

Sistema de información documental sobre desplazamiento forzado en Colombia basado en ontologías*

Sandra Patricia Arenas Grisales^{**}
Ana María Tangarife Patiño^{***}
Wilmar Arley Martínez Márquez^{****}
Jaime Alberto Bornacelly Castro^{*****}

Resumen

Presenta el proceso de elaboración del Sistema de información documental sobre desplazamiento forzado en Colombia basado en ontologías. La investigación se planteó como objetivo general construir una ontología que representa conceptualmente la terminología sobre el desplazamiento forzado en Colombia, de manera que permita representar y recuperar la información referencial y en texto completo de la documentación producida. Presenta una breve descripción de lo que son las ontologías y su aplicación para la gestión y recuperación de información. En la construcción de la ontología hubo tres momentos, en cada uno de los cuales se integraron y adaptaron diferentes metodologías propuestas desde la terminología, la documentación y la ingeniería informática. Para la creación del árbol de dominio, la confección del tesauro y la estructuración de las relaciones jerárquicas se implementó la metodología propia de la terminología; en la creación de la ontología las metodologías propuestas por Noy y Macguiness y la Universidad Politécnica de Madrid para la estructuración de ontologías y, la aplicación de diferentes lenguajes y herramientas informáticas. Finalmente se describe el sistema de información documental y la interfaz de consulta.

Palabras clave: desplazamiento forzado, ontologías, sistemas de información documental, terminología.

Cómo citar este artículo: Arenas Grisales, Sandra Patricia; *et al.* Sistema de información documental sobre desplazamiento forzado en Colombia basado en ontologías. *Revista Interamericana de Bibliotecología*. Ene. – Jun. 2009, vol. 32, no. 1; p. 129-152.

Artículo recibido: 13 de abril de 2009. Aprobado: 16 de junio de 2009.

* Artículo resultado de la investigación *Ontología del Desplazamiento Forzado en Colombia integrada a un sistema de Información*, financiada por el CODI, Universidad de Antioquia. Medellín-Colombia. Realizado entre 2007-2009

** Investigadora principal. Bibliotecóloga, Magíster en Ciencia Política. Docente, Escuela Interamericana de Bibliotecología, Universidad de Antioquia. Medellín, Colombia. sarena@bibliotecologia.udea.edu.co

*** Coinvestigadora. Bibliotecóloga. Jefe Centro de Documentación Instituto de Estudios Políticos, Universidad de Antioquia. Medellín, Colombia. anamariatangarife@gmail.com

**** Auxiliar de investigación. Filósofo. Docente. Instituto de Estudios Políticos, Universidad de Antioquia. Medellín, Colombia. apofainon@gmail.com

***** Estudiante en formación. Bibliotecólogo. Escuela Interamericana de Bibliotecología. Universidad de Antioquia. Medellín, Colombia. eibbornacelly@bibliotecologia.udea.edu.co

Abstract

This paper presents the elaboration of the ontology based system of documental information about forced displacement in Colombia. The research had as its main purpose to build an ontology that allows to represent and retrieve the reference information and the full text of documents on the subject. It presents a brief description of ontology and its application in information management and retrieval. The building of the ontology included three moments, which required the adoption of diverse methodologies of terminology, documentation, and computer engineering. The methodology of terminology was used to build the domain tree and the thesaurus and to structure the hierarchical relationships. The methodologies proposed by Noy and Macguiness and by the Universidad Politécnica de Madrid to structure ontologies were used to create the ontology and to determine the application of diverse computer's languages and tools. Finally, the paper describes the system of documental information, and the user's interface.

Keywords: Forced displacement, ontologies, system of documental information, terminology.

How to cite this article: ARENAS GRISALES, Sandra Patricia; et al. Ontology based system of documental information about forced displacement in Colombia. *Revista Interamericana de Bibliotecología*. Jan. - Jun. 2009, vol. 32, no. 1; p. 129-152

Introducción

El desplazamiento forzado es una de las expresiones claras del conflicto armado que se vive en nuestro país. De acuerdo con cálculos del Codhes, entre 1985 y septiembre de 2005 fueron desplazadas 3.662.842 personas¹. Para el Codhes, los desplazados por el conflicto interno colombiano entre enero y junio de 2008 alcanzaron 270.675, un 41% más que en el mismo periodo de 2007 y un récord que no se registraba desde 1985.² El desplazamiento es una problemática que afecta a casi todo el territorio nacional y las cifras actuales del CODHES demuestran que, lejos de superarse, la crisis humanitaria se acentúa en el país. Según el mismo informe, el desplazamiento intraurbano crece en forma alarmante; por su parte, la Personería de Medellín, registra que en el período enero-mayo de 2008, debieron desplazarse dentro de la ciudad 567 personas. Igual situación se presenta en Bogotá, Cali, Buenaventura y Santa Marta.³ Cerca del 10 % del total de la población colombiana se ha desplazado por violencia en las dos últimas décadas y 560 mil

-
1. *Desafíos para construir nación. El país ante el desplazamiento, el conflicto armado y la crisis humanitaria, 1995-2005*. Bogotá: Conferencia Episcopal de Colombia, Secretariado Nacional de Pastoral Social, Sección de movilidad humana; CODHES, 2006. p. 19
 2. Tapando el sol con las manos: Informe sobre desplazamiento forzado, conflicto armado y derechos humanos, enero-junio de 2008. [En línea] *Codhes Informa*. Ene.-Jun 2008, no. 74. Disponible en: http://www.codhes.org/index.php?option=com_docman&task=cat_view&gid=60&Itemid=50. [Consulta: febrero 13 de 2009]
 3. *Ibid.* p. 6

colombianos buscan refugio en diversos países del mundo, aduciendo falta de garantías a su vida, su integridad y su libertad.⁴

Debido a la trascendencia del fenómeno en términos sociales, políticos, económicos y de derecho humanitario, distintas organizaciones no gubernamentales, instituciones académicas y entidades estatales e internacionales producen documentación sobre el fenómeno del desplazamiento en Colombia; sin embargo, ésta no es recopilada eficazmente en un sistema de información único. Han surgido iniciativas como la Red Nacional de Investigación sobre el Desplazamiento Forzado en Colombia, REDIF, el proyecto de biblioteca digital en asuntos humanitarios del Instituto de Estudios Humanitarios, que coordina OCHA (Coordinación de Asuntos Humanitarios de Naciones Unidas) y la Cátedra Nacional de Desplazamiento Forzado, de la Universidad Nacional, que coordina Marta Nubia Bello.

