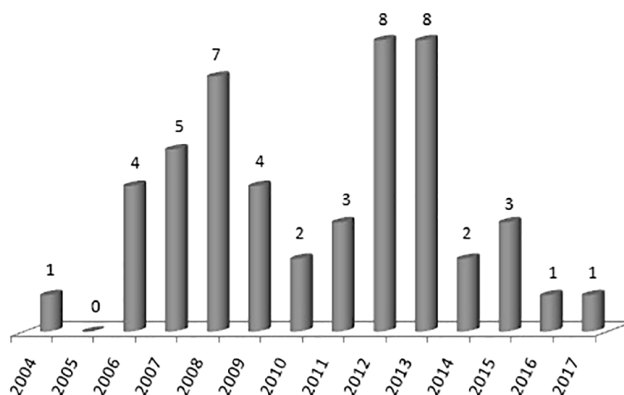


Figura 1. Cifras de documentos sobre procesos de *skosificación* por años de publicación

Características de los vocabularios en la documentación tratada

En los 49 documentos manejados se han podido identificar hasta 59 vocabularios diferentes (*Anexo 2*). De estos 59 vocabularios, 44 han visto reflejado su proceso de conversión en un único documento, mientras que los 15 restantes han sido objeto de estudio en varias ocasiones. Algunos de estos 15 SOC han llegado a ser tratados hasta en cinco documentos diferentes, como es el caso de los tesauros que tratan sobre arqueología y rehabilitación del patrimonio histórico británico, que pueden considerarse como ejemplos de *skosificación* en la literatura científica.

Los vocabularios identificados se han agrupado en seis tipos de SOC: esquemas de clasificación, listas de autoridades, listas de encabezamientos de materias, listas de palabras clave, ontologías y tesauros (Figura 2). Al considerar cada categoría en concreto destacan por sus elevados porcentajes los tesauros (aproximadamente uno de cada tres vocabularios) y los esquemas de clasificación (uno de cada cinco vocabularios), quedando el resto de tipos reducido solamente a nueve SOC que se distribuyen de la siguiente manera: tres listas de autoridades, tres listas de encabezamientos de materias, dos listas de palabras clave y una ontología.

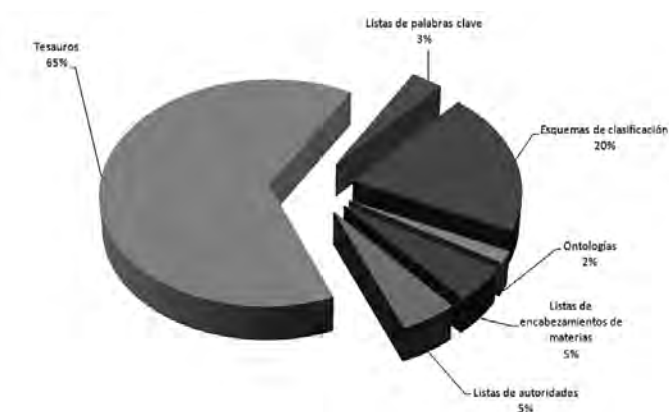


Figura 2. Porcentajes de los SOC convertidos ordenados por tipos

Los 59 SOC abarcan hasta 17 materias diferentes, aunque con una presencia desigual. Entre las categorías con mayor presencia, la multidisciplinar, con 16 vocabularios, resulta la más numerosa, seguida por la arqueología y conservación del patrimonio histórico con 10 SOC, y la ecología y medioambiente donde se encuentran siete vocabularios. El número de SOC que tratan alguna de las otras 14 disciplinas queda distribuido de la siguiente manera: con tres recursos (arte, comunicación, derecho, economía y geografía), con dos (agricultura, educación e historia) y con solamente uno (archivística, lengua griega, matemáticas, medicina, transporte y urbanismo).

En lo que se refiere a los objetivos para los que fueron creados los vocabularios de partida, punto A.4 de la fase 2 de la metodología, se puede comprobar que no todos los documentos estudiados recogen esta información porque se han centrado más en el proceso de la *skosificación* o en el vocabulario que la aplicación de dicho proceso ha dado como resultado, que en el SOC fuente. Sólo en 25 casos (42.4 %), se hace mención expresa a los objetivos

originales de los vocabularios, la mayoría de los cuales fueron creados para ser utilizados en entornos delimitados con anterioridad como bases de datos o colecciones de bibliotecas, no siempre coincidentes con los objetivos de sus versiones en SKOS.

La descripción detallada sobre el número de términos en los vocabularios, su estructura y la definición de sus relaciones, punto A.5 de la fase 2 de la metodología, es poco reflejada en la documentación. Un significativo nivel de detalle sólo se ofrece en 18 de los 49 documentos analizados, siendo el número total de términos la característica que más encuentra en las descripciones. Además, se ha podido constatar que en 17 de los 49 trabajos revisados no se proporciona ninguna descripción sobre el tamaño, la estructura y las relaciones entre términos de los SOC de partida.

Si se consideran de forma conjunta la expresión sobre los objetivos y la descripción física de los vocabularios, sólo se consigue encontrar un resultado positivo en un 15.25 % de documentos observados.

En cuanto a los formatos originales de los vocabularios o SOC antes de su *skosificación*, se declaran en tres cuartas partes de los documentos manejados. En los vocabularios que han sido tratados más de una vez los formatos de archivo han sido siempre los mismos, excepto en los casos del Tesauro de la Unesco, recogido en HTML (Pastor-Sánchez *et al.*, 2013) y XML (Francesconi *et al.*, 2007), y los Library of Congress Subject Headings, recogido en MARC21 (Harper, 2006) y MARCXML (Summers *et al.*, 2008).

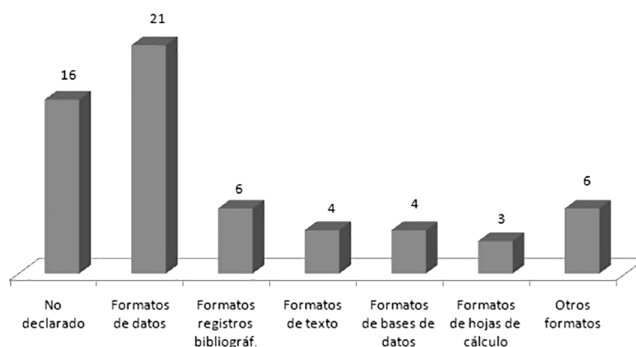


Figura 3. Tipos de formatos originales de SOC antes de la *skosificación*

En lo que se refiere a los formatos de los SOC de partida (*Figura 3*), se comprueba que la categoría de formatos de datos (XML, CSV y TSV) resulta ser la más numerosa con 21 vocabularios. En las otras categorías destacan seis en formatos de registros bibliográficos (MARC, MARC21 y MARCXML), cuatro en formatos de texto (TXT y TeX), cuatro en formatos de bases de datos (Acces y FirebirdDB), tres en formatos de hojas de cálculo (XLS) y seis distribuidos en otros formatos como HTML, JSON, OWL y PDF.

