

ISSN: 0718 – 1701



UTEM

Serie Bibliotecología y Gestión de Información N° 57, Mayo 2010

**Tesauros y Web Semántica: Diseño metodológico para
estructurar contenidos Web mediante SKOS-Core**

**Alonso Cavieres Abarca
Sergio Fredes Mena
Arturo Ramírez Novoa**



D · G · I

Departamento
de Gestión de
Información
Escuela de
Bibliotecología

Serie Bibliotecología y Gestión de Información es publicada desde Octubre de 2005 por el Departamento de Gestión de Información de la Universidad Tecnológica Metropolitana. Dr. Hernán Alessandri, 722, 6° piso, Providencia, Santiago, Chile, www.utem.cl

Sus artículos están disponibles en versión electrónica en E-prints in Library and Information Science: <http://eprints.rclis.org> y están indizados e integrados en la base de datos "Fuente Académica" de EBSCO Information Services.

Está registrada en el Sistema Regional de Información en Línea para Revistas Científicas de América Latina, El Caribe, España y Portugal (LATINDEX) y en Dialnet, portal de difusión de la producción científica hispana.

Sitio Web: <http://www.bibliotecarios.cl/servicios/serie-bibliotecologia-y-gestion-de-informacion/>

Dirección Editorial

- Héctor Gómez Fuentes, Director Departamento de Gestión de la Información
- Carmen Pérez Ormeño, Directora Escuela de Bibliotecología

Editor Jefe

Héctor Gómez Fuentes

Consejo Editorial

Académicos del Departamento de Gestión de Información

- Mariela Ferrada Cubillos
- Haydée Gutiérrez Vilches
- Guillermo Toro Araneda

Presidenta del Colegio de Bibliotecarios de Chile A. G.

Paola Roncatti Galdames

Representante Legal

Luis Pinto Faverio

Decano Facultad de Administración y Economía

Enrique Maturana Lizardi

Secretaria del Departamento de Gestión de Información

Rossana Flores Cuevas

Resumen en Inglés: Gemita Flores Cortés

Autorizada su reproducción con mención de la fuente.

LAS IDEAS Y OPINIONES CONTENIDAS EN LOS TRABAJOS Y ARTÍCULOS SON DE RESPONSABILIDAD EXCLUSIVA DE LOS AUTORES Y NO EXPRESAN NECESARIAMENTE EL PUNTO DE VISTA DE LA UNIVERSIDAD TECNOLÓGICA METROPOLITANA.

Tesauros y Web Semántica: Diseño metodológico para estructurar contenidos Web mediante SKOS-Core

Alonso Cavieres Abarca¹
aloncavier@gmail.com
Sergio Fredes Mena
Arturo Ramírez Novoa
Rafael Castillo Guerrero²
Héctor Gómez Fuentes

Resumen

Se revisan conceptos básicos sobre construcción de tesauros y aplicación de modelos de marcado en tesauros electrónicos. Se desarrolla un tesoro simple en base al contenido temático del sitio web "Chile para Niños", mediante el software ThManager y se transforma a SKOS-Core para una futura recuperación de información semántica.

Palabras claves: Web Semántica, Tesauros, Ontologías, Skos-Core, , Topic maps

Abstract

Basic concepts about thesauri construction and markup models for electronic thesauri are reviewed. Based on the analysis of the thematic content of the website "Chile para Niños" by means of the software ThManager, we developed a simple thesaurus and turned it into SKOS-Core for future semantic information retrieval.

Keywords: Semantic Web, thesauri, ontologies, SKOS-Core, Topic maps.

¹ Estudiantes de Seminario de Investigación de la Carrera de Bibliotecología y Documentación, Universidad Tecnológica Metropolitana.

² Profesores guía del Seminario de Investigación, Escuela de Bibliotecología de la Universidad Tecnológica Metropolitana.

TABLA DE CONTENIDO

INTRODUCCIÓN	6
1. Tesoros	7
1.1. Fundamentos conceptuales	7
1.2. Construcción de Tesoros	8
1.3 Automatización de Tesoros.....	15
1.4 Tesoros y Web semántica.....	19
2. Modelos de marcado para presentación de Tesoros en Internet.....	21
2.1 RDF	21
2.2 Topic Maps (TM).....	22
2.3 OWL	24
2.4 SKOS.....	24
2.5 ZTHEZ	25
2.6 RDF/XML.....	26
3. SKOS: Organización del conocimiento en la Web Semántica	28
3.1 SKOS: Contexto ontológico y web Semántica.....	29
3.2 Vinculación entre Tesoros y SKOS-CORE	30
3.3 Descripción de las propiedades y la relación definidas en SKOS-CORE	32
3.4 Ejemplificación de etiquetado SKOS-CORE.....	35
3.5 SKOS-CORE frente a Thes y Topic Maps.....	36
4.- Diseño metodológico para estructurar contenidos Web mediante SKOS-CORE	37
4.1 Sitio Web Chile para niños	37
4.2 Descripción del modelo por capas	39
4.3 Esquema de datos de red de términos.....	43

4.4 Implementación de software para creación de tesauro con etiquetado SKOS.....	46
4.5 Metodología de utilización	48
CONCLUSIONES	57
BIBLIOGRAFIA CONSULTADA.....	58

INTRODUCCIÓN

Organizar la información es una necesidad ante un mundo globalizado, donde los contenidos publicados en internet abundan y se multiplican de las más variadas formas. Estructurar los contenidos facilita una mayor comprensión de lo que se ve y se consulta en la web. Por esto resulta extraño que grandes sitios web clasifiquen su información con el mínimo interés y dedicación que esto debiera tener, principalmente porque los usuarios son los que finalmente agradecerán ese valor agregado y volverán a visitarlo.

Con esta afirmación, la propuesta de realizar un diseño metodológico para esquematizar una red de conceptos, tomando algunos de los preceptos de la Web Semántica, es una oportunidad para aprovechar aquellas potencialidades que aun puede proveer el uso de la información que se hace en internet.

Esta investigación utiliza este modelo en base a los sitios temáticos de la página "Chile para Niños", donde la información adquiere un valor significativo para hacer llegar a los pequeños los contenidos de manera sencilla y provechosa. Por ello, no es menor que las relaciones posibles entre las distintas temáticas que este sitio proporciona, estén sistematizadas de una manera coherente y útil con los objetivos que ella persigue, y ante esto, la aplicación práctica del modelo, lo cual conlleva la sistematización de los descriptores temáticos en complejas relaciones semánticas del tipo tesoro, contribuyan a objetivos básicos de la Arquitectura de información, como lo son la usabilidad, navegabilidad y la encontrabilidad de un Web en particular.

1. Tesoros

1.1 Fundamentos conceptuales

Un tesoro se define como un vocabulario controlado y estructurado formalmente, integrado por términos que guardan entre sí relaciones semánticas y genéricas: de equivalencia, jerárquicas y asociativas. Se trata de un instrumento de control terminológico que permite convertir el lenguaje natural de los documentos en un lenguaje controlado, ya que representa, de manera unívoca, el contenido de estos, con el fin de servir tanto para la indización, como para la recuperación de los documentos.

También se refiere a un léxico en orden alfabético de los términos que comprende un vocabulario especializado en una disciplina académica o campos de estudio, el cual muestra las relaciones lógicas y semánticas entre los términos, en particular una lista de encabezamientos de materia o descriptores utilizados como condiciones preferentes en la indización de la literatura en el área temática³. En recuperación de información (Information Retrieval) es un dispositivo de recall de naturaleza post coordinada⁴, el cual puede ser usado para localizar los términos más amplios y/o relacionados, si es que el usuario desea ampliar la recuperación, o términos específicos para delimitar una búsqueda más exacta. Un sistema bien diseñado de tesoro también permite al indizador mantener la coherencia y consistencia en la asignación de términos a los documentos, con lo cual se obtiene un control por sobre los puntos de acceso.

³ Thesaurus en: ODLIS Glosario [en: Línea]. [Consulta:4 de oct. 2009]. Disponible en: http://u.com/odlis/odlis_t.cfm

⁴ Iglesias Maturana, María Texia. Análisis y Recuperación de Información Documental. Santiago de Chile: IPS, 1992. ISBN: 956-7026-03-3

Algunos ejemplos prácticos de tesauros son los siguientes:

- a) Art & Architecture Thesaurus (AAT)
- b) GeoRef Thesaurus
- c) Legislative Indexing Terms: The CRS Thesaurus
- d) Thesaurus of ERIC Descriptors
- e) Thesaurus of Psychological Index Terms
- f) Thesaurus of Sociological Indexing Terms

1.2 Construcción de tesauros

Para el desarrollo y construcción de un tesoro será necesario conocer los aspectos reglamentarios básicos que orientan y ayudan a la creación del mismo, para construir un tesoro monolingüe es necesario conocer la Norma **UNE 50106-901**⁵, es en esta norma donde se establecen todos los aspectos que deberán ser considerados para la elaboración del tesoro de estas características. También debemos conocer la norma **UNE 50125**, para la construcción de un tesoro multilingüe.

Si bien esta investigación se basará en el estudio y posterior desarrollo de un tesoro monolingüe no está de más conocer los aspectos que rigen al tesoro multilingüe.

Ya hemos visto qué es un tesoro, para qué sirven y en qué contextos pueden aplicarse funcionalmente, pero, ¿Qué debemos tener en cuenta a la hora de desarrollarlo? Todo tesoro es en esencia un listado de palabras o términos empleados para representar conceptos, temas, materias o contenidos de los documentos, para realizar una normalización de los términos que permita mejorar el acceso de los usuarios a la

⁵ Norma UNE 50106-90, fue escrita y publicada en el año 1990, su equivalente en ISO es ISO 2788, escrita y publicada en 1986.

Norma UNE 50106-90 [en: Línea]. [Consulta:4 de oct. 2009]. Disponible en: http://bibliologia.info/archivos/une50106_90.pdf

información ubicada en distintos espacios sean estos físicos o virtuales. Es sabido que las relaciones establecidas o que pueden estar presentes para el desarrollo del tesoro se dividen en tres tipos: jerárquicas, asociativas y de equivalencia; además, sabemos que los tesauros establecen descriptores, que no son más que palabras, términos o expresiones escogidos entre un conjunto de sinónimos para representar un concepto susceptible de aparecer con cierta frecuencia en documentos indizables y en las distintas consultas realizadas por los usuarios. Generalmente corresponde a la etiqueta de un concepto y es la unidad básica de un tesoro.

Aspectos a considerar para la elaboración de un tesoro

Tomaremos como referencia principal la Norma UNE 50106-90, pues nuestro trabajo de investigación y esta norma abordan el tesoro Monolingüe como eje central, esta norma no pretende ser una regla que se deba seguir completamente, es más bien una guía de apoyo para esta tarea, las técnicas que se describen en esta norma no están limitadas a un determinado método de indización sea este pre o post coordinado.

Abreviaturas

Los tesauros establecen distintas abreviaturas para facilitar la estructura de los mismos, las siguientes son las distintas abreviaturas y sus significados:

- a) **NA** Nota de aplicación: se adjunta a algunos términos para precisar su significado dentro de un determinado lenguaje de indización.
- b) **USE** Usado por: el término que sigue a este símbolo es el término preferente que se elige entre varios términos sinónimos o cuasi sinónimos.
- c) **UP** Usado por: el término que sigue a este símbolo es un sinónimo o cuasi sinónimo no preferente.

- d) **TC** Término cabecera: se aplica a los términos que identifican el nombre de la clase más amplia a la que pertenece el concepto específico; se emplea algunas veces en la sección alfabética de los tesauros.
- e) **TG** Término genérico: se aplica a términos de significado más amplio.
- f) **TGG** Término genérico (genérico)
- g) **TGP** Término genérico (partitivo)
- h) **TE** Término específico: este símbolo precede al término referido a un concepto de significado más específico.
- i) **TEG** Término específico (genérico)
- j) **TEP** Término específico (partitivo)
- k) **TR** Término relacionado: el término que sigue se asocia a otro del que no es genérico, ni específico, ni sinónimo, ni cuasi sinónimo.

Las distintas relaciones distribuyen estas abreviaturas de la siguiente manera:

Relación de equivalencia

USE precede al término preferente
UP precede al término no preferente

Relación jerárquica

CS cabecera de serie de una jerarquía
TG término genérico

Puede lograrse una mayor distinción utilizando:

TGG Término genérico (genérico)
TGP Término genérico (partitivo)
TE Término específico o más restringido

Puede lograrse una mayor distinción utilizando:

TEG Término específico (genérico)
TEP Término específico (partitivo)

Relación asociativa

TR Término relacionado

Estas abreviaturas permiten al usuario del tesauro orientar su búsqueda entre términos (Descriptor) estructurados y relacionados entre sí, lo que permite al usuario localizar la información deseada de manera fácil, además, en base a las abreviaturas podemos inferir que no se deja al usuario solo con un término, sino que se le ofrecen distintas opciones a su búsqueda a través de los tres tipos de relaciones.