Como una alternativa de solución al problema de la dispersión e invisibilización de la información y a la polisemia conceptual y terminológica, se desarrolló la investigación *Ontología sobre el desplazamiento forzado en Colombia integrada a un sistema de información*. El objetivo general era construir una ontología donde se representara conceptualmente la terminología sobre el desplazamiento forzado en Colombia, de manera que permitiera la representación y recuperación de la información. Dicha ontología es el soporte estructural de un sistema de documentación en red sobre desplazamiento forzado, que permitirá el acceso a información referencial y en texto completo de la documentación producida sobre el tema.

Gracias a la colaboración del profesor Jaime Guzmán Luna, del grupo de Inteligencia Artificial de la Universidad Nacional, y de los asistentes Néstor Eduardo Rivera Vidal y Grissa Vianney Maturana González, que demostraron que «todo se puede», el proyecto pudo contar con la asesoría académica necesaria para ampliar el horizonte de posibilidades en la construcción de la ontología y utilizar algunas de las herramientas básicas para su desarrollo.⁵

El producto final de la investigación es un Sistema de Información Documental que, sobre la base de una ontología, tiene como objetivo recopilar y poner a disposición de los usuarios la documentación producida en Colombia sobre el desplazamiento forzado.

4. *Ibid.* p. 7

5. Véase: GUZMÁN LUNA, Jaime A.; TORRES PARDO, Durley; Ovalle, Demetrio A. SABIOS: una aplicación de la Web semántica para la gestión de documentos digitales. *Revista Interamericana de Bibliotecología*. Ene.-Jun. 2007, vol. 30, no. 1; p. 51-72.

En este artículo se presenta la metodología utilizada en la investigación, que combina métodos propios de la informática y la terminología, los resultados finales de la investigación y desarrollos futuros.

I. Metodología

La investigación se planteó como objetivos generales construir una ontología en donde se representara conceptualmente la terminología sobre el desplazamiento forzado, y diseñar un sistema de información en red sobre desplazamiento forzado que tuviera como estructura básica la ontología. Como objetivos específicos se propuso, en primer lugar, estructurar una ontología en donde se representaran conceptualmente los términos comúnmente usados entre los especialistas en desplazamiento forzado, con el fin de facilitar la comunicación entre ellos; en segundo lugar, compilar la información producida sobre desplazamiento forzado en diversos formatos y por diferentes instituciones, de forma que se constituyera en una herramienta de consulta para investigadores e instituciones estatales y ONG's interesadas en el tema; en tercer lugar, permitir el acceso a información referenciada y en texto completo (datos cuantitativos y cualitativos, investigaciones, informes, memorias de eventos, legislación, políticas públicas, fotografías, mapas, etc.) sobre el desplazamiento forzado para público interesado en el tema.

A continuación se presenta una breve descripción de lo que son las ontologías y los usos para la gestión y recuperación de información.

En áreas especializadas de conocimiento se hacen necesarias estructuras que ayuden a modelar dicho conocimiento, entre las que se encuentran herramientas como tesauros, taxonomías y ontologías. Las ontologías son representaciones de conocimiento que resultan de seleccionar un dominio o ámbito de conocimiento, y representar formalmente los conceptos y las relaciones entre ellos, todo ello construido en relación con un contexto de utilización.

Una definición más o menos consensuada de ontología es la propuesta por Gruber,⁶ quien plantea que es una descripción formal de conocimiento que debe ser compartido por un colectivo de usuarios e interpretado y leído automáticamente por las máquinas para garantizarle al usuario una efectiva recuperación. Otra de las definiciones compartidas por la comunidad académica es la de Borst⁷, quien

-
6. Definición propuesta en 1991, citada en la mayoría de los textos consultados. Un escrito de referencia sobre recientes apreciaciones y conceptualizaciones del autor sobre las ontologías se encuentra en una entrevista publicada en la revista Official Quarterly Bulletin of AIS Special Interest Group on Semantic Web and Information Systems. 2004, vol.1, Issue 3, [En línea]. Disponible en: <http://tomgruber.org/writing/sigsemis-2004.pdf> [Consulta: 12 de septiembre, 2008]
 7. LAMARCA LAPUENTE, María Jesús. Ontologías. En: *Hipertexto, el nuevo concepto de documento en la cultura de la imagen* [En línea]. 2008. Disponible en: <http://hipertexto.info/documentos/ontologias.htm> [Consulta: 18 de septiembre de 2008]

define una ontología como «la especificación formal de una conceptualización compartida». Otra, propuesta por Neches⁸, define una ontología como un vocabulario acerca de un dominio, y cuya extensión está dada por los términos, relaciones y reglas de combinación del lenguaje del dominio. Para Otros autores, como Weingad,⁹ una ontología es una base de datos que describe conceptos generales.

Estas definiciones permiten identificar los conceptos y elementos comunes de una ontología: el lenguaje, un dominio particular de conocimiento, la representación de conocimiento y las relaciones establecidas entre los términos que representan conceptos. Y todo ello con fines de recuperación de información, principalmente en la Web. Más que el vocabulario en sí mismo, en la ontología son esenciales los conceptos asociados a los términos, de ahí su papel central en el desarrollo de la Web semántica.

En resumen, y en palabras de Lozano Tello, ontología es una herramienta que define conceptos y relaciones propias de un dominio, de forma compartida y consensuada, representada de manera formal, legible y utilizable por el computador.¹⁰

Las ontologías son herramientas útiles para la comunicación entre expertos en un área determinada. Pueden ayudar en la generación de consenso entre especialistas sobre el vocabulario técnico que es usado en las interacciones. De igual forma, pueden ser usadas para la formalización del conocimiento eliminando las contradicciones e inconsistencias en un dominio para que sean automáticamente verificadas y validadas por computadores. Finalmente, como representación de conocimiento, las ontologías conforman un vocabulario de dominios que puede ser reutilizado en distintas aplicaciones y para propósitos específicos.

En esta investigación se optó por la realización de una ontología, pues es un recurso para la documentación, que proporciona la estructura conceptual de un campo de dominio, lo que lo hace muy eficiente en la generación de herramientas de indización, búsqueda y recuperación de información.¹¹

Al comparar las definiciones pareciera que el paralelismo entre ontologías y tesauros es evidente: ambos son instrumentos para la recuperación e indización de información, ambos definen los términos de un dominio particular y los relacionan.

.....
8. *Ibid.*

9. *Ibid.*

10. LOZANO TELLO, Adolfo. *Ontologías en la Web semántica*. [En línea]. Disponible en: <http://www.informandote.com/jornadasIngWEB/articulos/jiw02.pdf> [Consulta: 18 de septiembre de 2008]

11. ARANO, Silvia. La ontología: una zona de interacción entre la Lingüística y la Documentación. [En línea]. *Hipertext.net*. 2003, no. 2. Disponible en: <http://www.hipertext.net> [Consulta: 12 de febrero 2009].