Razones para la conversión a SKOS

Los motivos que han llevado a la conversión de los vocabularios no han sido declarados en casi la mitad de los documentos analizados. Entre los que sí explicitan las razones para el cambio, la causa principal ha sido la búsqueda de una mejor interoperabilidad de los SOC convertidos con otros vocabularios para que puedan adaptarse a la web semántica. Otras razones que también se han argumentado para el uso de SKOS en la transformación de vocabularios son las siguientes:

- servir para integrar información multilingüe y para la expansión semántica de los conceptos (Rodríguez y Priddy, 2011: 7),
- permitir una mejor reutilización del contenido de los vocabularios en otros contextos (De Boer *et al.*, 2016: 3)
- resultar de más fácil manejo que otras esquemas, como OWL, porque no es necesario dotar al vocabulario de una semántica y razonamiento formales (Caracciolo *et al.*, 2012: 73).

Instrumentos, procedimientos y lenguajes utilizados para la skosificación

La cantidad de información sobre los instrumentos, procedimientos y lenguajes usados en la *skosificación* ha resultado más bien escasa. En el análisis de estos parámetros el total de vocabularios a considerar es de 63 en lugar de 59 porque cuatro de ellos han sido objeto de conversión a SKOS más de una vez y en diferentes condiciones.

En el *Anexo 1* se puede comprobar que las vías de conversión a SKOS son declaradas en 71.43 % de los vocabularios analizados, mientras que la declaración de nombres de las herramientas utilizadas ha sido mucho menor (31.74 %). Dadas las diferencias metodológicas observadas entre los documentos al describir las labores realizadas en la sistematización de las vías de conversión a SKOS, se ha optado por agruparlas con base en la categoría del

lenguaje que han empleado. En este sentido, se han reunido los casos conforme a tres tipos de lenguajes: de programación, de transformación y de mapeo, según sea la técnica predominante que hayan empleado los autores para la *skosificación*.

Lenguajes de programación. Esta categoría tiene en cuenta elementos de software (scripts y aplicaciones) creados para la conversión de los vocabularios a SKOS. Los lenguajes de programación que se han observado, en orden de su frecuencia de uso en los procesos de transformación, son los siguientes: Python (seis), C (cinco), Java (tres), Perl (tres), PHP (dos) y Prolog (dos). En tres casos se opta por aplicaciones de consola: uno con skoshassetgenerator (Bell, 2013: 8) y los otros dos con Stellar console (Beer *et al.*, 2014: 25; Felicetti *et al.*, 2016: 37).

Los nombres de las aplicaciones y *scripts* más usados se suministran en contadas ocasiones. Sólo en 20 vocabularios se identifican de manera concreta las herramientas empleadas, entre las que destacan EH2SKOS en lenguaje C, Skosify en Python y SWI Prolog, por su uso en más de una ocasión.

Lenguajes de transformación. Esta denominación se emplea en los casos de vocabularios *skosificados* con base en el lenguaje XSLT, el cual permite la transformación de documentos XML en otros en ese mismo lenguaje. El uso de la transformación XSL se ha constatado en 12 ocasiones para la conversión de 10 vocabularios. Como XSLT exige que el archivo a convertir esté en XML, en seis vocabularios (Common Procurement Vocabulary, Eurovoc, European Commission Library Thesaurus, European Training Thesaurus, Library of Congress Subject Headings y UNESCO Thesaurus) se tuvo que adaptar el formato original para poder aplicar esta vía de transformación. Sólo se ha suministrado en una ocasión el nombre de la aplicación utilizada para la transformación XSL: XSLTProc (Álvarez Rodríguez, Rubiera Azcona y Polo Paredes, 2007).

En los trabajos sobre transformación de los tesauros de la English Heritage intentaron *skosificar* los SOC aplicando la tecnología XSL, pero al encontrar problemas con los tesauros de gran tamaño tuvieron que renunciar a ese método (Tudhope y Binding, 2008a y 2008b; May *et al.*, 2009).

Lenguajes de mapeo. Con este concepto nos referimos a los lenguajes que se utilizan para mapear entre bases de datos y RDF. Siete vocabularios emplean esta vía de transformación, seis a través del lenguaje D2RQ y uno con xR2RML. En los casos en que es utilizado D2RQ no se relaciona el nombre de la herramienta utilizada en el mapeo, mientras que en el caso de xR2RML sí se señala la herramienta empleada: Morph-xR2RML (Callou *et al.*, 2015: 21).

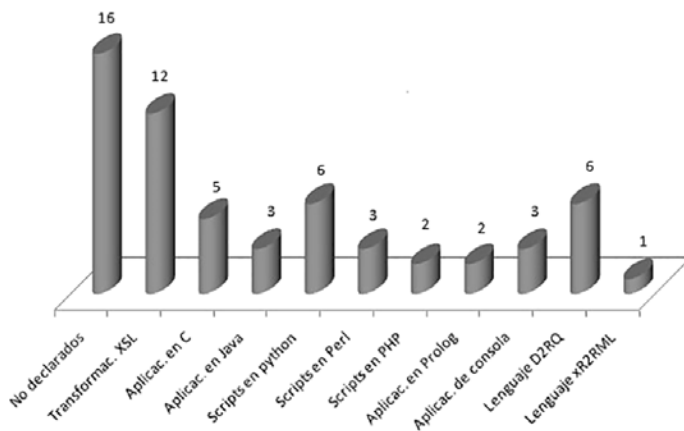


Figura 4. Vías de transformación de los SOC a SKOS

En cuanto al mapeo de equivalencias de elementos entre los vocabularios y SKOS, se puede destacar la ausencia de información en 18 documentos. En esos casos se debe suponer que se estableció dicho mapeo al realizar los procesos de *skosificación*, pero que la información no fue trasladada a la documentación sobre los correspondientes proyectos. En los documentos restantes (64 %) se muestran equivalencias entre los esquemas, aunque con distintos niveles de detalle como puede comprobarse en el *Anexo 2*. Los elementos de SKOS con una mayor presencia en los mapeos de equivalencias son la clase *skos:Concept* y las propiedades *skos:altLabel*, *skos:broader*, *skos:narrower*, *skos:prefLabel* y *skos:related*.