Relaciones básicas en el tesauro

a) La relación de equivalencia

Esta es la relación que se establece entre los términos preferente y no preferente (autorizado y no autorizado) cuando se considera, a efectos de la indización, que uno o más términos se refieren al mismo concepto.

b) La relación jerárquica

Esta es la relación básica que marca la diferencia entre un tesauro y una lista no estructurada de términos, como un glosario o diccionario. Se basa en grados o niveles de súper y subordinación, en que un término súper ordenado representa un todo o clase y los términos subordinados corresponden a sus miembros o partes.

c) La relación asociativa

Esta relación básica se define con más facilidad en términos de negación que en términos positivos. Incumbe a las relaciones que se establecen entre términos que no son equivalentes y que no pueden relacionarse tampoco jerárquicamente, aunque se asocien mentalmente hasta el punto de que la conexión entre ellos debe hacerse explícita en un tesauro, sobre la base de que tal conexión podría revelar términos alternativos que podrían ser útiles en la indización o la recuperación.

Estructura Básica

Veamos un ejemplo práctico de un descriptor desplegado correctamente:

Lo primero será definir el descriptor que se tomará como término aceptado, para este ejemplo deberemos optar entre aviones o aeroplanos, el término aceptado será el que más se acomode y sea el más usado por una comunidad, no tiene sentido definir descriptores que jamás serán buscados o utilizados por una comunidad de usuarios. En Centro America, por ejemplo, a los automóviles se les llama “Carros”, en España se les llama “Coche”, cada descriptor deberá asociarse a la comunidad para la que esta dirigida (por lo general los tesauros son especializados en alguna materia o área en particular).

Tomaremos para este ejemplo el término “Aviones” como término aceptado, por ser este término el más usado en nuestra (supuesta) comunidad de usuarios, lo siguiente será definir las relaciones que este término tiene, para esto utilizaremos las abreviaturas descritas anteriormente:

AEROPLANO

USE: avión
TR: aeropuerto
TR: avioneta

El término de cabecera deberá estar escrito en mayúsculas, los términos que acompañan las abreviaturas deberán escribirse en minúsculas, excepcionando los términos de nombres personales o acrónimos.

En los tesauros podremos encontrar dos métodos para controlar los el vocabulario:

- a) El alcance de los términos se restringe deliberadamente a significados seleccionados, a diferencia de los diccionarios, en los que cada

término puede ir acompañado de diferentes acepciones según el uso común, en un tesoro cada término se ajusta a un único significado, aquel que con mayor eficacia sirve las necesidades del sistema de indización. La estructura de un tesoro, especialmente la presentación de relaciones jerárquicas, indica con frecuencia el significado de un término. Si esto no resulta suficientemente explícito, puede añadirse una definición o nota de aplicación al término. La nota debe establecer el significado elegido y puede incluir también otros significados que se le atribuyen en lenguaje natural pero que deliberadamente se excluyen a efectos de la indización;

- b) cuando el mismo concepto puede expresarse mediante dos o más sinónimos, uno de ellos se elige como término preferente y se emplea siempre en la indización. Debe hacerse un reenvío desde cualquier sinónimo que pudiera servir al usuario como punto de acceso hacia el término preferente.

Organización del tesoro

La construcción de un tesoro no es algo fácil, por lo general es conveniente determinar el formato del tesoro antes de comenzar a recopilar los términos que serán incluidos en el tesoro, es decir, antes de comenzar a elaborar el tesoro, esto se extrapola a otras áreas, como al diseño, donde se deberá tener, primero, claro lo que se desea hacer. Existen distintos métodos para definir el formato de un tesoro:

- a) Método deductivo

Este método plantea la idea de elaborar el tesoro a medida que se vayan analizando los documentos que pertenecen al campo documental concreto que se quiere indexar y utilizando el mismo vocabulario que se encuentra en los documentos, a través de una extracción automática de palabras o términos que más se repitan en el lenguaje natural. Con este método podremos elaborar el tesoro y analizar a la vez los documentos,

pero constituye un vocabulario creciente en forma desordenada y difícil para relacionar términos.

b) Método Inductivo

Este método consiste en extraer todas las palabras que corresponden a un área documental determinada, partiendo por diccionarios, listas de clasificación, tablas de materias, etc. Este método facilita la elaboración de las relaciones semánticas, pero presenta la dificultad de no cubrir por completo el campo interesado.

c) Método Mixto

La idea es combinar los dos métodos anteriores, de esta forma obtendremos un tesoro específicamente adaptado a las necesidades del usuario.

Formato y contenido de un tesoro

Es difícil establecer una presentación estándar para los tesoros impresos, más difícil se hace esto con tesoros electrónicos, pero básicamente un tesoro debe contener y reflejar con claridad los siguientes aspectos:

- Portada,
- Sumario,
- Introducción,
- Presentación Gráfica, los términos deberán estar dispuestos como una figura que permita al usuario conectar los términos que se encuentran relacionados. Con dos tipos de presentación: árboles y flechados.
- Sección alfabética.

Los aspectos propios del tesoro deberán reflejar los siguientes aspectos:

- a) Objetivo del tesoro,
- b) Campo temático que cubre

1.3 Automatización de tesauros

La automatización y el desarrollo de la Web han ido evidenciando, por un lado, la adaptación de los tesauros a nuevos usos, integrándose, en conjunción con el lenguaje natural, en los Sistemas de Gestión de la Información (SGD) y en los sistemas de búsqueda en la red y, por otro, sus posibilidades hipertextuales a través de las relaciones para servir como elementos de navegación. En ese contexto, sería ideal que el tesoro automático debiese permitir al usuario la libre elección del tipo de relaciones que quiera establecer entre los términos.

Al momento de evaluar la calidad de un tesoro automatizado, los términos y las relaciones son aspectos que no deben dejarse de lado. En relación con los términos, cualquier tesoro automático debe ayudar al usuario a crear y gestionar la terminología adecuada a sus propósitos.

Al momento de automatizar, los requisitos básicos para cualquier software de automatización son, además del campo para el término, uno destinado a notas de alcance que definan su sentido y otro dedicado a anotaciones para facilitar la estructuración del vocabulario y su presentación sistemática. Algunos tesauros automáticos restringen la longitud del campo término, lo que puede resultar problemático por dos razones: cuanto más específico es el vocabulario hay una mayor probabilidad de que existan términos compuestos o nombres propios que pueden tener una longitud considerable y por el intercambio de datos, el que sería impracticable cuando el tesoro, fuente del cual se quieren importar términos, tenga una mayor longitud de campo.

Además de estos aspectos estructurales existen algunas rutinas de trabajo típicas de un tesoro automático como la captura de datos, la modificación o eliminación de términos y sus relaciones, o el control de la consistencia del vocabulario y sus relaciones.

No hay que olvidar que la elaboración de un tesoro es una operación intelectual, en la cual la automatización sólo puede aportar una ayuda para las tareas de carácter más rutinario y repetitivo. Su elaboración supone el trabajo continuo de un grupo de especialistas en la materia durante un periodo de tiempo bastante amplio, que dependerá del volumen del tesoro, de los campos a cubrir y de la profundidad de la indización.

Software para la construcción de Tesoros

En este apartado se podrá identificar algunas experiencias en la automatización de tesoros a partir de software creado especialmente para la construcción de estos.

- a) **B.E.A.T.** (Buen entorno para la administración de tesoros). BEAT es un programa desarrollado por Josep Sau, del Centre d'Informatica de la Universitat de Barcelona, para la elaboración, mantenimiento y consulta de thesaurus documentales sobre microordenadores compatibles PC, con sistema operativo ms-dos. Fue desarrollado en 1990. Cuenta con versiones en castellano, catalán e inglés. Con Beat puede crearse y actualizarse interactivamente un thesaurus de hasta 65.000 términos (manteniendo la misma velocidad de trabajo) con sus relaciones semánticas: equivalencia (USE/UF), jerarquía (BT/NT) o asociación (RT), así como notas de alcance o explicativas. Todo ello según la norma ISO 2788 (1979). Las relaciones inversas se crean automáticamente. El sistema normaliza los términos entrados para evitar duplicidades y controla la coherencia de las relaciones creadas.

- b) **Lexico:** Es un sistema de gestión diseñado expresamente para soportar la creación, mantenimiento y edición de vocabularios automáticos. Está basado en Java. Su interfaz es un navegador web estándar, con lo que se puede acceder y manipular el tesoro desde el www. Además los campos de nota descriptiva permiten acceder a información útil sobre el término: instrucciones de catalogación y recuperación, información histórica, etc. Pueden añadirse términos y relaciones provisionales pero permanecen temporalmente hasta que son aprobados. Empleado en los tesoros elaborados por la Library of Congress.

- c) **Multites:** Software específico que posibilita y optimiza el tratamiento automatizado de los términos así como la posterior edición electrónica del tesoro. Cuenta con un soporte de tesoros monolingües y multilingües, además de tesoros polijerárquicos. Tiene un soporte de las relaciones normalizadas más las establecidas por el usuario. Cuando se crea un nuevo tesoro, se definen automáticamente las relaciones y notas normalizadas. El sistema permite el cambio de estas relaciones y la definición de las que el usuario considere necesarias. Puede tener cuatro pares de relaciones de cada tipo (USE/UF sería un par. Es empleado, entre otros, en el Canadian Literacy Thesaurus, WorldBank Thesaurus.

- d) **SGAT:** Sistema Gestor de Automatización de Tesoros desarrollado por Juan Antonio Pastor Sánchez, Universidad de Murcia y utilizado para la construcción del Tesoro de Redes de Ordenadores.

- e) **THESmain:** Su principal aplicación es la visualización y mantenimiento del tesoro GEMET (General European Multilingual Environmental Thesaurus), redactado en 12 lenguas comunitarias.

- f) **Web Thesaurus Navigator:** Utilizado en el tesoro de ASIS

- g) **Wintertree Software Thesaurus:** Desarrollado de una aplicación específica para integrar un tesoro en aplicaciones Java.

EJEMPLOS DE TESAUROS MONOLINGUES APLICADOS A LA WEB

a) En Español:

- Tesoro ISOC de Psicología: Desarrollado por el Centro de Información y Documentación Científica de España (CINDOC) Permite la búsqueda por conceptos y por jerarquía.
- Tesoro de Derecho, desarrollado por el CINDOC.
- Tesoro Propiedad Industrial, desarrollado por el CINDOC.
- Tesoro ISOC de Economía, desarrollado por el CINDOC.
- Tesoro de Biología animal, desarrollado por el CINDOC.
- Tesoro de Medio Ambiente, de Alejandro de la Cueva.
- Tesoro de redes de ordenadores, de Fco Javier Martínez Méndez y Juan Carlos García Gómez Departamento de Información y Documentación (Universidad de Murcia).

b) En Inglés:

- HASSET: Humanities and Social Science Electronic Thesaurus, de la Universidad de Essex.
- SOSIG: Social Science Information Gateway.
- Global Legal Information Network (GLIN), de la Library of Congress.
- Legislative Indexing Vocabulary (LIV), de la Library of Congress.
- ERIC, de la U.S. Department of Education.
- NASA Thesaurus, de la NASA.
- Bioethics Thesaurus 1999, de la National Reference Center for Bioethics Literature (Georgetown University)

- Thesaurus of Parasitology, de la Jolene M. Mille (Kent State University)
- Art & Architecture Thesaurus browser (AAT), de Getty Research Institute. Vocabulary Program.
- Thesaurus for Graphic Materials I: Subject Terms, de la Library of Congress.
- Thesaurus for Graphic Materials II: Genre and Physical Characteristic Terms, de la Library of Congress.
- Getty Thesaurus of Geographic Names (TGN), de la Getty Research Institute. Vocabulary Program.

1.4 Tesauros y Web Semántica

La Web Semántica se define como aquel modelo que proporciona un marco común para que los datos sean compartidos y reutilizados a través de aplicaciones, empresas y fronteras comunitarias. Es un esfuerzo colaborativo liderado por el W3C con la participación de un gran número de investigadores y socios industriales. Está basado en Resource Description Framework (RDF) e integra una variedad de aplicaciones utilizando XML para la sintaxis y URI para las denominaciones. Si bien el proyecto de la W3C es único, sus ideas fuerza apuntan hacia dos caminos distintos:

Inteligencia Artificial (AI)

La Web Semántica es un conjunto de iniciativas destinadas a promover una futura Web cuyas páginas estén organizadas, estructuradas y codificadas de tal manera que los ordenadores sean capaces de efectuar inferencias y razonar a partir de sus contenidos.