Sin embargo, como lo señala Antonio García Jiménez, hay diferencias sustanciales entre los lenguajes documentales tradicionales y los nuevos (entendiendo las ontologías desde la documentación, como un lenguaje documental). Los cambios se producen en tres niveles: una mayor aproximación al lenguaje natural, modificaciones en las reglas morfosemánticas y preponderancia de las relaciones asociativas.¹²

En esta investigación no se habla en sentido estricto de ontologías desde el punto de vista de la informática, puesto que la investigación no tuvo inicialmente ese alcance, aunque es de reconocer que a partir del trabajo interdisciplinar con la ingeniería de sistemas y la aplicación e implementación de herramientas y lenguajes propios del ámbito informático, se llegó a una combinación metodológica y teórica de las ontologías desde la documentación, la terminología y la ingeniería.

1.1 Diseño e implementación de la ontología

En la construcción e implementación de la ontología hubo tres momentos, en cada uno de los cuales se integraron y adaptaron diferentes metodologías propuestas desde la terminología, la documentación y la ingeniería informática.

En una primera etapa, de orden conceptual, se siguieron los pasos para estructurar el árbol de dominio e identificar el discurso particular sobre el desplazamiento forzado. Un segundo momento, de tratamiento terminológico, tuvo que ver con la confección del tesauro y la estructuración de las relaciones jerárquicas, para lo cual se usaron herramientas como MultiTes para la edición y gestión de tesauro. Y en una tercera etapa, que tuvo que ver directamente con la implementación de la ontología y el sistema de información, se siguieron las metodologías propuestas por Noy y Macguiness¹³ y la Universidad Politécnica de Madrid¹⁴ para la estructuración de ontologías y la aplicación de diferentes lenguajes y herramientas informáticas en la implementación del sistema.

-
12. GARCÍA JIMÉNEZ, Antonio. Instrumentos de representación del conocimiento: tesauros versus ontologías. *Anales de Documentación*. 2004, no. 7; p. 85-86.
13. Entre las metodologías propuestas para la construcción y desarrollo de las ontologías se encuentra la sugerida por Noy y McGuinness que proponen el desarrollo de un modelo de ontologías a partir de los siguientes pasos: determinar el dominio, considerar la utilización de ontologías existentes, enumerar los términos importantes de la ontología, definir las clases y las jerarquías, definir las propiedades de las clases, definir las facetas de los slots y crear las instancias. Véase: NOY, Natalya F.; MCGUINNESS, Deborah L. *Ontology Development 101: A Guide to Creating Your First Ontology*. 2001 [En línea]. Disponible en: http://protege.stanford.edu/publications/ontology_development/ontology101.pdf [Consulta: 18 de septiembre de 2008]
14. Methontology, metodología desarrollada en la Universidad Politécnica de Madrid, sigue el proceso de desarrollo de la ontología y permite su permanente actualización, considerando los siguientes pasos: especificación, conceptualización, formalización, implantación y mantenimiento. Véase: FERNANDEZ, M.; GÓMEZ PÉREZ, A.; JURISTO, N. *Methontology: from ontological art towards ontological engineering*. [En línea]. Disponible en: http://www.l2f.inesc-id.pt/~joana/prc/artigos_06c%20METHONTOLOGY%20from%20Ontological%20Art%20towards%20_Ontological%20Engineering%20-%20Fernandez,%20Perez,%20Juristo%20-%20AAAI%20-%20201997.pdf [Consulta: 18 de septiembre de 2008]

Como se dijo anteriormente, la ontología sobre desplazamiento forzado en Colombia combina metodologías tanto de la ingeniería informática, para la construcción de ontologías propiamente dichas, como aquellas usadas en la terminología y la documentación para la construcción de lenguajes documentales. Por ello, el resultado es un híbrido entre la informática y la documentación, en donde se combinan la metodología sugerida desde el área de la ingeniería por Noy y McGuinness, y la Universidad Politécnica de Madrid, y la propuesta por Cardero y Cabré para la confección de productos terminológicos.¹⁵

1.1.1 Identificación del dominio

Para estructurar conceptualmente la ontología se seleccionó un corpus de textos producidos por algunas de las principales instituciones que en Colombia trabajan el tema del desplazamiento forzado: ACNUR, OIM, CODHES, Universidad de Antioquia, Universidad Nacional, Universidad Javeriana. De estas instituciones se logró reunir un total de 48 textos, que conforman el corpus textual de análisis en la investigación.

Derivado del análisis de dicho corpus documental se obtuvieron los términos que componen el fenómeno. Debido a lo vasto del tema y a las múltiples aristas que sobre el fenómeno se presentaban, era necesario elaborar un mapa conceptual sobre el mismo que sirviera inicialmente de guía para la comprensión del fenómeno y sus interrelaciones, pero también de estructura básica de la ontología. Es así como se hizo una lectura analítica de los textos más representativos de cada una de las instituciones, rastreando allí las distintas hipótesis interpretativas, las características propias del desplazamiento, los actores involucrados, las acciones políticas y militares de esos actores, la intervención estatal en el fenómeno y la acción social derivada de las formas organizativas de los desplazados.

Ello dio como resultado un mapa conceptual inicial que intentaba abarcar todo el fenómeno pero sin desarrollar cada uno de sus elementos; era, simplemente, una guía que permitía la ubicación de los términos de acuerdo con una clasificación inicial. A medida que la investigación avanzaba, el mapa conceptual iba perfilándose y llenándose de contenido. Con base en dicho mapa se realizaron las validaciones de los términos con los especialistas (6 sesiones), pues resultaba mucho más fácil para ellos visualizar conceptualmente el fenómeno a través de los mapas gráficos que con base en los listados de términos y sus relaciones. A medida que se analizaban los textos, el mapa lentamente iba ampliando sus categorías y relaciones y, como resultado del encuentro con los especialistas, muchos términos fueron descartados

15. CARDERO GARCÍA, Ana María; CABRÉ CASTELLVI, María Teresa. *Terminología y procesamiento*. México: Universidad Nacional Autónoma de México, 2003.

y otros nuevos salieron a flote y fueron incorporándose al mapa definitivo. Finalmente el mapa se estructuró a partir de cuatro facetas: causas del desplazamiento forzado, características del desplazamiento, políticas públicas e impactos.

1.1.2 Análisis lexicográfico del corpus

Para el análisis del corpus textual y la extracción de los términos propios del dominio, se usó el programa Cratilo 2, software desarrollado por los profesores Jorge Antonio Mejía, Francisco Javier Álvarez y John Albeiro Sánchez, del Instituto de Filosofía de la Universidad de Antioquia. Cratilo permite el análisis lexicográfico de textos, identificando las palabras que aparecen, su frecuencia y su ubicación en el texto. Mediante un proceso de reconocimiento, Cratilo identifica todas las palabras presentes en el texto y construye una base de datos que se convierte en el proyecto de análisis de la obra. En la investigación, cada proyecto lo constituyía el conjunto de textos de cada una de las instituciones seleccionadas.¹⁶

La información que se almacenaba en la base de datos de Cratilo podía ser exportada a formatos de texto y de hoja electrónica utilizando plantillas preestablecidas. Luego de una depuración de estas formas gráficas, eliminando términos no representativos como cifras, conectores, artículos, etc., se seleccionaron aquellas palabras con un valor semántico para los fines de la investigación, es decir aquellas que representaban algún concepto perteneciente a alguna de las categorías definidas en los mapas conceptuales.