En algunos casos el mapeo hacia SKOS ha generado dificultades para representar todas las particularidades estructurales que se encuentran en los vocabularios. A este respecto se ha advertido en la documentación revisada tres formas de suplir este problema:

- Creación de extensiones en elementos ya existentes o de nuevas propiedades de SKOS. En el caso de las extensiones, se ha encontrado una que permite especificar el género de la etiqueta de un concepto preferido, *skos-ehri:prefFemaleLabel*, creada para el EHRI Thesaurus (Rodríguez y Priddy, 2011: 8). Sobre la creación de nuevas propiedades, se puede observar *skos:notationProperty*, que representa diversos sistemas de codificación de un mismo lugar geográfico (concepto) en la *skosificación* del vocabulario ISO 3166 (Voss, 2007: 4).

- Combinación en el uso de elementos de SKOS para representar situaciones que con la aplicación directa del formato resultarían erróneas. Este método es aplicado en la *skosificación* del tesoro UNESCO, que es definido como un esquema de conceptos, mientras que las áreas de conocimiento y microtesoros son vistas como colecciones. Las primeras quedan asociadas al *schema* SKOS con la propiedad *skos:inScheme* y los microtesoros se vinculan a un área de conocimiento con la propiedad *skos:member* (Pastor-Sánchez *et al.*, 2013: 1031-1032). También se recoge este problema en la *skosificación* de Eurovoc, pero se resuelve de manera distinta considerando a los microtesoros y los campos temáticos como instancias de *skos:ConceptScheme*, lo que obliga a incluir la propiedad *owl:hasScheme* para poder afirmar los enlaces jerárquicos entre instancias de *skos:ConceptScheme* (Polo Paredes, Álvarez Rodríguez y Rubiera Azcona, 2008).
- Creación de clases y propiedades en un vocabulario para adaptar su estructura a SKOS. Esta alternativa se encuentra en la *skosificación* de MeSH, que resuelve características fuera de los estándares con *mesh:CompoundConcept* y *mesh:Qualifier*, subclases de *skos:Concept*, y propiedades como *mesh:considerAlso*, *mesh:publicMeSHNote* y *mesh:dateCreated*, subpropiedades de *skos:note* y *skos:editorialNote* respectivamente (Van Assem *et al.*, 2006: 13). Otros vocabularios han adaptado sus elementos con propiedades como *zbwext:useInsteadNote*, subpropiedad de *skos:note* en el Standard Thesaurus Wirtschaft (Neubert, 2009); *ipsv:obsoleteTerm*, subpropiedad de *skos:HiddenLabel* en el Integrated Public Sector Vocabulary (Van Assem *et al.*, 2006: 7) o *gtaa:hasLinkedTerm* y *gtaa:hasDebatedLine*, subpropiedades de *skos:related* en el Gemeenschappelijke Thesaurus Audiovisuele Archieven (Malaise y CHOICE Team, 2007: 27).

En general, la información que suministran los documentos estudiados es escasa o limitada para poder establecer unos patrones clave que definan los procesos de *skosificación*. De los 49 documentos examinados, que tratan la conversión de 59 vocabularios, solamente se han encontrado seis que presentan de forma detallada información sobre cada uno de los puntos que se explicitan en la fase 2 de la metodología: Álvarez Rodríguez, Rubiera Azcona y Polo Paredes, 2007; Casellas, 2012; Felicetti *et al.*, 2016; Pastor-Sánchez *et al.*, 2013; Polo Paredes, Álvarez Rodríguez y Rubiera Azcona, 2008; Van Assem *et al.*, 2006.

El artículo de Van Assem *et al.* (2006) evalúa de forma detallada la conversión de tres vocabularios a partir de un esquema en tres partes: análisis de sus características principales junto con su documentación y guías de uso, definición del mapeo de sus ítems de datos a SKOS y descripción del software usado en la conversión. En varios documentos tratados declaran que han tomado como modelo el mencionado esquema, y podría ser “el esquema” a tener en cuenta en un proceso de *skosificación* de SOC. Por otra parte, el trabajo debido a Pastor-Sánchez, Martínez-Méndez y Rodríguez-Muñoz (2009) ofrece también una descripción de todo el proceso de transformación a SKOS, incluyendo detalles interesantes sobre el software a utilizar y de la interfaz de navegación que se debe aplicar a los vocabularios.

CONCLUSIONES

Si consideramos de forma particular los resultados obtenidos de los estudios revisados, se pueden extraer las siguientes conclusiones:

- La *skosificación* de SOC ha sido abundante desde los orígenes del formato SKOS, incluso en la etapa en la que éste solamente era un borrador, aunque el tratamiento de dicho proceso ha sufrido un considerable descenso en los últimos años observados. No obstante, la disminución de publicaciones sobre la *skosificación* de vocabularios no resta a nuestro entender el interés por el uso de este formato que pueda despertar en los autores, como ya señalábamos en la introducción al hablar de la presencia de vocabularios en SKOS en BARTOC (<https://bartoc.org/>) y en la versión antigua de Datahub (<https://old.datahub.io/>), lo que da muestra de la madurez alcanzada por el formato SKOS frente a otras opciones de representación de vocabularios.
- Dos terceras partes de los SOC convertidos pueden inscribirse dentro de la categoría tesaurus. Esto puede resultar una obviedad, dado que este tipo de vocabularios controlados es el que mejor se adapta a la estructura de SKOS, pero también se puede considerar de forma que el mencionado formato sea una vía válida con la cual sea posible adaptar otros tipos de SOC a la web semántica, como se puede comprobar en el otro tercio de vocabularios transformados a SKOS en los estudios analizados.
- La información descriptiva sobre el estado y estructura de los vocabularios antes del proceso de conversión no es muy completa, en especial la concerniente a la estructura original de los vocabularios.