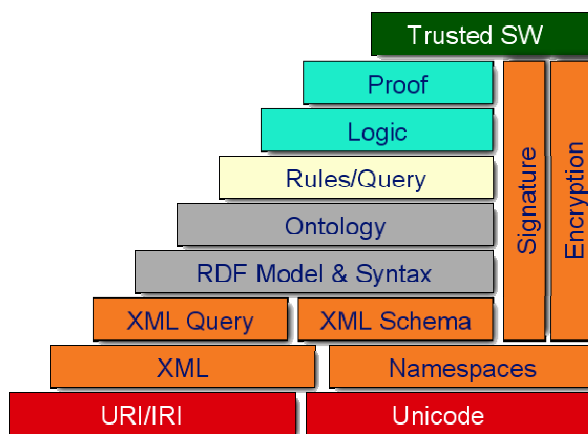
Procesamiento robusto

La Web Semántica es un conjunto de iniciativas destinadas a convertir la World Wide Web (WWW) en una gran base de datos capaz de soportar un procesamiento sistemático y consistente de la información.

La Web Semántica corresponde a la evolución o al estado siguiente de la Web sintáctica, es decir aquella Web en que el etiquetado de los contenidos vía HTML es el que le da la relevancia y pertinencia al recurso en cuestión y no su carácter semántico, inserto implícitamente en el argumento de búsqueda.

Infraestructura de la Web Semántica

Tim Berners-Lee, creador de la WWW, diagramó la infraestructura de la Web Semántica en 7 capas o niveles, como se muestra en el siguiente esquema:



Infraestructura de la web semántica

Fuente:

http://www.w3c.es/Presentaciones/2004/11-GiraEstandares-JA/bnes/sw_stack.png

1. Nivel de codificación de las páginas.
2. Nivel estándar con carga semántica.
3. Nivel, categorías de paginas interoperables RDF
4. Nivel sistema de ontologías OWL

Se puede apreciar que pasar por estas capas, pretende establecer confianza (Trust); esto se gana en la medida en que se incrementa el nivel, para así llegar a un nivel óptimo de la información recuperada.

2. Modelos de marcado para representación de tesauros en internet

2.1 RDF

Es un modelo para la estructuración de metadatos, que además permite describir cualquier recurso que pueda asignársele un URI (Uniform Resource Identifier); también podemos asignarle URI a los términos de un tesauro. Por ello puede considerarse un esquema para representar lenguajes jerárquicos y mapas de conocimiento. Al igual que XML, se trata de un estándar desarrollado por el W3C. La semántica funcional de RDF está formada por: un modelo de datos, una sintaxis y un esquema. La especificación del modelo y la sintaxis, es un estándar estructural de metainformación diseñado para servir como fundamento para la interoperabilidad en el procesamiento de metadatos. La especificación del esquema proporciona los recursos suficientes para crear modelos RDF que representen la estructura lógica de un tesauro [W3C-RDF-S]. Sin embargo, lo más significativo de RDF es la utilización de los namespaces que permiten la utilización de vocabularios distribuidos. Utilizar RDF/XML para el desarrollo de tesauros o vocabularios controlados en el entorno de información distribuida de una Intranet, implica "utilizar metadatos para describir metadatos", aprovechar la flexibilidad de XML para gestionar el conocimiento corporativo tanto a nivel estructural como semántico, así como por la versatilidad de los enlaces y la posibilidad de compartir distintos tesauros. Este es el primer lenguaje de ontologías ligados con la iniciativa de la Web Semántica.

2.2. Topic Maps (TM):

Este estándar está definido para SGML (Standard Generalized Markup Language). Un topic map es un documento, o un conjunto de documentos SGML o XML (Extensible Markup Language), interrelacionados en un espacio multidimensional en el que las localizaciones son un topic. Tiene como finalidad normalizar los elementos y la notación utilizada para estructurar la información mediante la construcción de una red de enlaces

semánticos que relacionen diferentes recursos informativos. Los elementos principales de los Topic Maps son los topics, las occurrences y las associations. Un topic es el elemento principal de un TM, es el término que expresa determinado concepto o idea. Ejemplos de topics pueden ser "persona", "idea" o "continente". Occurrences son enlaces a recursos informativos, por ejemplo una página Web. Son elementos opcionales que son relevantes a un topic dado, es decir un topic puede tener cero o muchas occurrences. La association es un enlace que establece una relación entre dos o más topics. Una forma de verlo es tomar una frase y considerar los sustantivos de esas frases como topics empleando los términos de unión entre los topics para denominar la asociación. Por ejemplo "la resina está en los pinos".

Una de las características de los TM es que permiten la posibilidad de adaptarse a diferentes perfiles de usuarios. Esto se realiza por medio de dos tipos de filtrado: el "scope", que filtra diferentes topics y "themes" (temas) para una determinada comunidad, y las facetas (facets), que se ocupan de filtrar recursos informativos, por ejemplo, ofrecer a determinado usuario sólo los recursos informativos que estén en un idioma determinado. Para la creación de TM existen muchos y variados métodos que pueden simplificar su creación. Se puede partir del procesamiento del lenguaje natural mediante frases en lenguaje libre para luego identificar estructuras verbales que pueden asociarse de forma automática con determinados "association types". Sin embargo, este método presenta muchos problemas debido a la polisemia, riqueza del lenguaje y ambigüedad. Más práctico resulta, mediante las propiedades de las asociaciones, obtener nuevas asociaciones distintas a las originales, o utilizar la fusión de recursos.

Similitudes entre RDF y Topic Maps

Al igual que RDF, toman como punto de partida la necesidad de describir de forma precisa los recursos de información. Otra similitud es que tanto los **Topic Maps** como los documentos RDF actúan como mediadores entre un repositorio de información y los usuarios del sistema.

Diferencias con los Tesauros

Desde su aparición, los tesauros fueron productos centralizados, incorporados a una única organización, y cuya presentación en formato electrónico es sólo una ventaja añadida, pero no es requisito imprescindible. Mientras que los Topic Maps son productos descentralizados, pensados para mejorar la cooperación entre distintas organizaciones y cuya naturaleza exige tener formato electrónico.

En la práctica, los tesauros suelen estar contruidos manual o semi-manualmente. Sin embargo, en los Topic Maps, y debido al número de relaciones que pueden tener, el tratamiento manual es imposible, puesto que es potencialmente infinito.

Una de las razones que explican el éxito de los tesauros es su simplicidad conceptual, con una estructura básica sencilla y limitados tipos de relación. Mientras que en los Topic Maps, el propio motivo que los originó, en su deseo de facilitar la fusión de diferentes recursos, impuso una estructura flexible, con gran riqueza de relaciones.

2.3 OWL

OWL permite, al igual que RDFS, establecer una jerarquía de clases, al definir relaciones de clases y subclases; definir propiedades; establecer asociaciones de clases, mediante dominio y rango de propiedades; y crear instancias de clases. También presenta funcionalidades adicionales en comparación a RDFS. Con OWL es posible: establecer sinónimos, al

definir clases y propiedades equivalentes; imponer un mayor número de restricciones a los valores de propiedades; y definir axiomas mediante lógica descriptiva que permiten realizar inferencia de la información. OWL se origina producto de la fusión entre las iniciativas norteamericana DAML y la europea OIL, en base a RDF Schema (RDFS), llamándose DAML+OIL; para luego enviar la especificación al W3C para su estandarización, donde fue renombrado como se conoce ahora.

2.4 SKOS

Sigla de *Simple Knowledge Organization System*, es una iniciativa del W3C en forma de aplicación de RDF que proporciona un modelo para representar la estructura básica y el contenido de esquemas conceptuales como listas de encabezamientos de materia, taxonomías, esquemas de clasificación, tesauros y cualquier tipo de vocabulario controlado. El origen del proyecto se remonta a la elaboración de un tesauro de actividades dentro del proyecto SWAD-Europe. La primera versión de SKOS Core se presentó en el año 2003. En SKOS los conceptos se identifican con referencias URI. Estos conceptos pueden etiquetarse en cadenas de texto en uno o varios idiomas, documentarse y estructurarse a través de relaciones semánticas de diversa tipología. Este modelo permite mapear conceptos de diferentes esquemas, así como definir colecciones ordenadas y agrupaciones de conceptos. También permite establecer relaciones entre las etiquetas asociadas a los conceptos. El uso de RDF en el desarrollo de SKOS permite obtener documentos en un formato que permita su lectura por parte de aplicaciones informáticas, así como su intercambio y su publicación en la Web. SKOS se ha diseñado para crear nuevos sistemas de organización o migrar los ya existentes adaptándolos a su uso en la Web Semántica de forma fácil y rápida. Proporciona un vocabulario muy sencillo y un modelo intuitivo que puede ser utilizado conjuntamente con OWL o de forma independiente. Por todo ello, SKOS se considera como un paso intermedio, un puente entre el

caos resultante del bajo nivel de estructuración de la Web actual y el riguroso formalismo descriptivo de las ontologías definidas con OWL.

El modelo de datos SKOS es en realidad una ontología definida con OWL Full. SKOS puede ser utilizado conjuntamente con OWL para expresar formalmente estructuras de conocimiento sobre un dominio concreto ya que SKOS no puede realizar esta función al no tratarse de un lenguaje para la representación de conocimiento formal.

2.5 ZTHES

Describe un modelo abstracto para la representación y búsqueda de tesauros tal y como se describen en la norma ISO 2788. La idea fundamental que reside tras esta propuesta es la de proponer un modelo que permita la implementación de tesauros para que se pueda acceder a ellos mediante el protocolo Z39.50 y SRW. Esta propuesta, debido al auge de los lenguajes de marcado, no puede dejar de ofrecer una DTD para la representación del tesauro, si bien se trata fundamentalmente de una iniciativa centrada en el protocolo Z39.50 con los condicionamientos que esto le supone.

2.6 RDF/XML

La aparición de XML establece una nueva forma de estructurar la información, pero la aparición de RDF establece una nueva mirada a la creación de una infraestructura Web Semántica que entregue soporte a la información que existe en Internet. RDF es un tipo de lenguaje que ayuda a representar información sobre recursos que debe ser procesada por máquinas, sin que esto suponga una pérdida de significado. RDF se basa en la idea de que se pueden identificar los elementos a partir de **URIs**, describiendo los recursos en términos de propiedades simples o pares propiedad-valor o de tríos (conocidas en términos de RDF como tripletes), es decir, sujeto-predicado-objeto, donde “el sujeto es el

recurso”, es decir aquello que se está describiendo. El predicado es la propiedad o relación que se desea establecer acerca del recurso. Por último, el objeto es el valor de la propiedad o el otro recurso con el que se establece la relación. La combinación de RDF con otras herramientas como RDF Schema y OWL permite añadir significado a las páginas, y es una de las tecnologías esenciales de la Web Semántica⁶.

RDF en el desarrollo de tesauros

Resulta evidente y de perogrullo afirmar que en los sistemas de información electrónica (de acceso web o no) no tienen ningún sentido los tesauros tradicionales impresos. Contar con una herramienta de control terminológico resulta muy importante, gracias a estas herramientas de control terminológico se podrá contar con un tesoro que pueda adaptarse a las necesidades de escalabilidad e interoperabilidad de la Internet. En Estados Unidos existe un convenio generalizado (patrocinado esencialmente por las recomendaciones del W3C⁷) para la estructuración y mantenimiento de tesauros, ya que este lenguaje permite, entre otras cosas, utilizar los *namespaces* que permiten la utilización de vocabularios distribuidos.

Utilizar **RDF/XML** para el desarrollo y elaboración de tesauros o vocabularios controlados en base a la web, supone utilizar metadatos para describir recursos. Además permite aprovechar la flexibilidad de XML para manejar el conocimiento, ya sea a nivel estructural o semántico.

6 Resource Description Framework, 2009, disponible en:
http://es.wikipedia.org/wiki/Resource_Description_Framework
7 RDF Thesaurus. W3C. Disponible en:
<http://www.w3.org/2001/sw/Europe/reports/thes/rdftthes.html>

Ejemplo de RDF

A continuación se presenta un ejemplo del lenguaje RDF y página sobre Juan Soto Pérez:

```
<http://es.personajessinmedio.org/Juan_soto_pérez>  
<http://purl.org/dc/elements/1.1/title> "Juan Soto Pérez" .  
<http://es.personajessinmedio.org/Juan_soto_pérez>  
<http://purl.org/dc/elements/1.1/publisher> "personajes sin medio".  
El ejemplo puede ser expresado mediante RDF/XML de la siguiente  
manera:  
<rdf: RDF  
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
xmlns:dc="http://purl.org/dc/elements/1.1/">  
<rdf:Description  
rdf:about="es.personajessinmedio.org/Juan_soto_pérez">  
<dc:title>Juan Soto Pérez</dc:title>  
<dc:publisher>personajes sin medio</dc:publisher>  
</rdf:Description>  
</rdf:RDF>
```

3. SKOS: Organización del conocimiento en la Web Semántica

SKOS, Simple Knowledge Organization System o Sistema de Organización del Conocimiento Simple, se define como un vocabulario para la representación de sistemas de organización del conocimiento, comúnmente conocidos como KOS (Knowledge Organization Systems)⁸; tales como tesauros, taxonomías, esquemas de clasificación y listas de encabezamiento de materias. SKOS está basado en RDF (Resource Description Framework - RDF-PRIMER) por lo que dichas representaciones pueden ser legibles por máquinas e intercambiarse entre aplicaciones de software, así como publicarse en la World Wide Web.