1.1.3 Elaboración del tesoro y determinación de las estructuras jerárquicas

El producto de este corpus terminológico integró un tesoro construido en Multites (versión prueba) para establecer en principio las relaciones de tipo jerárquico y asociativo que pudieran darse entre los términos. Este tesoro es la base para el procesamiento y la traducción a lenguaje de máquina, usando el Skos-Core, que es un lenguaje RDF para la representación de tesauros y sistemas de organización de conocimiento.

Multites es una herramienta para el diseño e implementación de tesauros, que permite la sistematización de los términos con sus relaciones jerárquicas y asociativas, así como el ingreso de notas u observaciones y la clasificación en las categorías correspondientes.

.....
16. ARENAS GRISALES, Sandra. Ontología sobre el desplazamiento forzado en Colombia. En: *Memoria del I Simposio Internacional sobre Organización del Conocimiento: Bibliotecología y Terminología*. Catalina Naumis Peña, coord. México: CUIB, 2009.

Al ingresar cada término y relacionarlo con otros, automáticamente este programa permite reconocer las relaciones ya previamente establecidas que afecten otros términos. En la **Figura 9** se presentan en detalle las relaciones que es posible generar a un término.

1.1.4 Implementación del sistema de información

La ontología, en la cual se describe cada uno de los conceptos del área del conocimiento, es creada con base en el lenguaje OWL¹⁷ y como complemento Skos-Core¹⁸, el cual proporciona un marco básico para la construcción de esquemas de conceptos, pero sin la definición semántica tan estricta que exige la utilización de OWL. Skos-Core es un schema RDF que permite la representación de tesauros y sistemas para la organización del conocimiento; a su vez, sirve para la construcción de esquemas de conceptos simples para su utilización en la Web. Las relaciones básicas que proporciona Skos-Core se pueden observar en la **Figura 1**. Skos-Core provee una serie de propiedades que permiten la representación del conocimiento de una forma flexible, describiendo así el contenido y la estructura organizacional de los conceptos.

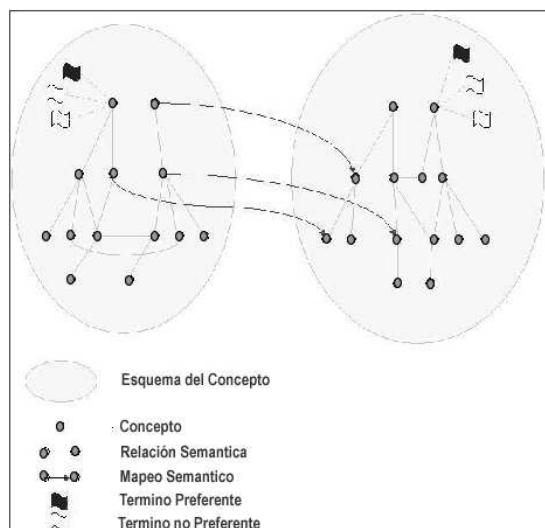


Figura 1 Relaciones básicas de Skos-Core

-
17. OWL (Ontology Web Language) es un lenguaje de etiquetado semántico para publicar y compartir ontologías, que permite definir el significado de términos en vocabularios y las relaciones entre ellos. Véase LAMARCA LAPUENTE, María Jesús. *Op cit.*
 18. SÁNCHEZ FERNÁNDEZ, L.; Fernández García, N. La Web semántica: fundamentos y breve estado del arte. *Novática*. 2005, no. 178; p. 6-11. [En línea]. Disponible en: <http://www.ati.es/novatica/2005/178/178-6.pdf> [Consulta: septiembre, 2008]

Cuando se tienen ontologías de gran tamaño (cientos o miles de conceptos y relaciones), se hace necesario ofrecer al usuario herramientas que no sólo permitan obtener vistas útiles de la ontología, sino que también posibiliten su manipulación y validación. Algunas de estas herramientas de manipulación y edición de ontologías permiten, a través de una interfaz gráfica, crear nuevos conceptos y relaciones entre éstos y exportar las ontologías a alguno de los lenguajes de representación semántica que las soporte.

Protégé, que es un programa informático de tipo open source, fue la herramienta utilizada para la edición de la ontología. Su función consiste en la creación, edición y mantenimiento de ontologías. Posee la capacidad de generar de manera automática el código fuente de cualquier ontología creada con Protégé en formato RDF/OWL.¹⁹

Una vez estructurada la ontología, se procede a la última fase: la implementación del sistema de información. En el desarrollo de la aplicación se hace uso de varios componentes necesarios para la correcta manipulación y posterior extracción de conocimiento de la ontología, que en principio es la base del software a desarrollar. El sistema es una aplicación Web, que permite la búsqueda de documentos e información referente al desplazamiento forzado en Colombia. El sistema se divide en tres módulos: módulo de conocimiento, módulo de búsqueda y módulo de visualización.

El módulo de conocimiento es la capa inferior dentro de la arquitectura del sistema. En éste son manejadas la base de contexto (descripción de los conceptos y sus relaciones) y la base documental (descripción de documentos digitales).

La *base de contexto*, como primer punto en el desarrollo de la ontología de Skos-Core, fue implementada con Protégé, describiendo claramente cada uno de los conceptos pertenecientes al área del conocimiento, sus propiedades y relaciones; a su vez se estructuraron las categorías o facetas en las cuales se agrupan los términos.

Para tener una mayor aproximación a la ontología de Skos-Core y las propiedades que éste pone a disposición, se puede observar la representación que se realiza en una de las colecciones: *Conflictos armados*, y algunos de los miembros que pertenecen a ésta, usando así propiedades tales como skos:prefLabel para los descriptores; skos:altLabel, para la relación de equivalencia o sinonimia, y skos:narrow y skos:broader para las relaciones de jerarquía (ver **Figura 2**).