- Los formatos de datos, principalmente XML y CSV, parecen ser los que mejor se adecuan a los procesos de transformación de vocabularios al formato SKOS.
- La justificación del cambio de formato que argumentan los autores de los trabajos analizados es casi exclusivamente la búsqueda de una mayor interoperabilidad de los vocabularios transformados hacia otros vocabularios controlados. Esta finalidad concuerda con lo expresado en la guía introductoria sobre SKOS elaborada por el W3C, en cuya introducción se señala que el objetivo del formato SKOS es que los vocabularios conceptuales puedan transferirse en un espacio compartido en el que se permita una reutilización más amplia y una mejor interoperabilidad (World Wide Web Consortium, 2009a).
- La definición del mapeo de equivalencias entre los elementos de los vocabularios y SKOS ha resultado algo limitada: las clases y propiedades más representadas en el formato SKOS se corresponden con las típicas relaciones semánticas que se establecen entre los términos de los tesauros. Algo ya esperado, dado que la mayor parte de los vocabularios son tesauros. Sin embargo, sí ha quedado demostrada la flexibilidad existente en el formato SKOS, el cual permite la asociación de subpropiedades con las que es posible definir relaciones que no están recogidas en las normas de tesauros y otros vocabularios controlados.
- El software empleado en los procesos de conversión de vocabularios es la parte que menos atención ha recibido en la bibliografía científica tratada en el presente estudio. En muchos casos la información dada por los autores se limita a señalar cuál es el lenguaje utilizado en la transformación, en alguno de los tipos que hemos establecido en el apartado de resultados (programación, mapeo o transformación), quedando incompletos muchos datos sobre las labores de transición a SKOS. Las explicaciones que se ofrecen sobre el funcionamiento de las aplicaciones de conversión o de las plantillas XSLT empleadas son muy puntuales.

Tras la revisión de tan abundante material se puede afirmar que el formato SKOS es visto por los autores como una buena opción para transformar los vocabularios de valores, los cuales pueden ser adaptados a un entorno de datos enlazados sin que sea necesario realizar profundos cambios en sus estructuras o en la forma en que se expresan sus conceptos. Aunque el afán que se demuestra en la publicación de los procesos de *skosificación* de vocabularios parece haber descendido en los tiempos más recientes, el modelo SKOS continúa manteniéndose como un importante formato para la representación de

vocabularios controlados. Es de esperar que dicho *schema* continúe siendo empleado en un futuro por los creadores y gestores de vocabularios controlados, si bien sería deseable que los trabajos que resulten de las actividades realizadas en los procesos para los cambios de formato de los vocabularios fueran más detallados. Es conveniente que los autores recojan toda la información sobre las tareas llevadas a cabo en cada una de las fases de los procesos de conversión de vocabularios para que lectores y usuarios puedan ver el desarrollo completo.

La abundante transformación de SOC al formato SKOS revela que el contenido de dichas herramientas continúa considerándose de utilidad para la web, y por ello siguen realizándose esfuerzos para su adaptación a las necesidades que demanda dicho entorno.

REFERENCIAS

- Akbari-Daryan, Saedeeh, Fariborz Khosravi y Mahdi Ebrahimi. 2016. "SKOSification of Trilingual Cultural Thesaurus (TCH) of National Library of Iran (NLI): A step in line with NLI's Linked Data strategy", en IFLA World Library and Information Congress, Columbus, Ohio (USA), 13-19 de agosto.
<http://library.ifla.org/2091/1/113-akbari-daryan-en.pdf>
- Agnoloni, Tommaso, Elisabetta Marinai, Ginevra Peruginelli, María Teresa Sagri y Daniela Tiscornia. 2013. "Annotation schema for legal doctrine: a case study on DoGi database". *Italian Journal of Library and Information Science* 4 (1): 227-249.
<https://doi.org/10.4403/jlis.it-5480>
- Álvarez Rodríguez, José María, Emilio Rubiera Azcona y Luis Polo Paredes. 2007. "Generación automática de ontologías en SKOS de clasificaciones estándar de productos: Common Procurement Vocabulary (CPV)", en II Jornadas sobre Ontologías Web y Semántica, Zaragoza, España, 11 al 14 de septiembre. Madrid: Thomson.
- Beer, Nikolaos, Kristin Herold, Wibke Kolbmann, Thomas Kollatz, Matteo Romanello, Sebastian Rose y Niels Oliver Walkowski. 2014. *Interdisciplinary Interoperability*. Göttingen: Georg-August-Universität (DARIAH-DE working papers, 3).
<http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2014-3.pdf>
- Bell, Lucy. 2013. *SKOS-HASSET*. Version 0.04. (17 de abril de 2013).
http://repository.jisc.ac.uk/5160/1/SKOS%2DHASSET_final_report.pdf
- Bergamin, Giovanni y Anna Lucarelli. 2013. "The Nuovo Sogettario as a service for the linked data world". *Italian Journal of Library and Information Science* 4 (1): 213-226.
<https://doi.org/10.4403/jlis.it-5474>
- Binding, Ceri, Phil Carlisle, Keth May, Douglas Tudhope y Andreas Vlachidis. 2012. "STAR, STELLAS and SKOS", en Computer Applications and Quantitative Methods in Archaeology (CAA) 2012 conference, Southampton, 26 - 29 March 2012.
http://hypermedia.research.southwales.ac.uk/media/files/documents/2012-05-11/SKOS_CAA2012_Glam1.ppt