SKOS ha sido diseñado para proporcionar un modo de migrar a la Web Semántica sistemas de organización del conocimiento ya existentes con un costo bajo. SKOS también proporciona un lenguaje conceptual de modelado muy sencillo e intuitivo para desarrollar y compartir nuevos sistemas de organización. Puede utilizarse por su cuenta, o en combinación con lenguajes más formales, como el Lenguaje de Ontologías Web (OWL). SKOS también puede contemplarse como una tecnología de transición que proporciona un nexo de unión entre el formalismo lógico riguroso de los lenguajes de ontologías como OWL y el mundo caótico, informal y débilmente estructurado de las herramientas colaborativas basadas en Web, ejemplificadas por las aplicaciones de etiquetado social o Folksonomias.

El objetivo de SKOS no es sustituir vocabularios conceptuales originales en su contexto inicial de uso, sino que puedan implementarse en un

⁸ SKOS-Core [En línea]. [Consultado: 8 de Noviembre, 2009]. Disponible en: <http://www.w3.org/2001/sw/Europe/events/200406-esp/trabajo-final-extratesauros/node6.html>

espacio compartido, basado en un modelo simplificado, que haga posible su reutilización y una mejor interoperabilidad.

3.1 SKOS: Contexto Ontológico y Web Semántica

Las ontologías han sido tradicionalmente usadas como modelo de representación del conocimiento en inteligencia artificial. Esta se define según (Gruber, 2006)⁹ como una especificación explícita y formal de una conceptualización. De manera más simple, una ontología ha de concebirse como un entendimiento común y compartido de un dominio, que puede comunicarse entre humanos y sistemas computacionales. Ésta última característica y el hecho de que puedan compartirse y reutilizarse en diferentes aplicaciones, explica en parte el gran interés suscitado en los últimos años en la creación e integración de ontologías.

Objetivos de una Ontología

- Compartir la comprensión común de la estructura de información entre personas o agentes del software,
- Permitir la reutilización del conocimiento perteneciente a un dominio,
- Permitir representar de manera explícita los supuestos de un dominio,
- Separar el conocimiento de un dominio, del conocimiento que se pueda denominar operacional y
- Hacer posible analizar el conocimiento de un campo.

De cara a la web y retomando la definición técnica de ontología dada anteriormente, se encuentra que la web es mucho más extensa que una base de conocimiento, por lo que es necesario el uso de lenguajes estándares que permitan implementar ontologías en este ambiente. Esto

⁹ GRUBER T. R. Toward principles for the design of ontologies used for knowledge sharing. En: Formal ontology in conceptual analysis and knowledge representation. The Netherlands: Kluwer Academic Publishers, 1993.

obligó el desarrollo de diversos trabajos destacándose en los últimos años los trabajos realizados por la Web Semántica, entre los que se encuentran el lenguaje RDF, que implementa un modelo básico para establecer propiedades sobre los recursos y el lenguaje de ontologías OWL.

3.2 Vinculación entre tesauros y SKOS-CORE

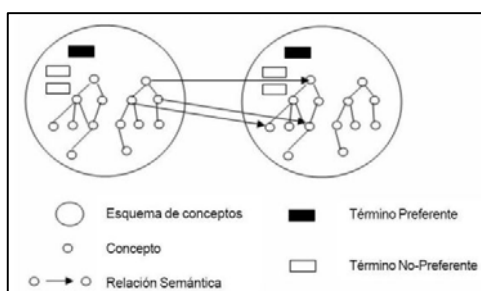
Los tesauros documentales, tal como se definió anteriormente, son instrumentos de control terminológico utilizados para convertir el lenguaje natural a un lenguaje normalizado. Se estructuran en listados de términos que tienen entre ellos relaciones genéricas, y que se aplican a un dominio concreto del conocimiento. Asociadas a los tesauros existen diferentes normas internacionales las cuales definen lineamientos para su construcción y mantenimiento.

En un tesoro existe manejo de relaciones entre los descriptores. Es importante resaltar tres tipos de relaciones básicas:

- a) Relaciones de equivalencia: es la relación entre descriptores o términos preferidos y los no descriptores o términos no usados en la indización referentes a un mismo concepto.
- b) Relación jerárquica: es la relación vertical entre todos los descriptores de una misma clase, expresada en términos de subordinación de los conceptos en un doble sentido.
- c) Relaciones de asociación: indican la relación o uniones en la significación de los descriptores (dos descriptores que son susceptibles de evocarse mutuamente por asociación de ideas). Todas las relaciones son recíprocas (A es a B lo que B es a A).

Si bien la aparición del XML marca el comienzo de una tendencia en la implementación de tesauros, la aparición del RDF supone un punto de inflexión en cuanto a la creación de una infraestructura semántica. Por ello no es raro suponer que la primera propuesta para la conversión de un tesoro se hubiera orientado a utilizar el RDF como mecanismo básico para su implementación, siendo luego utilizado el OWL, donde su versión OWL-LITE ha sido la más recomendada para transformar tesauros en el ambiente de la Web Semántica. Finalmente, una propuesta más concreta para la representación de tesauros documentales en la Web Semántica ha sido el SKOS-CORE. Además de ser una propuesta formal validada por el W3C, está pensada para ser un complemento al OWL, ya que proporciona un marco básico para la construcción de conceptos. Estas dos características hacen del SKOS-CORE una herramienta valiosa y útil para recopilar el conocimiento almacenado en un tesoro.

En el SKOS-CORE, la idea base del uso de un RDF Schema reside en su capacidad para permitir la definición de conceptos y esquemas de conceptos. Un concepto se define como una unidad de pensamiento que puede ser definida o descrita. A su vez un esquema de conceptos no es otra cosa que una colección de ellos.



Relaciones Básicas de SKOS-CORE
Fuente: [http://www.w3.org/...](http://www.w3.org/)

Las relaciones que se encuentran entre los descriptores de los tesauros tradicionales pueden ser precisadas entre los conceptos del esquema de conceptos, definidos en el SKOS-CORE, con la salvedad de que tales relaciones pasan a ser de carácter semántico.

3.3 Descripción de las propiedades y relaciones definidas en SKOS-CORE

Vocabulario SKOS

a) Conceptos y esquemas de conceptos

- skos:Concept
- skos:ConceptScheme
- skos:inScheme
- skos:hasTopConcept

b) Etiquetas léxicas

- skos:prefLabel
- skos:altLabel
- skos:hiddenLabel

c) Relaciones semánticas

- skos:semanticRelation
- skos:broaderTransitive
- skos:narrowerTransitive
- skos:broader
- skos:narrower
- skos:related

d) Documentación

- skos:note
- skos:scopeNote
- skos:historyNote
- skos:changeNote
- skos:definition
- skos:editorialNote

- skos:example

e) Colecciones de conceptos

- skos:Collection
- skos:OrderedCollection
- skos:member
- skos:memberList
- Propiedades de mapeado
- skos:mappingRelation
- skos:exactMatch
- skos:broadMatch
- skos:narrowerMatch
- skos:relatedMatch

f) Notaciones

- skos:notation

La codificación de las etiquetas correspondientes a los Términos Preferentes y No Preferentes pertenecientes a un Concepto, skos:Concept, se realiza mediante las propiedades skos:preflabel, para los términos preferentes y skos:altLabel, para los no preferentes o descriptores. Esta segunda etiqueta representa la relación de sinonimia o equivalencia del tesoro. La utilización de notas aclaratorias es otra necesidad que se representa con el skos:ScopeNote y las definiciones de conceptos se representan mediante skos:definition.

Adicionalmente, el SKOS-CORE permite representar las relaciones básicas entre los conceptos manejadas por la mayoría de los tesauros mediante la familia skos:semanticRelation, bajo la cual se encuentran las relaciones jerárquicas definidas por las propiedades skos:narrower y skos:broader, para términos específicos y generales respectivamente y skos:related para términos relacionados.

El SKOS-CORE también posee la forma de especificar otros tipos de propiedades que representan una forma de ampliar las relaciones semánticas entre conceptos, estas propiedades sobrepasan lo actualmente considerado en un tesauro tradicional, es así como aparecen: (a) skos:broaderGeneric y skos:narrowerGeneric, para definir únicamente relaciones de subfunción entre dos conceptos, donde un concepto es subclase de otro, (b) skos:broaderInstative y skos:narrowerInstative, donde Instative expresa que un concepto es una instancia de otro, y (c) skos:broaderPartitive y skos:narrowerPartitive, donde Partitive indica que un concepto es parte de otro (es una forma de especificar que un concepto es parte de otro).

Por último, este también permite expresar refinadamente la relación de asociación con las propiedades skos:relatedHasPart y skos:relatedPartOf, las cuales expresan que un concepto tiene parte de otro concepto o que un concepto esta relacionado con parte de otro concepto respectivamente.

Propiedad SKOS-Core	Tipo / Actua	Definición	Relación
Skos:prefLabel Skos:altLabel Skos:externalID	Terminos Terminos Conceptos	Preferente u Descriptur No-preferente o No-descriptor Identificadores para los conceptos	Equivalencia
Skos:copeNote Skos:definition	Conceptos Conceptos	Definición de concepto, notas aclaratorias Definición de concepto	
Relaciones Semánticas	Tipo / Actua	Definición	Relación
Skos:narrower Skos:broader Skos:related	Terminos Terminos Terminos	Terminos específicos Terminos generales Terminos relacionados	
Generic	Tipo / Actua	Hereda de subclassof	Subsuncion
Skos:broaderGeneric Skos:narrowerGeneric	Terminos Terminos	Herencia de terminos específicos Herencia de terminos generales	
Instative	Tipo / Actua	Hereda de type	Un concepto es instancia de otro
Skos:broaderInstative Skos:narrowerInstative	Conceptos Conceptos	Un concepto es instancia de un concepto específico Un concepto es instancia de un concepto general	Skos:relatedHasPart Skos:relatedOf

Descripción de las propiedades y relaciones definidas en SKOS-CORE

Fuente: Cuadro resumen generado [http://www.w3.org/...](http://www.w3.org/)

3.4 Ejemplificación de etiquetado SKOS-CORE

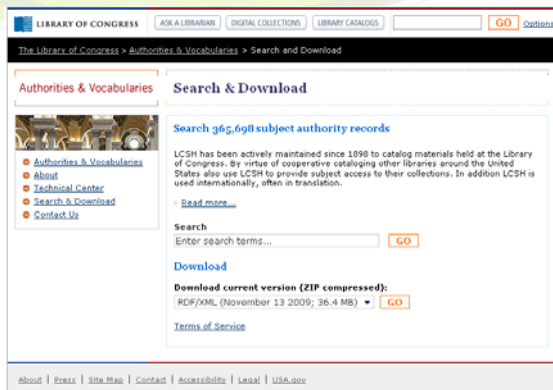
```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:skos="http://www.w3.org/2004/02/skos/Core#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  >
  <skos:ConceptScheme rdf:about="http://spines.org/thesaurus">
    rdf:about="http://spines.org/thesaurus">
    <dc:title>SPINES</dc:title>
    <dc:description> Tesauro de política científica. </dc:description>
    <dc:creator> UNESCO </dc:creator>
  </skos:ConceptScheme>
</rdf:RDF>
```