19. NOY, Natalya F.; MCGUINNESS, Deborah L. *Op cit.*

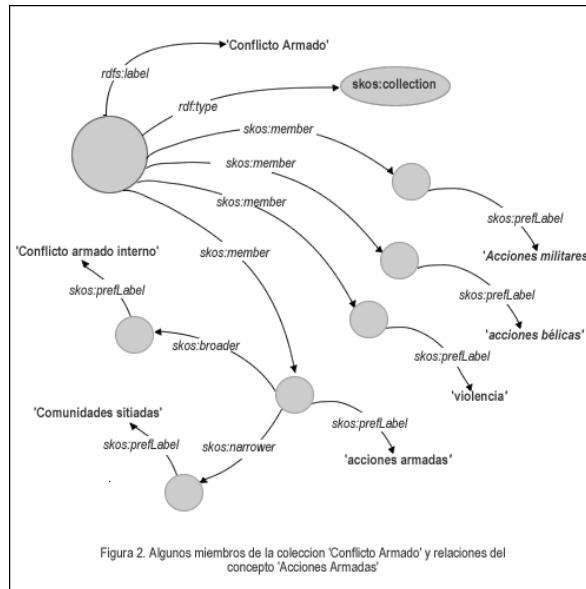


Figura 2 Relaciones de jerarquía usando Skos-Core

Con la ayuda de la herramienta Protégé se facilitó la creación y la edición de los términos, permitiendo así sacar el máximo provecho a la capacidad de representación que provee la ontología de Skos-Core con cada una de sus propiedades. En la **Figura 3** se observa la representación del término *Estrategias de guerra*, con cada una de sus propiedades.

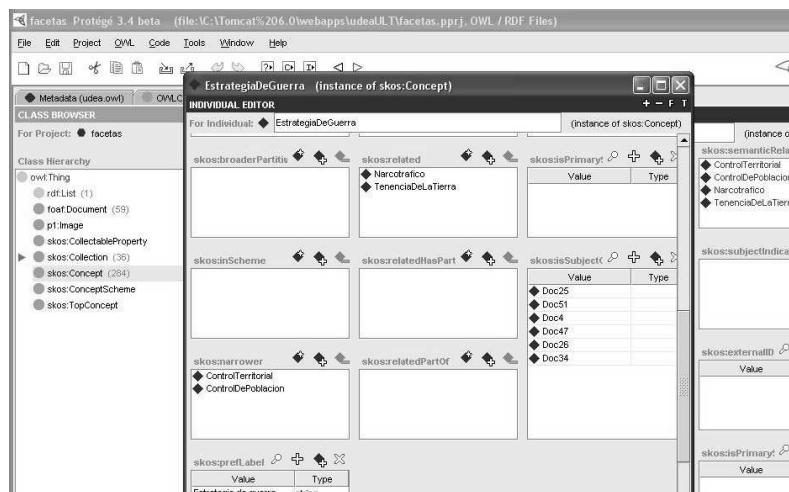


Figura 3 Representación del concepto *Estrategias de guerra*, en Protégé.

La *base documental* almacena las características de los documentos manejados por el sistema y posee atributos o metadatos de descripción tales como título, autor, formato, entre otros. Éstos son almacenados para enlazar cada documento con los términos asociados en la base de contexto.

Cada uno de los atributos, es decir, la información que identifica un documento, al igual que los términos que son representados en Skos-Core, tiene su representación en los lenguajes FOAF²⁰ y Dublin Core.²¹ Estos lenguajes facilitan la identificación de los documentos. Algunas de estas propiedades son: foaf:document, utilizada para la creación de recursos de tipo documento; dc:creator, para referir el autor del documento; dc:source, para indicar la ruta del documento; dc:format, para describir el formato, entre otras.

A continuación se muestra una propiedad para recuperación de documentos en el Sistema: skos:subject. Esta propiedad indica los términos que están asociados a un documento. En la **Figura 4** se presenta un ejemplo del término *Estrategias de guerra* asociado al documento «Actores armados y sus estrategias de guerra»

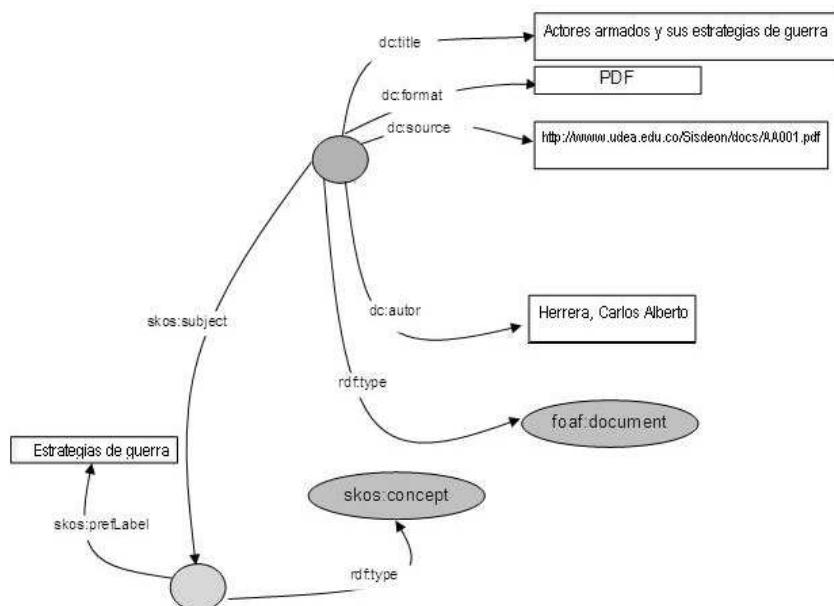


Figura 4 Representación de un documento.

20. *The Friend of a Friend (FOAF) project* [En línea] Disponible en: <http://www.foaf-project.org/> [Consulta: 20 de enero de 2008]
21. *The Dublin Core Metadata Initiative* [En línea] Disponible en: <http://dublincore.org/> [Consulta: 28 de enero de 2008]

El *módulo de búsqueda* permite la recuperación de información documental. La recuperación puede efectuarse a partir de una búsqueda alfabética, simple (término libre) y compleja (asociando varios términos a la vez). Para ello el sistema se vale de las etiquetas asignadas a los términos, tales como prefLabel, altLabel y hiddenLabel.

La búsqueda alfabética ofrece el listado de todos los términos. Para ello hace uso de la etiqueta Skos:Collection. La búsqueda libre le permite al usuario acceder a los documentos a partir de la digitación de un término concreto; y la búsqueda compleja permite seleccionar múltiples términos y conjugarlos con operadores booleanos (AND y OR) y elaborar así ecuaciones de búsqueda en un nivel mayor de complejidad, con lo cual puede obtener resultados más precisos.

Las consultas y las búsquedas son realizadas con la ayuda de SPARQL y Pellet. SPARQL es utilizado para formular la consulta y Pellet para razonar sobre la ontología y encontrar los documentos que correspondan con los criterios de búsqueda.