- Callou, Cécile, Franck Michel, Catherine Faron-Zucker, Chloé Martin y Johan Montagnat. 2015. "Towards a shared reference thesaurus for studies on history of zoology, archaeozoology and conservation biology", en 12th Extended Semantic Web Conference (ESWC), Portoroz (Slovenia), 31 de mayo a 4 de junio. <https://hal.archives-ouvertes.fr/hal-01146638/document>
- Caracciolo, Caterina, Ahsan Morshed, Armando Stellato, Gudrun Johannsen, Yves Jaques y Jacques Keizer. 2012. "Thesaurus Maintenance, Alignment and Publication as Linked Data. The AGROVOC Use Case". *International Journal of Metadata, Semantics and Ontologies* 7 (1): 65-75.
- Casellas, Nuria. 2012. "Linked Legal Data: A SKOS Vocabulary for the Code of Federal Regulations". *Semantic Web Journal* s. n. (Artículo rechazado, pero publicado). <http://www.semantic-web-journal.net/content/linked-legal-data-skos-vocabulary-code-federal-regulations>
- De Boer, Victor, Jan Wielemaker, Judith van Gent, Michiel Hildebrand, Antoine Isaac, Jacco Van Ossenbruggen y Guus Schreiber. 2012. "Supporting linked data production for cultural heritage institutes: The Amsterdam Museum Case Study", en 9th Extended Semantic Web Conference (ESWC 2012), Heraklion, Greece, 27 - 31 de mayo. Berlin: Springer. https://doi.org/10.1007/978-3-642-30284-8_56
- De Boer, Victor, Matthias Priem, Michiel Hildebrand, Nico Verplancke, Arjen de Vries y Johan Oomen. 2016. "Exploring Audiovisual Archives Through Aligned Thesauri", en 10th Metadata and Semantics Research Conference (MTSR 2016), 22 - 25 november. Berlin: Springer. https://doi.org/10.1007/978-3-319-49157-8_19
- Felicetti, Achille, Ilena Galluccio, Cinzia Luddi, Maria Letizia Mancinelli, Tiziana Scarselli y Antonio Davide Madonna. 2016. "Integrating Terminological Tools and Semantic Archaeological Information: the ICCD RA Schema and Thesaurus", en Proceedings of the Workshop on Extending, Mapping and Focusing the CRM, Poznań, Poland, September 17. <http://ceur-ws.org/Vol-1656/paper3.pdf>
- Francesconi, Enrico, S. Faro, E. Marinai, V. Sandrucci, F. Bargellini, M. Biasiotti y L. Peruginelli. 2007. *Report on execution and results of the interoperability tests: EUROVOC Studies - LOT2. Deliverable 2.3*. Firenze: Italian National Research Council. <http://eurovoc.europa.eu/drupal/sites/all/files/D2.3-Final.pdf>
- Grimouard, Claire Sibille de. 2014. "The Thesaurus for French Local Archives and the Semantic Web". *Procedia - Social and Behavioral Sciences* 147: 206-212. <https://doi.org/10.1016/j.sbspro.2014.07.153>
- Guzmán Luna, Jaime A., Durlay Torres Pardo y Alba Nubia López García. 2006. "Desarrollo de una ontología en el contexto de la web semántica a partir de un tesoro documental tradicional". *Revista Interamericana Bibliotecología* 29 (2): 79-95.
- Harper, Corey A. 2006. "Encoding Library of Congress Subject Headings in SKOS: Authority Control for the Semantic Web", en DC 2006: DCMI International Conference on Dublin Core and Metadata Applications, Colima, México, 3-6 de octubre.
- Hodge, Gail. 2000. *Systems of Knowledge Organization for digital libraries: beyond traditional authority files*. Washington: Council on Library and Information Resources. <http://www.clir.org/pubs/reports/pub91/contents.html>

- Ion, Patrick y Wolfram Sperber. 2010. "MSC 2010 in SKOS - the Transition of the MSC to the Semantic Web". *European Mathematical Society Newsletter* 84: 55-57.
- Judaica Europeana. 2008. D3.4: JUDAICA controlled vocabularies expressed in RDF/SKOS. ECP-2008-DILI-538025.
https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Judaica/Deliverables/D3.4_JUDAICA_controlled_vocabularies.doc
- Lange, Cristoph, Patrick Ion, Anastasia Dimou, Charalampos Bratsas, Wolfram Sperber, Michael Kohlhase y Ioannis Antoniou. 2012. "Bringing Mathematics to the Web of Data: The Case of the Mathematics Subject Classification", en Extended Semantic Web Conference (ESWC 2012), Heraklion, Greece, 27 - 31 de mayo.
https://doi.org/10.1007/978-3-642-30284-8_58
- Larrenova, Olga y Vasily Pavlov. 2017. "The publication of a knowledge organization classification system as linked data vocabulary", en NKOS 2017: 17th European Networked Knowledge Organization Systems (NKOS) Workshop, Thessaloniki, Greece.
<http://ceur-ws.org/Vol-1937/paper6.pdf>
- Leroi, Marie Veronique y Johann Holland. 2010. *Athena: Guidelines for mapping into SKOS, dealing with translations. Deliverable 4.2 for the ECP-2005-CULT-038099 project.*
<http://www.athenaeurope.eu/getFile.php?id=684>
- Malaise, Veronique y CHOICE Team. 2007. "SKOS: a model for metadata representation and interoperability Dutch Cultural Heritage Institution thesaurus conversion use case", en Delos Multimatch Workshop, Tirrenia, Italia, 15 de febrero.
- Martínez-Conde, María Luisa. 2013. "La conversión a SKOS de la Lista de Encabezamientos para Bibliotecas Públicas". *Boletín ANABAD LXIII* (3): 107-116.
- Martino, Monica de y Riccardo Albertoni. 2011. "A multilingual/multicultural semantic-based approach to improve Data Sharing in a SDI for Nature Conservation. International". *Journal of Spatial Data Infrastructures Research* 6: 206-233.
<https://doi.org/10.2902/1725-0463.2011.06.art10>
- May, Keith, Douglas Tudhope, Ceri Binding y Andreas Vlachidis. 2009. "The Semantic Technologies for Archaeological Resources (STAR) project's use of SKOS", en ATHENA WP4 SKOS Workshop.
<http://www.athenaeurope.org/getFile.php?id=288>
- Méndez, Eva y Jane Greenberg. 2012. "Datos enlazados para vocabularios abiertos y marco general de HIVE". *El Profesional de la Información* 21 (3): 236-244.
- Miles, Alistair J., Nikki Rogers y Dave Beckett. 2004. *Migrating Thesauri to the Semantic Web: Guidelines and case studies for generating RDF encodings of existing thesauri. SWAD-Europe.*
<https://www.w3.org/2001/sw/Europe/reports/pdf/8.8.pdf>
- Neubert, Joachim. 2009. "Bringing the "thesaurus for economics" on to the web of linked data", en Proceedings of the Linked Data on the Web Workshop (LDOW2009), Madrid, Spain, 20 de abril.
http://www.assalweb.org/assal_nueva/documentos/Estandares/ldow2009_paper7.pdf
- Pastor-Sánchez, Juan Antonio, Francisco Javier Martínez-Méndez y José Vicente Rodríguez-Muñoz. 2009. "Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives". *Information Research* 14 (4): paper 422.
<http://InformationR.net/ir/14-4/paper422.html>