Un concepto:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/Core#">
  <skos:Concept rdf:about="http://spines.org/concept/0001">
    < rdf:about="http://spines.org/ concept/0001">
    <skos:externalID> A.01.0001 </skos:externalID>
    <skos:prefLabel> Capital </skos:prefLabel>
    <skos:altLabel> Activo </skos:altLabel>
    <skos:altLabel> Riqueza </skos:altLabel>
    <skos:inScheme rdf:resource="http://spines.org/thesaurus"/>
      rdf:resource="http://spines.org/thesaurus"/>
  </Concept>
</rdf:RDF >
```

al público su base de datos de encabezamientos de materia en formato SKOS de manera gratuita. Esta se puede visualizar y descargar en diversos formatos a partir de la siguiente URL:

<http://id.loc.gov/authorities/search/>



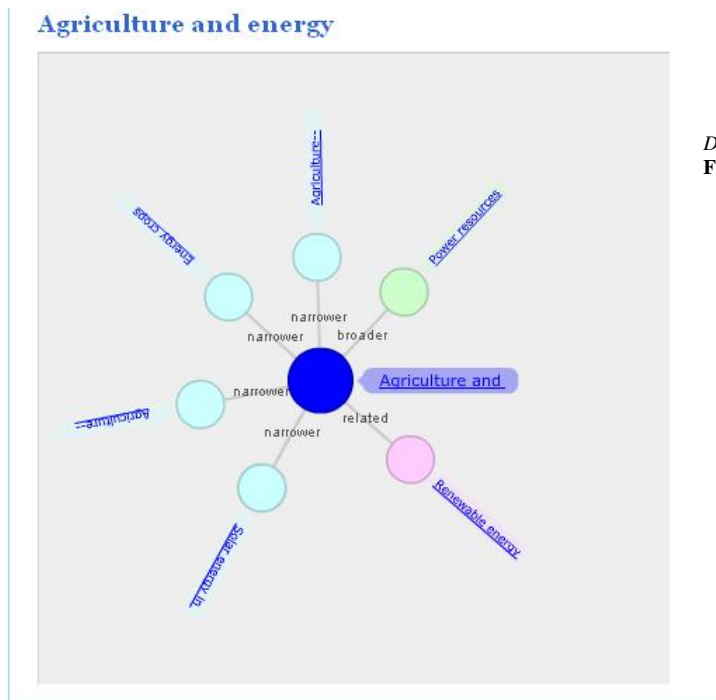
Pantalla de búsqueda

Fuente: <http://id.loc.gov/authorities/search/>



Aciertos: Agriculture and energy

Fuente: <http://id.loc.gov/authorities/search/>



Despliegue visual de resultados

Fuente: <http://id.loc.gov/authorities/search/>



3.5 SKOS-CORE frente a Zthes y Topic Maps

Tanto Topic Maps como Zthes vienen de líneas de trabajo más antiguas que se han visto en la necesidad de adaptarse a los nuevos tiempos marcados por el ritmo y las modas impuestas por la aparición de XML. Esto no quiere decir que estos proyectos sean mejores ni peores sino simplemente que encuentran un campo de aplicación más reducido y suponen soluciones concretas a problemas concretos, por lo cual están lejos de convertirse en un estándar de facto como suele ocurrir con las tecnologías propuestas por el Consorcio. Esta es precisamente la razón por la que se ha optado en este trabajo por un desarrollo basado en RDF/OWL/SKOS-Core.

4. Diseño metodológico para estructurar contenidos Web mediante SKOS-Core

4.1 Sitio web Chile para Niños

El sitio "Chile para Niños" tiene un propósito similar al sitio *Memoria Chilena*, de la Dirección de Bibliotecas, Archivos y Museos (DIBAM), pero está enfocado a un público infantil, de 4 a 9 años de edad, por lo que el lenguaje y las imágenes buscan una navegación fácil y amena, abordando cada temática por medio de juegos -como rompecabezas y sopa de letras-, fotografías, música, literatura, diccionarios e ilustraciones. Estos recursos se articulan como ejercicios que incentivan habilidades en torno al lenguaje, las matemáticas, la percepción visual, la abstracción y la capacidad para relacionar conceptos, entre otras.

La decisión de escoger este sitio para realizar una propuesta de diseño metodológico se debe, entre otras cosas, a la facilidad con la cual se pueden agrupar los sitios temáticos y al lenguaje directo y sencillo que el sitio emplea, al estar dirigido a niños.

Descripción general del sitio:

El sitio posee un menú principal compuesto por un Inicio, un Club, donde están los amigos de Memoriosa, quien es una niña que está presente en todo el sitio, la sección Juegos, donde se clasifican y organizan todos los juegos desplegados en los sitios temáticos que luego serán señalados en este informe, un Baúl, que registra todas las actividades realizadas por el Club, la Biblioteca, donde hay material de archivo digitalizado para compartir, como revistas infantiles, sitios de interés, libros animados, entre otros; un Índice de temas, que explica y contextualiza el contenido de cada uno de los sitios temáticos señalados en la barra vertical izquierda, y una sección de Links amigos del sitio. En la barra vertical izquierda hay un cuadro donde introduce al sitio con "Tesoros de Chile". Luego de ello se presentan los 34 Sitios temáticos, señalados a continuación:

-- Patrimonio Natural:

1. Cetáceos de Chile
2. Fauna terrestre
3. Paleontología
4. Parques Nacionales
5. Volcanes de Chile

-- Patrimonio Cultural:

6. Artesanía de Chile
7. Arturo Prat
8. Chiloé
9. Cocina Chilena
10. Estancias magallánicas
11. Familias chilenas
12. Fiestas y Juegos
13. Fútbol
14. Gabriela Mistral
15. Javiera Carrera
16. Lautaro
17. Museos de Chile
18. Música para Niños
19. Nicanor Parra

20. Oficios Tradicionales
21. Pablo Neruda
22. Padre Hurtado
23. Pueblo Mapuche
24. Pueblos del Norte
25. Pueblos del Sur
26. Rapa Nui
27. Roberto Matta
28. Salitreras
29. San Pedro de Atacama
30. Símbolos patrios
31. Sewell
32. Valparaíso
33. Vicente Huidobro
34. Violeta Parra

Cada uno de estos sitios se compone de juegos, música, literatura y recursos que ayudan al niño a entender mejor se qué se trata cada tema.

Para efectos metodológicos, se determinó seleccionar cuatro de estos sitios, que tuvieran una cierta relación entre sí y que tuvieran características comunes. Para tal efecto, los poetas Gabriela Mistral, Nicanor Parra, Pablo Neruda y Vicente Huidobro, fueron escogidos para realizar el diseño metodológico.

4.2 Descripción del modelo por capas

Para poder realizar el modelo de una red semántica de conceptos, relacionados entre los cuatro sitios temáticos escogidos, se determinó efectuar primero una descripción a través de rótulos de las capas que componen el sitio:

- La primera capa ha quedado plenamente establecida al momento de señalar los 34 sitios temáticos sobre los cuales se estructura el sitio "Chile para Niños".
- La segunda capa, corresponde a los contenidos que cubre cada uno de los cuatro sitios temáticos seleccionados.

- La tercera capa, está determinada por los conceptos destacados de cada uno de los artículos que se relacionan con el contenido articulado de los sitios temáticos seleccionados. De esta manera, se expone a continuación la segunda y tercera capa, explicadas previamente:

- **Gabriela Mistral**

- Dame la mano y danzaremos: poeta, profesora, escuelas, Temuco, Punta Arenas, Santiago, seudónimos, concurso de poesía, prosa, amor, soledad, muerte, Dios, educación, literatura, Premio Nobel de Literatura, Premio Nacional de Literatura.

- Visitar: Valle de Elqui, astronomía, Museo, poemas.

Literatura.

Poesía: Rocío, Todas íbamos a ser reinas, Hallazgo, ¿En dónde tejemos la ronda?, Miedo, Dame la mano, Piececitos,

El papagayo, Los que no danzan, Todo es ronda.

Prosa: La leyenda del viento, El árbol, El brasero, Elogio del libro, Proyectos..., Dos rizos de Jesús (cuento de Navidad)

Juegos: Asociación, Comprensión lectora, Juego del rompecabezas, Completar la poesía.

Audiovisual.

Entrevista.

- **Nicanor Parra**

- El anti-poeta: Nicanor Parra: poeta, músico, costurera, profesor, matemáticas, física, literatura, antipoesía.

- Algunas de sus obras: Poemas y antipoemas, La cueca larga, Versos de salón, Chistes para desorientar a la policía.

- ¿Qué son los artefactos? Religión, política, arte, ecología, sociedad de consumo.

- La familia Parra.

- ¿Y qué premios ha recibido Nicanor Parra? Premio Nacional de Literatura, Premio Juan Rulfo, Premio Reina Sofía de Poesía Iberoamericana, Premio Nobel de Literatura.

- El discurso de recibimiento del premio Juan Rulfo.

Literatura: Artefactos, Versos de salón, Chistes para desorientar a la policía, Poemas y antipoemas.

Fotos.

Juegos: Recordando a Nicanor Parra, Rompecabezas, Artefactos.

Música: Nicanor Parra.

- **Pablo Neruda**

- Pablo Neruda, el poeta: poeta, obrero ferroviario, maestra de escuela, pedagogía en francés, poesía, Premio Nobel de Literatura, cónsul, embajador

- Recorramos la poesía de Pablo Neruda: romanticismo, nostalgia, melancolía.

- Mini diccionario: Antigüedad, litoral central, cónsul, embajador, senador, Premio Nobel.

Literatura: Libro de las preguntas Neruda, Picaflor, Oda a la abeja, Lluve.

Música: Neruda.

Juegos: Memorice, Rompecabezas, Recordemos las poesías.

- **Vicente Huidobro**

- Diccionario de personajes: André Breton, Carlos Díaz Loyola, Eduardo Anguita, Faride Zerán, Gerardo Diego, Guillaume Apollinaire, Henriette Petit, Juan Gris, Kasimir Malevich, Pablo Neruda, Pablo de Rokha, Pablo Picasso, Pierre Reverdy, René de Costa, Robert Delaunay, Sara Malvar, Sonia Delaunay, Tristan Tzara, Volodia Teitelboim.

- El creacionismo: movimiento de vanguardia.

- Altazor o el viaje en paracaídas.

- Vicente Huidobro y la política.

- La pelea de Huidobro con Neruda: Pablo Neruda.

- Los 4 amores de Huidobro.

- La muerte del poeta.

- La Fundación Vicente Huidobro.

- ¿Te gustaría visitar la casa y tumba de Vicente Huidobro?

Literatura

Fotos

Juegos

Música

Poemas pintados

A partir de esta rotulación, se determinó que, al momento de crear la red semántica, los conceptos tuvieran relación. Es así como sólo algunos de los rótulos señalados en los cuatro sitios temáticos fueron extraídos para ser utilizados, y a la vez, fueron agregados otros que englobaran a los cuatro en general. De esta manera, el término común para los cuatro fue "Poesía", que no está explícitamente mencionado en los artículos de estos pequeños sitios. Se expone a continuación los conceptos que

finalmente serán los que se reflejarán en la red semántica de conceptos que configurarán este tesoro:

- Poesía
- Poeta
- Poetisa
- Gabriela Mistral
- Pablo Neruda
- Nicanor Parra
- Vicente Huidobro
- Premio Nacional de Literatura
- Premio Nobel de Literatura
- Política

4.3 Esquema de datos de red de términos

La simbología utilizada para la representación de los tesauros tiene un paralelo en formato SKOS

Term	
Broader terms	TG
Narrower terms	TE
Scope Note	NA
Used for	USE
Related terms	TR

La representación del etiquetado SKOS en base a los términos seleccionados del sitio www.chileparaninos.cl se estructura de la siguiente manera:

Term: POLÍTICA

Broader terms:

PABLO NERUDA
GABRIELA MISTRAL

Term: PREMIO NACIONAL DE LITERATURA

Broader terms:

PABLO NERUDA
GABRIELA MISTRAL

Term: NICANOR PARRA

Scope Note:

Antipoeta Chileno

Broader terms:

POETA

Narrower terms:

PREMIO NACIONAL DE LITERATURA

Related terms:

POESÍA

Term: PABLO NERUDA

Scope Note:

Poeta chileno

Broader terms:

POETA

Narrower terms:

PREMIO NOBEL
PREMIO NACIONAL DE LITERATURA
POLÍTICA

Related terms:

POESÍA

Term: POETA

Broader terms:

POESÍA

Narrower terms:

VICENTE HUIDOBRO
PABLO NERUDA