El *módulo de visualización* permite mostrar la ontología al usuario a través de la Web. Para ello se usó JSP (Java Server Pages), por ser una tecnología de servidor eficaz y una base para desarrollar aplicaciones Web, ya que utiliza Java para implementar la lógica compleja que requiere el manejo de bases de datos²². Gracias a los *servlets*, que son una estructura que sirve a JSP²³, se pueden hacer páginas dinámicas, y éstas actúan como interfaces de aplicación que también procesan el almacenamiento de datos. Éstas, junto con Java y su tecnología JDBC para el manejo de sistemas de administración de bases de datos, son una alternativa para que el usuario pueda visualizar los razonamientos hechos por el Pellet. Para tener una idea más clara de la presentación que se tiene al momento de consultar por un término específico, se presenta en la **Figura 5** el término *Estrategias de guerra*.

.....
22. *JavaServer Pages Technology*. [En línea] Disponible en: <http://java.sun.com/products/jsp/index.jsp> [Consulta: 18 de septiembre de 2007]

23. HANNA, Phil. *JSP Manual de referencia*. España: McGraw-Hill, 2002.

Sistema de información
documental sobre desplazamiento forzado en
Colombia

EIB
INSTITUTO DE ESTUDIOS POLITICOS

Quienes somos | Ontología | La investigación | Mapas | Contacto | Inicio

Término : Estrategia de guerra

Faceta: A2 Estrategia de guerra

Término(s) específico(s): Control de población, Control territorial

Término(s) relacionado(s): Narcotráfico, Tenencia de la tierra

Nombre(s) alternativo(s): Métodos de guerra

Definición: Si bien el Derecho Internacional Humanitario utiliza preferentemente el término métodos de guerra, esta ontología utilizará el término estrategias de guerra, entendiendo por éste las armas y medios armados, las tácticas o estrategias utilizados contra el enemigo durante el conflicto.

Figura 5 Visualización del término *Estrategias de guerra*

1.2 Validación con los especialistas

En cada una de las etapas se realizaron sesiones de trabajo con especialistas en el tema, para validar los términos que surgían del proceso, sus definiciones y relaciones. La validación se hizo sobre la base de mapas conceptuales, pues se consideró que era mucho más clara la relación entre los términos cuando se presentaban en forma gráfica.

Las sesiones tuvieron una duración aproximada de dos horas. En los primeros 15 minutos se hizo una breve presentación de la investigación y sus propósitos y a continuación se expuso el objetivo del encuentro y la metodología que se iba a aplicar.

Se le entregó a cada asistente una copia de los mapas, diferenciándolos por facetas, para que durante unos quince minutos los observaran y proceder a la discusión. El propósito fundamental era que los asistentes pudieran deducir e interpretar los mapas antes de que fueran explicados, con el fin de corroborar si su estructura era lo suficientemente clara para representar gráficamente el fenómeno del desplazamiento forzado y sus implicaciones y relaciones. La única aclaración que se hizo fue acerca de ciertos conectores y convenciones, la ubicación de los términos y sus relaciones. Se les dijo que firmaran sus copias y rayaran y escribieran sobre ellas sus observaciones, lo que sirvió luego de insumo para la elaboración del acta de la sesión.

A continuación se dio inicio a las discusiones. Sin excepción, los expertos comprendieron rápidamente que se trataba de una conceptualización del fenómeno que pretendía abarcar la mayoría de sus aspectos. En algunos casos pidieron explicaciones sobre las relaciones y jerarquías que se establecieron, sin embargo se consideraron siempre los cambios y adiciones que sugirieron.

Fueron sesiones de trabajo donde se evidenció la pluralidad de opiniones, pero también donde finalmente se logró llegar a consensos importantes que permitieron confiar en que el resultado final sería acogido por los otros expertos.

Para cada nueva sesión se presentaron los mapas elaborados inicialmente y, después de la discusión, los mapas producto de las sesiones pasadas, en procura de consensos o disensos con relación a los cambios o las adiciones propuestas por los demás expertos.

En general puede afirmarse que todos coincidieron en resaltar la importancia de este tipo de ejercicio, pues en muchos casos cada uno de ellos era especialista en algún aspecto y reconocía la existencia de los demás, pero pocas veces se habían tomado el trabajo de realizar mapas conceptuales que abarcaran el fenómeno en toda su dimensión.

2. Resultados obtenidos

Como resultado de esta investigación se obtuvieron básicamente dos productos: uno, la ontología que constituye la estructura o base conceptual del dominio, que define cada uno de los conceptos y los relaciona. El otro es el sistema de información, que, teniendo como estructura la ontología, proporciona una base documental especializada en el tema. A continuación se describe la interfaz de consulta del Sistema de Información Documental sobre Desplazamiento Forzado en Colombia.

Para acceder a la página Web del «Sistema de información documental sobre desplazamiento forzado en Colombia» del Instituto de Estudios Políticos de la Universidad de Antioquia, es necesario ingresar a la dirección Web: <http://iep.udea.edu.co/sisdeon>. Allí el usuario encuentra una pequeña página introductoria y el enlace para acceder al Sistema como tal.

En el Sistema se pueden encontrar unas pestañas en el margen superior, donde se visualiza la información general sobre el proyecto. En el margen izquierdo se encuentra el menú principal para la navegación. Esta información permanece siempre visible para el usuario (ver **figura 6**).



Figura 6 Pantallazo inicial del Sistema de Información Documental sobre Desplazamiento Forzado

Menú de la parte superior de la página principal

Quiénes somos: incluye los datos de los integrantes del proyecto.

Ontología: se ofrece información sobre qué es una ontología y su funcionamiento dentro del sistema.

La investigación: aparece información general del proyecto y acerca de las entidades que lo financiaron.

Mapas: se pueden visualizar los mapas conceptuales, con los términos clasificados por facetas.

Contacto: esta pestaña se encuentra en desarrollo y permitirá la comunicación con los administradores del sistema de información, de tal forma que se puedan sugerir términos y relaciones entre estos que enriquezcan la Ontología.

Inicio: devuelve a la página principal del sistema.

Menú de la parte izquierda de la página principal

Facetas: al dar clic sobre esta opción, se muestran las facetas que componen la ontología, que son: *causas del desplazamiento, características del*

desplazamiento, políticas públicas, e impactos, cada una con los términos que la componen y sus relaciones. Al dar clic a cualquier término perteneciente a la faceta, se desplegará la información específica de dicho término (Ver **Figura 7**).

The screenshot shows the website's header with the logo of the Universidad de Antioquia and the EIB (Instituto de Estudios Políticos). The main title is 'Sistema de información documental sobre desplazamiento forzado en Colombia'. Below the title is a navigation bar with links: Quienes somos, Ontología, La investigación, Mapas, Contacto, and Inicio. On the left sidebar, there are links for Inicio, Términos, Facetas, Búsqueda avanzada, Ayuda, and Enlaces de interés. A search bar is also present. At the bottom of the sidebar are buttons for 'Enviar consulta' and 'Ir a término'. The central content area is titled 'Facetas' and lists categories starting with 'A1 Consecuencias del Conflicto Armado Interno' and ending with 'D4 Marco Institucional'. Each category is a link.