- Pastor Sánchez, Juan Antonio. 2013. "ISO-THES: Ampliando Skos a partir de la norma de tesauros ISO 25964". *Anuario ThinkEPI* 7: 189-193.
- Pastor-Sánchez, Juan Antonio, Francisco Javier Martínez-Méndez, Rosana López-Carreño y José Vicente Rodríguez-Muñoz. 2013. "UNESKOS: Publicación como Linked Open Data de la Nomenclatura Internacional de Ciencia y Tecnología y del Tesouro UNESCO", en I Congresso ISKO Espanha e Portugal y XI Congresso ISKO España, Oporto, Portugal, 7-9 de noviembre.
<http://eprints.rclis.org/24272/1/ICongressoISKOEspanhaePortugal000211987.pdf>
- Polo Paredes, Luis, José María Álvarez Rodríguez y Emilio Rubiera Azcona. 2008. "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, 2 de junio.
<http://www.josemalvarez.es/web/mypapers/siedl2008.pdf>
- Rajabi, Enayat, Miguel Ángel Sicilia y Salvador Sánchez Alonso. 2013. "A Simple Approach towards SKOSification of Digital Repositories", en *MTSR 2013: 7 th Metadata and Semantics Research Conference*, Thessaloniki, Greece, 19 - 23 de noviembre. Springer.
https://doi.org/10.1007/978-3-319-03437-9_8
- Rajbhandari, Sachit y Johannes Keizer. 2012. "The AGROVOC Concept Scheme: A Walkthrough". *Journal of Integrative Agriculture* 11 (5): 694-699.
- Rodriguez, Kepa Joseba y Mike Priddy. 2011. *European Holocaust Research Infrastructure Theme: INFRA-2010-1.1.4. Multilingual Search Interface. Deliverable D19.3. EHRI Project*.
<https://www.ehri-project.eu/sites/default/files/downloads/Deliverables/D%2019%203%20Multilingual%20Search%20Interface.doc>
- Romanello, Matteo. 2012. *Skosifying an Archaeological Thesaurus - Computers for the Classics*.
<https://c4tc.wordpress.com/2012/10/08/skosifying-an-archaeological-thesaurus>
- Saorín Pérez, Tomás y Juan Antonio Pastor-Sánchez. 2012. "Gestión documental y de contenidos web". *Anuario ThinkEPI* 6: 232-239.
- Sathn, Lena Louise. 2015. "Vocabulary alignment for archaeological Knowledge Organization Systems", en 14th Workshop on Networked Knowledge Organization Systems at International Conference on Theory and Practice of Digital Libraries (TPDL), Poznan, Poland, 14-18 de septiembre.
<https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2015/content/NKOS2015-presentation-stahn.pdf>
- Slavic, Aida y Marie Balikova. 2011. "Linking the Web of Knowledge: UDC's role in bridging information services and resources", en 12. konference Archivy, knihovny, muzea v digitálním sv t , Prague, 30 de noviembre y 1 de diciembre.
<http://www.skipcr.cz/akce-a-projekty/dokumenty/akm-2011/Slavic.pdf>
- Solomou, Georgia y Theodore Papatheodorou. 2010. "The Use of SKOS Vocabularies in Digital Repositories: The DSpace case", en *IEEE Fourth International Conference on Semantic Computing (ICSC)*, Pittsburgh (USA), 22-24 de septiembre.
https://www.researchgate.net/profile/Georgia_Solomou/publication/224193184_The_Use_of_SKOS_Vocabularies_in_Digital_Repositories_The_DSpace_Case/links/00463536110b38796b000000/The-Use-of-SKOS-Vocabularies-in-Digital-Repositories-The-DSpace-Case.pdf

- Spampinato, Daria y Ignazio Zangara. 2013. "Classical Antiquity and semantic content management on Linked Open Data", en Proceedings of the 1st International Workshop on Collaborative Annotations in Shared Environment: metadata, vocabularies and techniques in the Digital Humanities, Florence, Italy, 10 de septiembre. <http://doi.org/10.1145/2517978.2517992>
- Summers, Ed, Antoine Isaac, Clay Redding y Dan Krech. 2008. "LCSH, SKOS and Linked Data", en DC 2008: Proceedings of the International Conference on Dublin Core and Metadata Applications, Berlin, 22-26 de septiembre. <http://dcpapers.dublincore.org/pubs/article/viewFile/916/912>
- Tordai, Anna, Borys Omelayenko y Guus Schreiber. 2007. "Semantic excavation of the city of books", en Proceedings of the Semantic Authoring, Annotation and Knowledge Markup Workshop (SAAKM2007) located at the 4th International Conference on Knowledge Capture (KCap 2007), Whistler, British Columbia, Canada, October 28-31.
- Troncy, Raphael. 2008. "Bringing the IPTC News Architecture into the Semantic Web", en The 7 International Semantic Web Conference (ISWC 2008), Karlsruhe, Alemania, 26-30 de octubre. https://doi.org/10.1007/978-3-540-88564-1_31
- Tudhope, Douglas y Ceri Binding. 2008a. "Experiences with Knowledge Organization System Services from the STAR Project". *Signum* 36 (5).
- Tudhope, Douglas y Ceri Binding. 2008b. *Making KOS Machine Understable. Project Delos: A Network of Excellence in Digital Libraries*. Additional Report for Work Package 5. <http://hypermedia.research.southwales.ac.uk/media/files/documents/2008-07-05/Additional-report-wp5.pdf>.
- Tudhope Douglas, Ceri Binding y Keith May. 2008. "Semantic interoperability issues from a case study in archaeology", en Semantic Interoperability in the European Digital Library, Proceedings of the First International Workshop Semantic Interoperability in the European Digital Library (SIEDL), Tenerife, 2 de junio. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.556.7043&rep=rep1&type=pdf>
- Tuominen, Jouni, Matias Frosterus, Kim Viljanen y Eero Hyvönen. 2009. "ONKI SKOS Server for Publishing and Utilizing SKOS Vocabularies and Ontologies as Services", en ESWC 2009 Heraklion Proceedings of the 6th European Semantic Web Conference on The Semantic Web: Research and Applications, Heraklion, Greece, 31 de mayo - 4 de junio. https://doi.org/10.1007/978-3-642-02121-3_56
- Van Assem, Mark, Véronique Malaise, Alistair Miles y Guus Schreiber. 2006. "A Method to Convert Thesauri to SKOS", en The Semantic Web: Research and Applications. 3rd European Semantic Web Conference, ESWC 2006, Budva, Montenegro, 11-14 de junio. https://doi.org/10.1007/11762256_10
- Van Gendt, Marjolein, Antoine Isaac, Lourens van der Meij y Stefan Schlobach. 2006. "Semantic Web Techniques for Multiple Views on Heterogeneous Collections: A Case Study", en X European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2006), Alicante (Spain), 17 - 22 de septiembre. Berlin: Springer. https://doi.org/https://doi.org/10.1007/11863878_36