NICANOR PARRA

Term: POESÍA

Scope Note:

Género literario que emplea el lenguaje por sus cualidades estéticas y evocativas

Narrower terms:

POETISA

POETA

Term: PREMIO NOBEL

Broader terms:

GABRIELA MISTRAL

PABLO NERUDA

Term: POETISA

Broader terms:

POESÍA

Narrower terms:

GABRIELA MISTRAL

Term: GABRIELA MISTRAL

Scope Note:

Poetisa chilena, premio nobel de literatura 1945

Broader terms:

POETISA

Narrower terms:

POLÍTICA

PREMIO NOBEL

PREMIO NACIONAL DE LITERATURA

Industrial cooperation

Related terms:

POESÍA

Term: VICENTE HUIDOBRO

Broader terms:

POETA

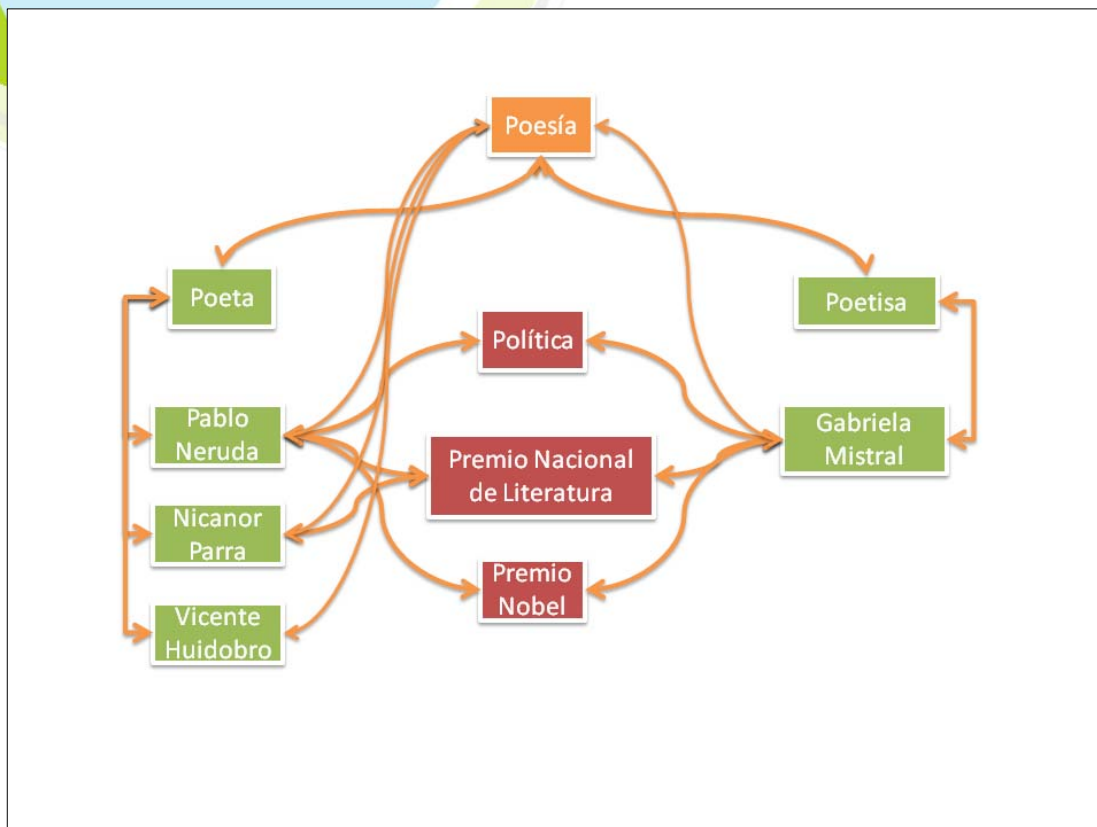
Scope Note:

Poeta Chileno

Related terms:

POESÍA

Figura: Representación del tesoro en mapa conceptual



4.4 Implementación de software para creación de tesoro con etiquetado SKOS

El proceso de edición y creación de un tesoro con etiquetado SKOS puede apoyarse en herramientas que gestionen el contenido del tesoro de manera automática. Existe software que permiten editar y crear tesauros, pero no todos exportan el contenido en etiquetas SKOS. Para efectos de esta investigación se utilizó un software open source llamado **Thmanager**, el cual puede ser descargado desde <http://sourceforge.net/projects/thmanager>. Este programa permite editar, estructurar, gestionar y visualizar un tesoro, y además permite exportarlo a etiquetado SKOS. Esta es una iniciativa de la W3C para la

representación de la organización del contenido como los tesauros, los esquemas de clasificación, taxonomías, listados de encabezamientos de materia y otros tipos de vocabulario controlado.

Las herramientas que apoyan la edición y creación de tesauros son desarrolladas para distintos tipos de lenguaje, sean estos en xml, html y javascript entre otros. Thmanager, a diferencia de otros editores de tesauros como thw32 o Protégé, permite desarrollarlos de manera rápida y sencilla.

Protégé es un software open source diseñado para desarrollar Ontologías del lenguaje robustas, exige pre-requisitos en el ordenador como graphViz, que permite desarrollar mapas de visualización de la ontología, requiere de práctica y de conocimientos especializados en ontologías y semántica, su lenguaje de programación de ontologías es OWL (Ontology Web Language).

Thew32 es un software open source diseñado para sistemas operativos Windows de 32 bits, es ligero y sencillo en cuanto al uso de sus herramientas, permite exportar los tesauros en formato HTML, JavaScript y XML. Este software pertenece a Tim Craven, quien además ha desarrollado otros software orientados al lenguaje y vocabularios controlados.

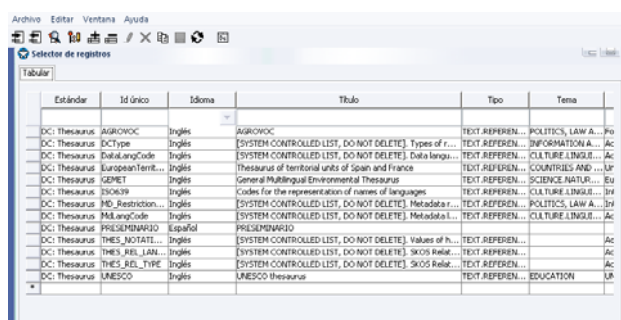
Thmanager, a diferencia de los anteriores programas, permite transformar el tesoro a formato SKOS y es de uso sencillo.

Thmanager es un software de uso simple, para el cual es necesario conocer ciertas herramientas del software para diseñar un tesoro y poder exportar el contenido.

En este capítulo se analiza este software en detalle y se entregan instrucciones para crear y editar un tesoro utilizando este programa.

4.5 Metodología de utilización

Al abrir el software se encontrará frente a una pantalla con una cantidad de tesauros que incluye Thmanager como demostraciones de tesauros hechos con este software, los que se pueden editar, eliminar, duplicar o exportar. Estos tesauros aparecen ordenados alfabéticamente.



The screenshot shows the Thmanager application window with a menu bar (Archivo, Editar, Ventana, Ayuda) and a toolbar. The main area displays a table with the following columns: Estándar, Id único, Idioma, Título, Tipo, and Tema. The table contains several rows of thesauri data, including AGROWOC, DOCTYPE, DublinLangCode, EuropeanTerr..., GEMET, SKOS3P, MD_Restriction..., MilLangCode, PRESEMVARIO, THES_NOTATE..., THES_REL_LAN..., THES_REL_TYPE, and UNESCO.

Estándar	Id único	Idioma	Título	Tipo	Tema
DC: Thesaurus	AGROWOC	Inglés	AGROWOC	TEXT REFEREN...	POLITICS, LAW A... Po
DC: Thesaurus	DOCTYPE	Inglés	[SYSTEM CONTROLLED LIST, DO NOT DELETE] Types of f...	TEXT REFEREN...	INFORMATION A... Ac
DC: Thesaurus	DublinLangCode	Inglés	[SYSTEM CONTROLLED LIST, DO NOT DELETE] Data langua...	TEXT REFEREN...	CULTURE LINGUI... Ac
DC: Thesaurus	EuropeanTerr...	Inglés	[Thesaurus of territorial units of Spain and France	TEXT REFEREN...	COUNTRIES AND ... UP
DC: Thesaurus	GEMET	Inglés	[General Multilingual Environmental Thesaurus	TEXT REFEREN...	SCIENCE NATUR... Eu
DC: Thesaurus	SKOS3P	Inglés	[Codes for the representation of names of languages	TEXT REFEREN...	CULTURE LINGUI... In
DC: Thesaurus	MD_Restriction...	Inglés	[SYSTEM CONTROLLED LIST, DO NOT DELETE] Metadata...	TEXT REFEREN...	POLITICS, LAW A... In
DC: Thesaurus	MilLangCode	Inglés	[SYSTEM CONTROLLED LIST, DO NOT DELETE] Metadata...	TEXT REFEREN...	CULTURE LINGUI... Ac
DC: Thesaurus	PRESEMVARIO	Español	PRESEMVARIO		
DC: Thesaurus	THES_NOTATE...	Inglés	[SYSTEM CONTROLLED LIST, DO NOT DELETE] Values of n...	TEXT REFEREN...	Ac
DC: Thesaurus	THES_REL_LAN...	Inglés	[SYSTEM CONTROLLED LIST, DO NOT DELETE] SKOS Relat...	TEXT REFEREN...	Ac
DC: Thesaurus	THES_REL_TYPE	Inglés	[SYSTEM CONTROLLED LIST, DO NOT DELETE] SKOS Relat...	TEXT REFEREN...	Ac
DC: Thesaurus	UNESCO	Inglés	UNESCO thesaurus	TEXT REFEREN...	EDUCATION UR

Ilustración 1, interfaz nativa de thmanager

Existen dos formas de creación de un nuevo tesoro, para la primera es necesario dirigirse a la opción Editar y hacer clic en la opción Insertar (ver ilustración 2)



Ilustración 2, menú de edición

Una vez hecho esto aparecerá una ventana en donde se pide el nombre del nuevo tesoro y definir el idioma (ver ilustración 3)



Ilustración 3, insertar un nuevo tesoro

El tesoro que se creó aparecerá en la lista de tesauros en la pantalla de inicio de Thmanager, y para comenzar la edición del tesoro se debe hacer clic sobre el tesoro y dirigirse a la opción **Archivo > Editor del Tesoro** (ver ilustración 4).

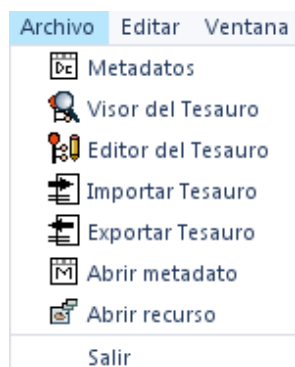


Ilustración 4, menú archivo

El editor de contenido de Thmanager se abrirá y se comenzará con la edición del nuevo tesoro. La estructura de Thmanager permite hacer relaciones de conceptos a distintos niveles jerárquicos y es posible relacionar términos aunque estos no pertenezcan a la jerarquía directa.

Thmanager permite crear el tesoro utilizando distintas herramientas, la primera es el editor de tesauros, donde los términos serán ingresados y relacionados. Para ingresar un nuevo término al tesoro es necesario, en primer lugar, definir el idioma con el que se presentara el tesoro y para hacerlo es necesario estar dentro del editor del tesoro.

Una vez que el idioma esté definido, es posible ingresar los términos correspondientes al tesauro que se creará, para hacerlo será necesario hacer clic sobre el botón “insertar un nuevo término” (ver ilustración 5), donde se desplegará una ventana que pedirá definir la URI correspondiente al término creado.

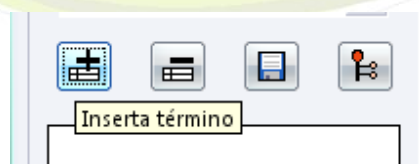


Ilustración 5, menú de editor de tesauro

Thmanager sólo permite relacionar términos en base a tres relaciones: la primera corresponde a la relación de término padre, lo que indica que el término que se edita está subordinado jerárquicamente a otro término, también acepta relación de término hijo, lo que indica que el término que se está editando tiene un término que se subordina a este, la última relación es la de término relacionado, lo que permite relacionar términos sin necesidad de que estos sean se subordinen entre sí.

Los términos ingresados al tesauro deberán ser descritos y relacionados en base a una serie de pasos del editor de tesauros de Thmanager. Las URI permiten ubicar el término en base a una línea de caracteres que se definen para el termino en particular, lo que permite relacionar un término con otro a través de la asociación de las URI; en la ilustración 6 se observa que la URI del término “abuelo” es <http://abuelo>, esta URI permite asociar otro término como “papa” al término “abuelo”

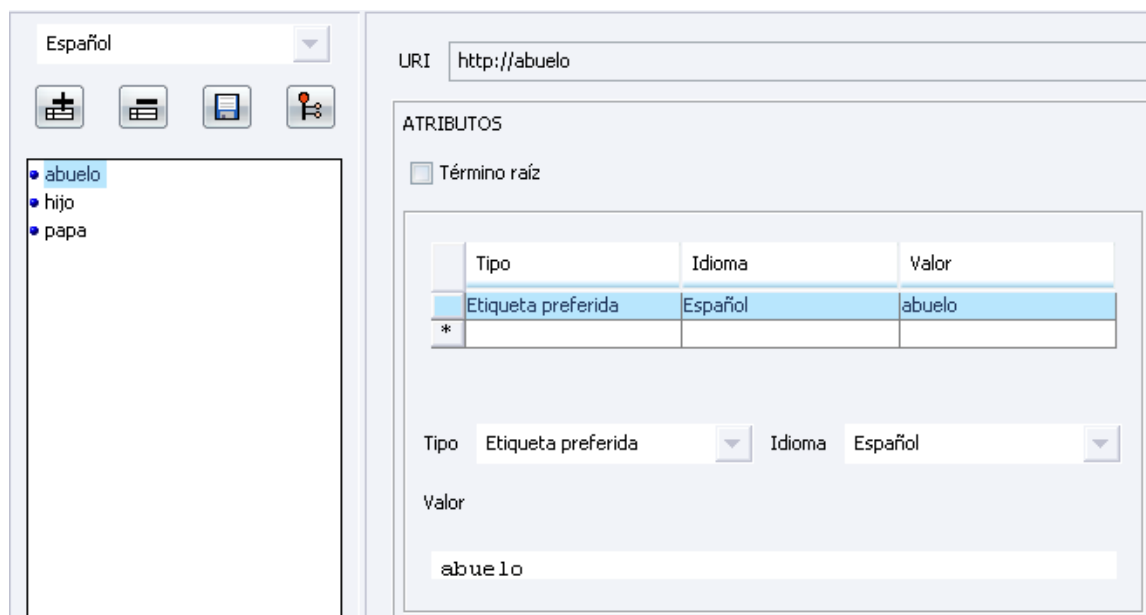


Ilustración 6, ejemplificación de los atributos de los términos

Para establecer las distintas relaciones semánticas de los términos es necesario ubicar las URI de los términos que fueron creados. En la siguiente ilustración se aprecia que el término abuelo tiene relacionado como término “hijo” la URI correspondiente al término “papa”, lo que indica que en la jerarquía de términos creados el término “papa” está subordinado al término “abuelo”, para establecer la relación entre el término “papa” y el término “abuelo”, es necesario definir en las relaciones del término “papa” a quien se subordinado, lo que permite ascender o descender en la jerarquía de términos.

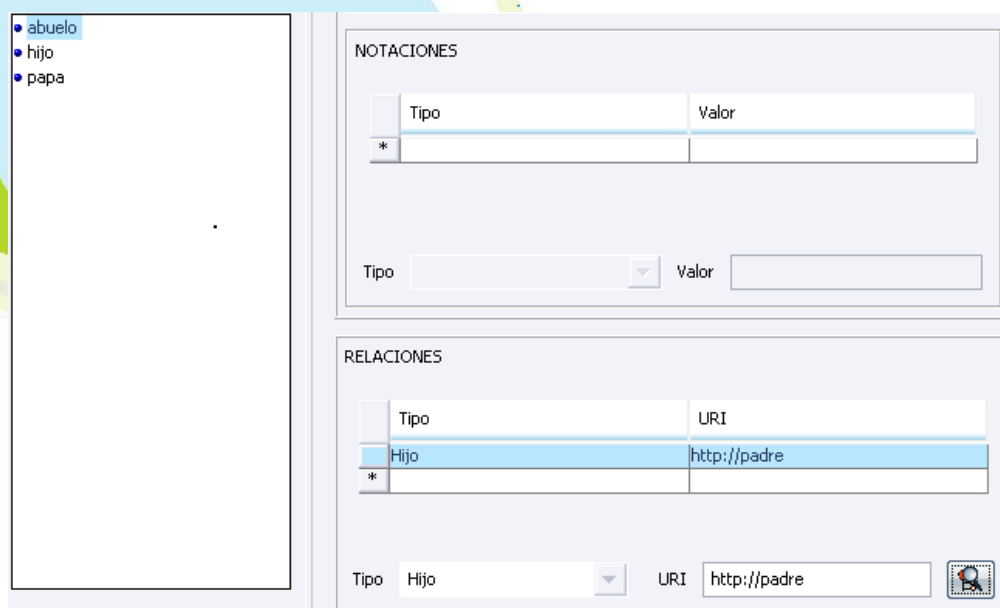


Ilustración 7, presentación de la relación entre los términos

Una vez que las relaciones son establecidas, es necesario verificarlas; para hacerlo, Thmanager integra un visor de tesauros, lo que permite navegar a través de los términos relacionados, para activarla es necesario dirigirse a **archivo > visor del tesauro** (ver ilustración 4), donde se desplegará un visor de tesauros en donde aparecen todas las relaciones establecidas en el editor de tesauros (ver ilustración 8), en donde el término “papa” tiene como término padre “abuelo” y como término hijo a “hijo”.

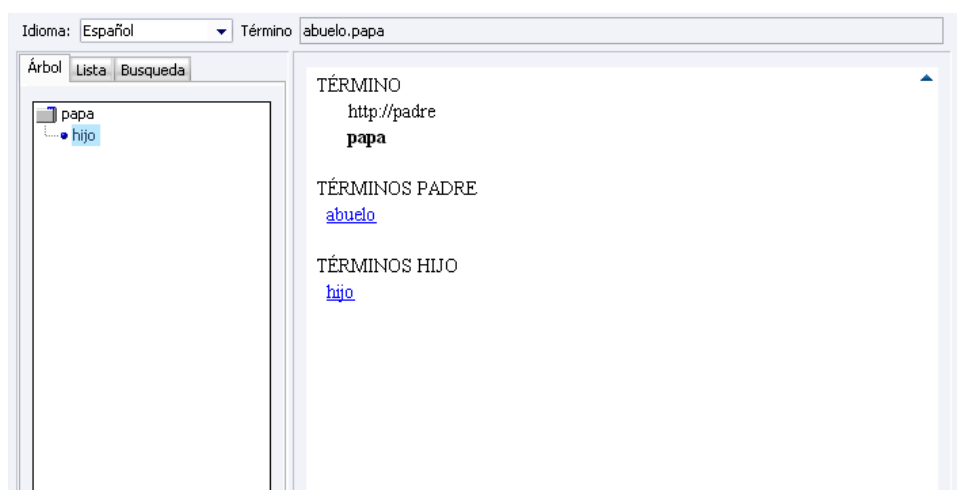


Ilustración 8, Visualización del tesauro creado

La exportación del tesaurus creado en Thmanager se hará en formato SKOS y para hacerlo será necesario ir a la opción **Archivo > Exportar Tesaurus** (ver ilustración 4), al exportar el tesaurus se debe definir la ubicación en donde el archivo se alojará y el nombre que le daremos al tesaurus (ver ilustración 9)



Ilustración 9, proceso de exportación del tesaurus

La exportación alojará dos archivos en la ubicación definida, el tesaurus se encuentra definido dentro del archivo skos.xml.

Esquema skos-core producto de la conceptualización del sitio web en estudio