Figura 7 Pantallazo de las facetas de la ontología

Términos: Al dar clic se despliegan en la parte central de la página todos los términos en orden alfabético, mostrando sólo los que comienzan con la letra A, y dando en la parte superior derecha la opción de buscar por las otras letras del abecedario (Ver **Figura 8**).

The screenshot shows the same website layout as Figure 7. The central content area is titled 'Términos' and lists terms starting with 'Acceso a tierras' and ending with 'Asentamiento nucleado'. Above this list is a horizontal menu with letters from A to Z: ABCDEFGHIJKLMNOPQRSTUVWXYZ. The left sidebar contains the same navigation links as Figure 7, along with a search bar and buttons for 'Enviar consulta' and 'Ir a término'.

Figura 8 Listado de términos

Al dar clic en cualquier término, el sistema muestra su contenido, permitiendo visualizar los términos generales, específicos que tiene, así como la definición y documentos asociados (ver **Figura 9**).

The screenshot shows a search results page for the term "Estrategia de guerra". The top navigation bar includes the University of Antioquia logo, the system title, and the EIB logo. The main content area displays the term definition, associated facets, specific terms, related terms, alternative names, and a definition section. A "Documentos" button is visible on the right.

Faceta:	A2 Estrategia de guerra
Término(s) específico(s):	Control de población Control territorial
Término(s) relacionado(s):	Narcotráfico Tenencia de la tierra
Nombre(s) alternativo(s):	Métodos de guerra
Definición:	Si bien el Derecho Internacional Humanitario utiliza preferentemente el término métodos de guerra, esta ontología utilizará el término estrategias de guerra, entendiendo por éste las armas y medios armados, las tácticas o estrategias utilizados contra el enemigo durante el conflicto.

Figura 9 Visualización de la información del término

Documentos asociados a términos

Cuando un término tiene documentos asociados, aparece en la parte superior derecha de la pantalla un enlace llamado *documentos*, que lleva a un listado de textos asociados al término, proporcionando información bibliográfica del documento. También se muestran otros términos asociados al documento, que pueden servir para navegar dentro del tema y ampliar las posibilidades de búsqueda del usuario. Al dar clic en el nombre del documento este se abre en un pdf descargable.

Búsqueda de términos

En la parte izquierda de la pantalla aparece un campo para escribir el término que se desea buscar con dos opciones:

Enviar consulta: busca todos los términos que tienen la palabra ingresada y los despliega para acceder a cualquiera de ellos dando clic.

Ir a término: encuentra el término específico que se busca, desplegando directamente la información sobre éste si está bien escrito y/o existe; de lo contrario muestra las opciones y da la posibilidad de acceder a otros cercanos.

Búsqueda avanzada

La búsqueda avanzada permite seleccionar varios términos y utilizar los operadores boléanos para especificarla. Para realizar este tipo de búsqueda, se ingresa por el enlace *búsqueda avanzada*, se digita el término completamente o se selecciona de la lista que se va desplegando con los términos que corresponden con el criterio de búsqueda.

El sistema de búsqueda avanzada funciona agregando palabras a una especie de «canasta» de términos, es decir, se buscan términos y posteriormente se recogen para realizar las búsquedas avanzadas (Ver **Figura 10**).



Figura 10 Ejemplo de búsqueda avanzada

Este tipo de búsqueda tiene dos posibilidades, la búsqueda AND y OR:

AND: permite buscar documentos que tengan asociados obligatoriamente todos los términos que se insertaron para realizar la búsqueda.

OR: permite buscar documentos que tengan asociados por lo menos uno de los términos que se insertaron para realizar la búsqueda (Ver **Figura 11**).

Figura 11 Ejemplo de búsqueda con el uso de operadores

Para visualizar la búsqueda se la clic en *Enviar búsqueda OR* o *Enviar búsqueda AND*, y el sistema despliega los documentos asociados a dichos términos (Ver **Figura 12**).

Figura 12 Documentos asociados al término

Después de realizar la búsqueda, para limpiar la «canasta» de términos, damos clic en *nueva búsqueda*, así el buscador quedara limpio para continuar buscando documentos asociados a términos.

Enlaces de interés

Esta pestaña se encuentra en construcción y allí aparecerá información relativa a las instituciones gubernamentales, organismos internacionales, instituciones académicas, organizaciones no gubernamentales, medios de comunicación, otras bases de datos y fuentes de información y documentación sobre desplazamiento forzado.

3. Conclusiones y trabajos futuros

- Se construyó un Sistema de Información Documental sobre el Desplazamiento Forzado, basado en una ontología, donde se representa y define conceptualmente la terminología sobre el tema facilitando los procesos de recuperación de información en la Web. El sistema permite búsquedas semánticas sobre el dominio específico y la navegación a través de los conceptos y sus relaciones, facilitándole al usuario encontrar nuevos conceptos asociados a su pesquisa inicial.
- El proyecto representa una confluencia de las metodologías propias de la terminología y las tecnologías usadas por la Web semántica en un dominio específico. Dicha confluencia puede dar como resultado una mayor confiabilidad en la indización y recuperación de la información, debido a que la metodología aplicada por la terminología confiere validez y rigurosidad en la selección de los términos, lo cual no se explicita claramente en la metodología usada en la construcción de las ontologías.
- El Sistema hará parte del Observatorio sobre el Desplazamiento Forzado, del Instituto de Estudios Políticos, como complemento a la información cualitativa y cuantitativa que se proporciona. Está dirigido a todo tipo de público, desde investigadores hasta público en general, quienes encontrarán un repositorio de documentos y conocimiento sobre el tema bien estructurado, estandarizado, formalizado y validado por un grupo de expertos en el tema.
- Se espera difundir la ontología entre documentalista y encargados del proceso de análisis de información en centros de documentación y bibliotecas académicas, con el fin de normalizar el lenguaje documental propio del desplazamiento y facilitar la recuperación de la información en otras bases de datos.
- Se propone extender la aplicación de la ontología a otros dominios conexos que respondan a las líneas de investigación que desarrolla actualmente el Instituto de Estudios Políticos.