- Villazón Terrazas, Boris Marcelo. 2011. "A method for reusing and re-engineering non-ontological resources for building ontologies". PhD Thesis. Madrid: Universidad Politécnica de Madrid.
<http://oa.upm.es/6338/1/BorisVillazonTerrazas.pdf>
- Voss, Jakob. 2007. "Encoding changing country codes for the Semantic Web with ISO 3166 and SKOS", en 2nd International Conference on Metadata and Semantics Research (MTSR07). Corfu, Greece, 1-2 de octubre. Boston: Springer.
- World Wide Web Consortium. 2009a. *SKOS simple knowledge organization system primer*. W3C Recommendation 18 August 2009. Antoine Isaac y Ed Summers (eds.).
<https://www.w3.org/TR/skos-primer/>
- World Wide Web Consortium. 2009b. *SKOS simple knowledge organization system reference*. W3C Recommendation 18 August 2009. Alistair Miles y Sean Bechhofer (eds.).
<http://www.w3.org/TR/skos-reference>
- Xian Guo Jian, Zhao Ruixue, Zhu Liang y Kou Yuantao. 2012. "Conversion and consumption of Chinese agricultural thesaurus as SKOS", en Proceedings of International Symposium on Agricultural Ontology Service (AOS2012), Kuching, Malaysia, 3-4 de septiembre.
- Zapilko, Benjamin, Johann Schaible, Philipp Mayr y Brigitte Mathiak. 2013. "TheSoz: A SKOS Representation of the Thesaurus for the Social Sciences". *Semantic Web Journal* 4 (3): 257-263.
- Zapilko, Benjamin y York Sure. 2009. *Converting the TheSoz to SKOS: GESIS Technical Report 2009/07*.
https://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2009/TechnicalReport_09_07.pdf

Para citar este texto:

- Mochón Bezares, Gonzalo, Eva María Méndez Rodríguez y Gema Bueno de la Fuente. 2020. "Conversión normalizada (SKOS) de sistemas de organización del conocimiento interoperables en la web". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 34 (82): 53-86.
<http://dx.doi.org/10.22201/iibi.24488321xe.2020.82.58071>

Anexo 1 Información sobre el proceso de conversión de los vocabularios

Nombre del vocabulario	Referencias que tratan el vocabulario	Extensión de archivo	Vía de transformación	Herramientas utilizadas
AGROVOC Thesaurus	Caracciolo <i>et al.</i> , 2012; Rajbhandari y Keizer, 2012	OWL		
Amsterdam Museum Thesaurus	De Boer <i>et al.</i> , 2012			
Archaeological Sciences Thesaurus	Tudhope y Binding, 2008a; Tudhope y Binding, 2008b; Tudhope, Binding y May, 2008; May <i>et al.</i> , 2009; Binding <i>et al.</i> , 2012	CSV	Aplicación en C	EH2SKOS
Aria	Van Gendt <i>et al.</i> , 2006			
Australian Public Affairs Information Service	Miles, Rogers y Beckett, 2004	XML	Transformación XSL	
Bibliopolis Thesaurus	Tordai, Omelayenko y Schreiber, 2007			
Bibliotheca Iuris Antiqui Thesaurus	Spampinato y Zangara, 2013			
Building materials Thesaurus	Tudhope y Binding, 2008a; Tudhope y Binding, 2008b; Tudhope, Binding y May, 2008; May <i>et al.</i> , 2009; Binding <i>et al.</i> , 2012	CSV	Aplicación en C	EH2SKOS
Camps	Rodriguez y Priddy, 2011	XSLX	Script en Python	Skosify
Charda	Sathn, 2015	XSLX	Script en Python	
Chinese Agricultural Thesaurus	Xian Guo jian <i>et al.</i> , 2012			

Nombre del vocabulario	Referencias que tratan el vocabulario	Extensión de archivo	Vía de transformación	Herramientas utilizadas
Common Procurement Vocabulary	Álvarez Rodríguez, Rubiera Azcona y Polo Paredes, 2007	CSV	Transformación XSL	Procesador XSLTProc
Common Procurement Vocabulary	Polo Paredes, Álvarez Rodríguez y Rubiera Azcona, 2008	CSV	Transformación XSL	
Dottrina Giurudica Classificazione materie giuridiche	Agnoloni <i>et al.</i> , 2013	XML		
e.Jewish.info Thesaurus	Judaica Europeana, 2008			
EHRI Thesaurus	Rodriguez y Priddy, 2011		Script en Python	Skosify
English Heritage Aircraft Type Thesaurus	Miles, Rogers y Beckett, 2004	CSV	Aplicación en Java	Jena 2
Environmental Application Reference Thesaurus	Martino y Albertoni, 2011	FDB	Lenguaje D2RQ	
European Commission Library Thesaurus	Francesconi <i>et al.</i> , 2007	XML	Transformación XSL	
European nature information system species	Martino y Albertoni, 2011	MDB	Lenguaje D2RQ	
European Training Thesaurus	Francesconi <i>et al.</i> , 2007	XML	Transformación XSL	
EUROVOC Thesaurus	Francesconi <i>et al.</i> , 2007	XML	Transformación XSL	
EUROVOC Thesaurus	Polo Paredes, Álvarez Rodríguez y Rubiera Azcona, 2008	XML	Transformación XSL	
Evidence Thesaurus	Tudhope y Binding, 2008a; Tudhope y Binding, 2008b; Tudhope, Binding y May, 2008; May <i>et al.</i> , 2009; Binding <i>et al.</i> , 2012	CSV	Aplicación en C	EH2SKOS
Federal Register Thesaurus for Indexing Terms	Casellas, 2012	TXT	Aplicación en Java	