```
<?xml version="1.0" encoding="UTF-8" ?>
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:skos="http://www.w3.org/2004/02/skos/core#">
- <rdf:Description rdf:about="http://POLITICA">
  <skos:broader rdf:resource="http://PABLONERUDA" />
  <rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
  <skos:prefLabel xml:lang="es">POLITICA</skos:prefLabel>
  <skos:broader rdf:resource="http://GABRIELAMISTRAL" />
  <skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
  </rdf:Description>
- <rdf:Description rdf:about="http://PREMIONACIONAL">
```

```
<skos:broader rdf:resource="http://PABLONERUDA" />
<rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
<skos:broader rdf:resource="http://GABRIELAMISTRAL" />
<skos:prefLabel xml:lang="es">PREMIO NACIONAL DE
LITERATURA</skos:prefLabel>
<skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
</rdf:Description>
- <rdf:Description rdf:about="http://NICANORPARRA">
<skos:broader rdf:resource="http://POETA" />
<rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
<skos:related rdf:resource="http://POESIA" />
<skos:notation
rdf:datatype="http://iaaa.cps.unizar.es#code">http://www.chileparan
inos.cl/temas/nicanor-parra/index.html</skos:notation>
<skos:prefLabel xml:lang="es">NICANOR
PARRA</skos:prefLabel>
<skos:narrower rdf:resource="http://PREMIONACIONAL" />
<skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
<skos:scopeNote xml:lang="es">Antipoeta
Chileno</skos:scopeNote>
</rdf:Description>
- <rdf:Description rdf:about="http://PABLONERUDA">
<skos:broader rdf:resource="http://POETA" />
<rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
<skos:notation
rdf:datatype="http://iaaa.cps.unizar.es#code">http://www.chileparan
inos.cl/temas/pabloneruda/index.html</skos:notation>
<skos:related rdf:resource="http://POESIA" />
<skos:prefLabel xml:lang="es">PABLO
NERUDA</skos:prefLabel>
<skos:narrower rdf:resource="http://PREMIONACIONAL" />
<skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
<skos:narrower rdf:resource="http://PREMIONOBEL" />
<skos:narrower rdf:resource="http://POLITICA" />
<skos:scopeNote xml:lang="es">Poeta chileno</skos:scopeNote>
</rdf:Description>
- <rdf:Description rdf:about="http://POETA">
<skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
<skos:broader rdf:resource="http://POESIA" />
```

```

<rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
<skos:narrower rdf:resource="http://VICENTEHUIDOBRO" />
<skos:narrower rdf:resource="http://PABLONERUDA" />
<skos:narrower rdf:resource="http://NICANORPARRA" />
<skos:prefLabel xml:lang="es">POETA</skos:prefLabel>
</rdf:Description>
- <rdf:Description rdf:about="http://POESIA">
<skos:prefLabel xml:lang="es">POESIA</skos:prefLabel>
<skos:definition xml:lang="es">Género literario que emplea el
lenguaje por sus cualidades estéticas y
evocativas</skos:definition>
<rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
<skos:narrower rdf:resource="http://POETA" />
<skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
<skos:narrower rdf:resource="http://POETISA" />
</rdf:Description>
- <rdf:Description rdf:about="http://PREMIONOBEL">
<skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
<skos:broader rdf:resource="http://GABRIELAMISTRAL" />
<skos:prefLabel xml:lang="es">PREMIO
NOBEL</skos:prefLabel>
<skos:broader rdf:resource="http://PABLONERUDA" />
<rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
</rdf:Description>
- <rdf:Description rdf:about="http://POETISA">
<skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
<skos:broader rdf:resource="http://POESIA" />
<skos:narrower rdf:resource="http://GABRIELAMISTRAL" />
<skos:prefLabel xml:lang="es">POETISA</skos:prefLabel>
<rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
</rdf:Description>
- <rdf:Description rdf:about="http://GABRIELAMISTRAL">
<skos:scopeNote xml:lang="es">Poetisa chilena, premio nobel de
literatura 1945</skos:scopeNote>
<skos:related rdf:resource="http://POESIA" />
<rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
<skos:narrower rdf:resource="http://POLITICA" />
<skos:notation
rdf:datatype="http://iaaa.cps.unizar.es#code">http://www.chileparan
inos.cl/temas/gabrielamistral/index.html</skos:notation>

```