Referencias bibliográficas

1. ARIAS ORDOÑEZ, José. Web semántica: significado, componentes y funcionamiento. *Bibliotecas y Tecnologías de la Información*. Jul.-Sep. 2006, vol. 3, no. 3; p. 5-12, 14-16.
2. CARLAN, Eliana. *Ontología e Web semántica*. Brasilia: 2006. (Tesis de grado. Faculdade de Economia, Administração, Contabilidade e Ciência da Documentação e Informação, Departamento de Ciência da Informação e Documentação, Curso de Graduação em Biblioteconomia).
3. CASTELLS, Pablo. *Aplicación de técnicas de la Web semántica*. Madrid: Escuela Politécnica Superior, Universidad Autónoma de Madrid. [En línea]. Disponible en: <http://arantxa.ii.uam.es/~castells/publications/coline02.pdf> [Consulta: 28 de febrero de 2008]
4. CASTELLS, Pablo y SAIZ, Francisco. Ontologías para la Web semántica. En: *La Web semántica: tecnologías y aplicaciones*. Madrid: Escuela Politécnica Superior, Universidad Autónoma de Madrid, 2003.
5. CURRÁS, Emilia. *Ontologías, taxonomía y tesauros: manual de construcción y uso*. España: Trea, 2005.
6. FERNÁNDEZ BREIS, Jesualdo Tomás. *Un Entorno de Integración de Ontologías para el Desarrollo de Sistemas de Gestión de Conocimiento*. España: 2003. (Tesis de doctorado. Universidad de Murcia, Departamento de Ingeniería de la Información y las Comunicaciones).
7. GARCÍA JIMÉNEZ, Antonio. Instrumentos de representación del conocimiento: tesauros versus ontologías. *Anales de Documentación*. 2004, no. 7; p. 79-95.
8. GARCÍA MARCO, Francisco Javier. Ontologías y organización del conocimiento: retos y oportunidades para el profesional de la información. *El Profesional de la Información*. Nov.-Dic. 2007, vol. 16, no. 06; p. 541-550.
9. GIRALDO, Juan D.; GUZMÁN, Jaime. Desarrollo de un módulo de visualización y navegación de información basada en ontologías para el sistema de recuperación semántica SABIOS. *Revista Avances en Sistemas e Informática*. Jun. 2006, vol. 3, no. 1; p. 57-61.
10. GUZMÁN LUNA, Jaime A.; TORRES PARDO, Durley. Sabios: una aplicación de la Web semántica para la gestión de documentos digitales. *Revista Interamericana de Bibliotecología*. Ene.-Jun. 2007, vol. 30, no. 1; p. 51-72.

11. HERNÁNDEZ RAMÍREZ, Haliuska; SAIZ NOEDA, Maximiliano. Ontologías mixtas para la representación conceptual de objetos de Aprendizaje. *Procesamiento del lenguaje natural*. 2007, no. 38; p. 99-106. [En línea]. Disponible en: <http://dialnet.unirioja.es/servlet/articulo?codigo=2281146> [Consulta: 23 de febrero de 2009]
12. LORENTE CASAFONT, Mercè. Ontología sobre economía y recuperación de información. *Hipertext.net*. 2005, no. 3. [En línea]. Disponible en: <<http://www.hipertext.net>> [Consulta: 23 de febrero de 2009].
13. MOREIRA, Alexandra. *Tesauros e ontologias: estudo de definições presentes na literatura das áreas das ciências da Computação e da informação, utilizando-se o método analítico-sintético*. Belo Horizonte: 2003. (Tesis de Maestría. Escola de Ciência da Informação, Universidade Federal de Minas Gerais) [En línea]. Disponible en: <http://www.r020.com.ar/recursos.php>tag=literatura> [Consulta: junio 15 de 2009]
14. PEDRAZA JIMÉNEZ, Rafael; CODINA, Lluís. Web semántica y ontologías en el procesamiento de la información documental. *El Profesional de la Información*. Nov.-Dic. 2007, vol. 16, no. 6; p. 569-578.
15. PEIS REDONDO, Eduardo *et al.* *Ontologías, metadatos y agentes: recuperación «semántica» de la información*. [En línea]. Disponible en: <http://www.nosolousabilidad.com/hassan/jotri2003.pdf> [Consulta: 23 de febrero de 2009]
16. PÉREZ AGÜERA, José Ramón. *Automatización de tesauros y su utilización en la Web semántica*. Madrid: Universidad Complutense de Madrid, Departamento de Biblioteconomía y Documentación, 2004.
17. PINTO MOLINA, María. *Organización, filtración y representación de la información*. [En línea]. 2004. Disponible en: http://www.mariapinto.es/e-coms/or_con_elect.htm#or1 [Consulta: 18 de febrero de 2009]
18. RAMOS, Esmeralda, *et al.* Aplicación de visualización de una ontología para el dominio del análisis del semen humano. *Ingeniería y Ciencia*. Jun. 2007, vol. 3, no. 5; p. 43-66.
19. REDONDO, Ana Belén; ROBLEDO, Montse y ROMERO, Mari Carmen. *Recuperación de información: ontologías*. [En línea]. Disponible en: http://es.geocities.com/recupdeinformacion_ontologias/sobreontoligias.htm#sumario [Consulta: 18 de febrero de 2009]
20. RODRÍGUEZ PEROJO, Keilyn; LEÓN, Rodrigo Ronda. *Web semántica: un nuevo enfoque para la organización y recuperación de información*

en el Web. [En línea]. Disponible en: http://www.r020.com.ar/paperlandia/detalle.php?id_handler=6680 [Consulta: 18 de febrero de 2009]

21. RUBERT LIBRELOTTO, Giovani *et al.* *TM-Builder: Um Construtor de Ontologias baseado em Topic Maps.* [En línea]. Disponible en: <http://www.di.uminho.pt/~jcr/XML/publicacoes/artigos/2003/clei.pdf> [Consulta: 18 de febrero de 2009]
22. SÁNCHEZ JIMÉNEZ, Rodrigo; GIL URDICIAIN, Blanca. Lenguajes documentales y ontologías. *El Profesional de la Información.* Nov.-Dic. 2007, vol. 16, no. 6; p. 551-560.
23. TALAVERA PEREIRA, Rosalba; MARCANO AULAR, Yelitza Josefina. Los lenguajes de representación semántica y su uso en la construcción de ontologías. *RCS; Revista de Ciencias Sociales.* Ene.-Abr. 2007, vol. 13, no. 01; p. 59-71.
24. TELLO LOZANO, Adolfo. Ontologías en la Web semántica. En: *I Jornadas de Ingeniería Web '01.* [En línea]. Disponible en: <http://www.informandote.com/jornadasIngWEB/articulos/jiw02.pdf> [Consulta: 18 de febrero de 2009]
25. TRAMULLAS SAZ, Jesús. *Agentes y ontologías para el tratamiento de la información: clasificación y recuperación en Internet.* Zaragoza: Universidad de Zaragoza, Departamento Ciencias de la Documentación [En línea]. Disponible en: <http://tramullas.com/papers/isko99.pdf> [Consulta: 18 de febrero de 2009]
26. SÁNCHEZ CUADRADO, Sonia; MORATO LARA, Jorge Luis. De repente, ¿Todos hablamos de ontologías? *El Profesional de la Información.* Nov.-Dic. 2007, vol. 16, no. 6; p. 562-568.