Nombre del vocabulario	Referencias que tratan el vocabulario	Extensión de archivo	Vía de transformación	Herramientas utilizadas
Gemeenschappelijke Thesaurus Audiovisuele Archieven	Van Assem <i>et al.</i> , 2006; Malaise y CHOICE Team, 2007	Texto	Script en Perl	
General Management Environmental Thesaurus	Miles, Rogers y Beckett, 2004; Francesconi <i>et al.</i> , 2007	XML	Transformación XSL	
Ghettos	Rodriguez y Priddy, 2011	XSLX	Script en Python	Skosify
Habitat Types	Martino y Albertoni, 2011	MDB	Lenguaje D2RQ	
HASSET Thesaurus	Bell, 2013		Aplicación de consola	Clase "skoshas-segenerator"
Iconclass	Van Gendt <i>et al.</i> , 2006			
iDAI.gazetteer	Sathn, 2015	JSON	Script en Python	
Integrated Public-Sector Vocabulary	Van Assem <i>et al.</i> , 2006	XML	Aplicación en Prolog	SWI Prolog
International Press Telecom. Council News Code	Troncy, 2008			
International Standard for country codes ISO 3166	Voss, 2007			
Internat. Union for Conservation of Nature Class.	Martino y Albertoni, 2011	PDF	Lenguaje D2RQ	
Lib. of Congress Subj. Headings	Harper, 2006	MARC21	Transformación XSL	
Lib. of Congress Subj. Headings	Summers <i>et al.</i> , 2008	MARCXML	Script en Python	RDFlib
Lista de encabezamientos de materia para bibliotecas públicas	Martínez-Conde, 2013	MARC21		

Nombre del vocabulario	Referencias que tratan el vocabulario	Extensión de archivo	Vía de transformación	Herramientas utilizadas
Mathematics Subject Classification	Ion y Sperber, 2010; Lange <i>et al.</i> , 2012	TeX	Script en Perl	
MDA Object Type Thesaurus	Tudhope y Binding, 2008a; Tudhope y Binding, 2008b; Tudhope, Binding y May, 2008; May <i>et al.</i> , 2009; Binding <i>et al.</i> , 2012	CSV	Aplicación en C	EH2SKOS
Medical Subject Headings	Van Assem <i>et al.</i> , 2006	XML	Aplicación en Prolog	SWI Prolog
Nature 2000 AI Taxonomy	Martino y Albertoni, 2011	MDB	Lenguaje D2RQ	
Nuovo Soggetario Thesaurus	Bergamin y Lucarelli, 2013			
Organic Edumet Ontology	Rajabi, Sicilia y Sánchez-Alonso, 2013	OWL	Lenguaje D2RQ	
Ra Thesaurus	Felicetti <i>et al.</i> , 2016	TSV	Aplicación de consola	Stellar Console
Roman thesaurus	Sathn, 2015	MARC21	Script en Python	
Russian Library Bibliographic Classification	Larrenova y Pavlov, 2017			
Standard Thesaurus Wirtschaft	Neubert, 2009	Texto	Aplicaciones en Perl y C	RDF Helper, Redland RDF libraries
TAXREF	Callou <i>et al.</i> , 2015	TSV	Lenguaje xR2RML	Morph-xR2RML
Tesaurus del Habitat del CEHAP	Guzmán Luna, Torres Pardo y López García, 2006			
Thesaurus for the Social Sciences	Zapilko y Sure, 2009; Zapilko <i>et al.</i> , 2013	XML	Transformación XSL	

Nombre del vocabulario	Referencias que tratan el vocabulario	Extensión de archivo	Vía de transformación	Herramientas utilizadas
Thesaurus of Greek Terms	Solomon y Papatheodorou, 2010	XML	Transformación XSL	
Thesaurus of Monument Types	Tudhope y Binding, 2008a; Tudhope y Binding, 2008b; Tudhope, Binding y May, 2008; May <i>et al.</i> , 2009; Binding <i>et al.</i> , 2012	CSV	Aplicación en C	EHSKOS
Thésaurus pour la description et l'indexation des archives locales	Grimouard, 2014			
Thesaurus VRT	De Boer <i>et al.</i> , 2016	XML	Script en Python	Skosify
Trilingual Cultural Thesaurus	Akbari-Daryan, Khosravi y Ebrahimi, 2016	MARC		
UNESCO Nomenclature	Pastor-Sánchez <i>et al.</i> , 2013	PDF	Script en PHP	
UNESCO Thesaurus	Pastor-Sánchez <i>et al.</i> , 2013	HTML	Script en PHP	
UNESCO Thesaurus	Francesconi <i>et al.</i> , 2007	XML	Transformación XSL	
Universal Decimal Classification	Slavic y Balikova, 2011			
Yleinen suomalainen asiasanasto	Tuominen <i>et al.</i> , 2009	MARCXML	Aplicación en Java	Apache Jena
Zenon thesaurus	Romanello, 2012; Beer <i>et al.</i> , 2014	MARC21	Aplicación de consola	Stellar Console

Anexo 2 Elementos de SKOS con mapeo definido a elementos de cada vocabulario

	skos:affLabel	skos:broader	skos:broaderTransitive	skos:changeNote	skos:Collection	skos:Concept	skos:ConceptScheme	skos:definition	skos:editorialNote	skos:example	skos:hasTopConcept	skos:hiddenLabel	skos:historyNote	skos:inScheme	Skos-XI:literalForm	skos:member	skos:narrower	skos:notation	skos:note	skos:prefLabel	skos:related	skos:scopeNote	skos:topConceptOf
AGROVOC Thesaurus																							
Amsterdam Museum Thesaurus																							
Archaeolog. Sciences Thesaurus																							
Aria Catalogue																							
Australian Public Affairs Information Service	•	•					•							•						•	•	•	
Bibliopolis Thesaurus	•	•						•									•			•	•		
Bibliotheca Iuris Antiqui Thes.	•	•															•			•	•		
Building materials Thesaurus																							
Camps																							
Chardra	•	•										•		•						•	•		•
Chinese Agricultural Thesaurus	•	•												•						•	•		•

	skos:altLabel	skos:broader	skos:broaderTransitive	skos:changeNote	skos:collection	skos:Concept	skos:ConceptScheme	skos:definition	skos:editorialNote	skos:example	skos:hasTopConcept	skos:hiddenLabel	skos:historyNote	skos:inScheme	Skos-xl:literalForm	skos:member	skos:narrower	skos:notation	skos:note	skos:prefLabel	skos:related	skos:scopeNote	skos:topConceptOf
Library Congress Subject Head.																							
Lista de encabezamientos de materias para bibliotecas publicas																							
Mathematics Subject Class.																							
MDA Object Types Thesaurus																							
Medical Subject Headings																							
Nature 2000 A1 Taxonomy																							
Nuovo Soggetario Thesaurus																							
Organic EduNet Ontology																							
Ra Thesaurus																							
Roman thesaurus																							
Russian Library Bibliog. Class.																							
Standard Thesaurus Wirtschaft																							
TAXREF																							
Tesaurus del Habitat del CEHP																							
Thes. for the Social Sciences																							
Thesaurus of Greek Terms																							
Thesaurus of Monument Types																							