```
<skos:prefLabel xml:lang="es">GABRIELA
MISTRAL</skos:prefLabel>
<skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
<skos:narrower rdf:resource="http://PREMIONOBEL" />
<skos:broader rdf:resource="http://POETISA" />
<skos:narrower rdf:resource="http://PREMIONACIONAL" />
</rdf:Description>
- <rdf:Description
  rdf:about="http://iaaa.unizar.es/thesaurus/PRESEMINARIO">
  <skos:hasTopConcept rdf:resource="http://POESIA" />
  <rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#ConceptSche
me" />
  </rdf:Description>
- <rdf:Description rdf:about="http://VICENTEHUDOBRO">
  <skos:inScheme
rdf:resource="http://iaaa.unizar.es/thesaurus/PRESEMINARIO" />
  <skos:notation
rdf:datatype="http://iaaa.cps.unizar.es#code">http://www.chileparan
inos.cl/temas/vicentehuidobro/index.html</skos:notation>
  <skos:broader rdf:resource="http://POETA" />
  <skos:prefLabel xml:lang="es">VICENTE
HUDOBRO</skos:prefLabel>
  <rdf:type
rdf:resource="http://www.w3.org/2004/02/skos/core#Concept" />
  <skos:related rdf:resource="http://POESIA" />
  <skos:definition xml:lang="es">Poeta chileno</skos:definition>
</rdf:Description>
</rdf:RDF>
```


Conclusiones

Se postuló un modelo y se aplicó un software para estructurar un tesoro con los conceptos agrupados en el sitio "Chile para niños". De esta forma el rotulado de dichos contenidos pudo verificarse a través de un mapa que establece relaciones que ayuden a comprender en su contexto los artículos y la respectiva información que éstos entregan.

El uso de la Web Semántica y de un sistema de lenguaje organizado permitió desarrollar con cierta facilidad la agrupación y el rotulado de dichos contenidos, facilitando su acceso y comprensión entre quienes pretenden aspirar a aplicarlo. Su encriptado en Flash y la variedad de contenidos dispersos en el sitio jugó en contra al momento de profundizar en otras relaciones que ayudaran a aumentar el mapa y de esta forma los conceptos entre sí.

El uso que se haga de los buscadores, en la respuesta ante las demandas de información relevante, en un sitio Web, será un desafío a considerar al momento de diseñar nuevos modelos de recuperación de información, y por consiguiente, de sistemas de lenguajes controlados y normalizados ante el usuario.

El sitio analizado en este trabajo es pequeño y focalizado en un grupo etario. Cuando el sitio es un portal de noticias, donde el público es más amplio y los contenidos se crean permanentemente, ¿será posible profundizar y abarcar de una mejor forma este diseño metodológico? Quizás la respuesta sea no, pero es ante esos desafíos que se hace necesario un conocimiento acabado de posibles alternativas que solucionen esta carencia, que de paso dificulta y agrieta la relación que tienen los sitios web con sus usuarios, que cada vez cuentan con menos tiempo y espacio para cubrir sus cada vez más altas necesidades de información"

Bibliografía consultada

1. BEAT, gestor de thesaurus sobre PC [en línea] [Consulta: 5 de octubre, 2009] Disponible en:
http://www.elprofesionaldelainformacion.com/contenidos/1992/mayo/beat_gestor_de_thesaurus_sobre_pc.html
2. CARABALLO Pérez, Yeter. Los Topic Maps y su relación con las redes sociales [en línea] [Consulta: 18 de octubre de 2009] Disponible en:
http://bvs.sld.cu/revistas/aci/vol16_4_07/aci091007.html
3. EÍTO Brun, Ricardo. Topics maps y la indización de recursos electrónicos en la web [en línea] [Consulta: 18 de octubre de 2009] Disponible en:
<http://www.elprofesionaldelainformacion.com/contenidos/2003/marzo/15.pdf>
4. GUZMÁN Luna, Jaime a; Torres Pardo, Durley and López García, Alba Nubia. Desarrollo de una ontología en el contexto de la Web Semántica a partir de un tesoro documental tradicional. Rev. Interam. Bibliot [En: línea]. 2006, vol.29, n.2 [Consultado: 8 de Noviembre, 2009], pp. 79-95. Disponible en:
<http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-09762006000200005&lng=en&nrm=iso>. ISSN 0120-0976.
5. IGLESIAS Maturana, María Texia. Análisis y Recuperación de Información Documental. Santiago de Chile: IPS, 1992. ISBN: 956-7026-03-3
6. KRSULOVIC-MORALES, Ernesto y GUTIÉRREZ, Claudio. Propuesta para la Creación de una Ontología sobre Departamentos Universitarios de Computación en Chile [en línea] [Consulta: 18 de octubre de 2009] Disponible en: <http://www.dcc.uchile.cl/~cguiterr/articulos/ontologia.pdf>

7. LIBRARY of Congress [En línea]. Authorities and vocabularies [Consultado: 8 de Noviembre, 2009]. Disponible en: <http://id.loc.gov/authorities/search/>
8. MARZAL, Miguel Ángel... et al. Mapas Conceptuales y Presentación Gráfica del Tesauro: Aplicación a las Bibliotecas Educativas [en línea]. [Consulta: 5 de octubre, 2009] Disponible en <http://cmc.ihmc.us/cmc2006Papers/cmc2006-p70.pdf>
9. MÉNDEZ Rodríguez, Eva M^a. Metadatos y Tesauros: aplicación de XML/RDF a los sistemas de organización del conocimiento en Intranets [en línea] [Consulta: 18 de octubre de 2009] Disponible en: <http://www.bib.uc3m.es/~mendez/publicaciones/fesabid00/fesabid002.pdf>
10. MILES, Alistair [En línea]. The web and SKOS. [Consultado: 8 de Noviembre, 2009]. Disponible en: <http://www.iskouk.org/>
11. MOREIRO, J. A., SÁNCHEZ Cuadrado, S., MORATO, J. Panorámica y tendencias en topic maps [en línea] [Consulta: 18 de octubre de 2009] Disponible en: <http://www.hipertext.net/web/pag229.htm>
12. MOREIRO González, José... [et al.] De los tesauros a los Topics Maps: nuevo estándar para la representación y la organización de la información [en línea] [Consulta: 18 de octubre de 2009] Disponible en: <http://www.periodicos.ufsc.br/index.php/eb/article/viewFile/178/5470>
13. MOYA Martínez, Gregorio, GIL Leiva, Isidoro. Evaluación de softwares de gestión de tesauros [en línea] [Consulta: 5 de octubre, 2009] Disponible en: <http://www.cinfo.cu/Userfiles/file/Cinfo/cinfo2001/v32n3a2001/TESAUROS.PDF>

14. NORMA UNE 50106-90 [en: Línea]. [Consulta:4 de oct. 2009]. Disponible en:
http://www.google.com/url?sa=t&source=web&ct=res&cd=1&url=http%3A%2F%2Fbibliologia.info%2Farchivos%2Fune50106_90.pdf&ei=tGHOSqu sKsyY8AbK5IXzAw&usq=AFQjCNGvPIHlyIQSOay0W8BralWd2NdTYg&sig2=e2xvGqloSwEAzpf68lzFiA
15. NOTA del Grupo de Trabajo del W3C de 18 de Agosto de 2009. [En línea]. Manual de SKOS (Simple Knowledge Organization System, Sistema para la Organización del Conocimiento simple). [Consultado: 8 de Noviembre, 2009]. Disponible en: <http://skos.um.es/TR/skos-primer/#secpref>
16. ODLIS Glosario [en: Línea]. [Consulta: 4 de oct. 2009]. Disponible en: http://lu.com/odlis/odlis_t.cfm
17. PAREJA Lora, Antonio [En línea]. Anotación Semántica y Recuperación de Información [Consultado: 8 de Noviembre, 2009]. Disponible en: www.fesabid.org/santiago2007/descargas/mesas/apareja_lora.pdf.
18. PASTOR Sánchez, Juan Antonio. Diseño de un sistema colaborativo para la creación y gestión de tesauros en Internet basado en SKOS [en línea] [Consulta: 18 de octubre de 2009] Disponible en: http://www.tesisenxarxa.net/TESIS_UM/AVAILABLE/TDR-0403109-113737//PastorSanchez.pdf
19. PÉREZ Agüera, José Ramón. Automatización de tesauros y su utilización en la Web Semántica [en línea] [Consulta: 18 de octubre de 2009] Disponible en: <http://www.ub.es/bid/13perez2.htm>
20. RDF Thesaurus. WC3. Disponible en: <http://www.w3.org/2001/sw/Europe/reports/thes/rdfthes.html>

21. RESOURCE Description Framework, Wikipedia. [En: línea]. Fecha consulta 20-10-2009. Disponible en:
http://es.wikipedia.org/wiki/resource_description_framework
22. ROSA, Antonio de la. Tesauros, tesauros automaticos, tesauros automaticos online [en línea] [Consulta: 5 de octubre, 2009] Disponible en:
http://www.elprofesionaldelainformacion.com/contenidos/1999/julio/tesauros_tesauros_automaticos_tesauros_automaticos_online.html
23. ROSENFELD, Louis, Morville, Peter. Information Architecture: for the World Wide Web. Beijing: O' Reilly, 200-.
24. SIMPLE Knowledge Organization System [En línea]. En: Wikipedia. [Consultado: 8 de Noviembre, 2009]. Disponible en:
http://es.wikipedia.org/wiki/Simple_Knowledge_Organization_System
25. SISTEMA de Bibliotecas DuocUC incorpora aquabrowser [en línea] [Consulta: 5 de octubre, 2009] Disponible en:
<http://www.slideshare.net/hreyes/sistema-de-bibliotecas-duocuc-incorpora-aquabrowser>
26. SKOS-Core frente a Zthes y Topic Maps [En línea]. [Consultado: 8 de Noviembre, 2009]. Disponible en:
<http://www.w3.org/2001/sw/Europe/events/200406-esp/trabajo-final-extratesauros/node12.html>
27. SKOS-Core [En línea]. [Consultado: 8 de Noviembre, 2009]. Disponible en:
<http://www.w3.org/2001/sw/Europe/events/200406-esp/trabajo-final-extratesauros/node6.html>
28. SOFTWARE para la construcción de tesauros. [en línea] [Consulta: 5 de octubre, 2009] Disponible en:
<http://www3.unileon.es/dp/abd/tesauro/pagina/software/software.htm>

29. UNIVERSIDAD Carlos III de Madrid. [en línea]. Fecha consulta: 19-10-2009. Disponible en:
<http://www.bib.uc3m.es/~mendez/publicaciones/fesabid00/fesabid002.pdf>
30. VENTANA indiscreta. Glosario [en: Línea]. [Consulta:4 de oct. 2009].
Disponible en: <http://web.ftapies.com/thesaurus/spip.php?article128>

Serie Bibliotecología y Gestión de Información.

Títulos publicados 2010

- Nº 53 Usos de Twitter en la Educación Superior. Guillermo Toro Araneda
- Nº 54 Videojuegos, videojugadores y bibliotecas públicas. Francisco Bustos Naulin - José Guzmán Seguel
- Nº 55 Gestión de recursos culturales en bibliotecas comunitarias- Una propuesta de servicio. Antonio Gorosito López – Paula Szafran Maiche
- Nº 56 Recursos de información en español sobre prevención y actuación en casos de desastre en bibliotecas y archivos. Mariela Ferrada Cubillos

Disponible en : <http://eprints.rclis.org>

NORMAS DE PUBLICACION

• Objetivos

La **Serie Bibliotecología y Gestión de Información** tiene por objetivo difundir la productividad, académica, las investigaciones y las experiencias de profesionales del área de la de Bibliotecología y Ciencia de la Información y del sector afin al mundo del libro y la lectura.

• Alcance y política editorial

Los trabajos a ser considerados en la Serie Bibliotecología y Gestión de Información, deben ser inéditos, no publicados en otras revistas o libros. Excepcionalmente el Comité Editorial podrá aceptar artículos que no cumplan con este requisito.

- **Arbitraje:** Los artículos recibidos serán sometidos a evaluación, a recomendación del Director de la Serie, donde el Comité Editorial enviará los trabajos a árbitros independientes para su aceptación o rechazo. En este último caso, se emitirá un informe al autor/a donde se señalen las razones de la decisión. El Comité Editorial podrá solicitar trabajos a autores de reconocido prestigio, quienes no serán sometidos al proceso de evaluación por árbitros.

• Forma y preparación de manuscritos

- **Extensión:** El artículo deberá tener una extensión entre 12 y 100 páginas, tamaño carta, espacio 1,5, cuerpo 12, incluidos gráficos, cuadros, diagramas, notas y referencias bibliográficas.

- **Idiomas:** Se aceptan trabajos en castellano, portugués e inglés, los cuales serán publicados en su idioma original.

- **Resumen y palabras claves:** El trabajo deberá tener un resumen en español e inglés en la primera página, de no más de 200 palabras, que sintetice sus propósitos y conclusiones más relevantes. De igual modo, deben incluirse tres palabras claves, que en lo posible no se encuentren en el título del trabajo, para efectos de indización bibliográfica.

- **Nota biográfica:** En la primera página, en nota al pie de página, deben consignarse una breve reseña curricular de los/as autores/as, considerando nacionalidad, título y/o grados académicos, desempeño y/o afiliación profesional actual y sus direcciones de correo electrónico, para posibles comunicaciones de los/las lectores/as con los autores/as.

- **Referencia bibliográfica:** Utilizar para las referencias bibliográficas la modalidad de (Autor, año) en el texto, evitando su utilización a pie de página. Ejemplo: (González, 2006). Agregar al final del texto, la bibliografía completa. Sólo con los/las autores/as y obras citadas, numeradas y ordenadas alfabéticamente. Para el formato de la bibliografía, utilizar la "Guía para la presentación de referencias bibliográficas de publicaciones impresas y electrónicas" disponible en formato electrónico en :

<http://eprints.rclis.org/archive/00005163/01/ReferenciasBibliograficas.pdf>

- **Derechos:** Los derechos sobre los trabajos publicados, serán cedidos por los/as autores/as a la **Serie**.

- **Investigadores jóvenes:** El Comité Editorial considerará positivamente el envío de trabajos por parte de profesionales y/o investigadores/as jóvenes, como una forma de incentivo y apoyo a quienes comienzan su carrera en investigación.

- **Ejemplares de cortesía:** Los/as autores/as recibirán un ejemplar de cortesía del trabajo publicado.

• Envío de manuscritos

Todas las colaboraciones deberán ser enviadas impresas en duplicado. Los autores/as podrán remitir sus artículos en CD, o al correo electrónico: hector.gomez@utem.cl , en programa Word (office).