

Modelo para la evaluación de ontologías. Aplicación en Onto-Satcol

Jose A. Senso*, Amed Abel Leiva-Mederos**, Sandor Eduardo Domínguez-Velasco**

Resumen: Se analizan los referentes teóricos y conceptuales de la evaluación de ontologías para conocer los procedimientos utilizados en la evaluación de estos sistemas y establecer nuevas pautas para calibrar el sistema ontológico empleado por el programa Satcol, especializado en Ingeniería de Puertos y Costas. En el trabajo se describen las características de la ontología Onto-Satcol y se evalúa la misma mediante el uso de varios indicadores (léxicos, de recuperación de información y de la estructura sintáctica). Mediante un experimento llevado a cabo por 6 expertos, y con la ayuda de la herramienta Protex, se identifican inconsistencias semánticas, estructurales y errores en la organización del conocimiento de dicha ontología.

Palabras clave: Ontologías, evaluación de ontologías, ontologías de dominio, análisis, evaluación de ontologías, Ingeniería de puertos y costas.

A model for ontology evaluation. The case of Onto-Satcol

Abstract: *This paper analyzes the conceptual and theoretical framework for the evaluation of ontologies, in order to understand the procedures used in the evaluation of these systems and to establish new guidelines for evaluating the system employed by the ontological program, SATCOL, that specializes in Port and Coastal Engineering. This paper describes the characteristics of the Onto-SATCOL ontology and evaluates it by using several indicators (lexical, information retrieval, and syntactic structure). Through an experiment conducted by six experts aided by the tool, Protex, semantic and structural inconsistencies are identified, as are errors in the ontology's organization of knowledge.*

Keywords: *Ontologies, Ontology Evaluation, Domain ontologies, Ontological analysis and evaluation, port engineering, coastal engineering.*

1. Introducción

En la representación del conocimiento es cada vez más común el uso de las ontologías como herramientas que permitan relacionar los conceptos que aparecen dentro de un área de conocimiento. Estas especificaciones formales cuentan con varios mecanismos y fases para ser construidas, así como múltiples proyec-

* Departamento de Biblioteconomía y Documentación. Universidad de Granada.

** Universidad Central «Marta Abreu» de Las Villas, Cuba. Correo-e: jsenso@ugr.es; amed@uclv.edu.cu

Recibido: 25-07-2010; 2.ª versión: 30-07-2010; aceptado: 20-10-2010.

tos y publicaciones que respaldan la validez de todas las metodologías propuestas. En general, de todas ellas se puede concluir lo mismo: no hay un único camino para construir una ontología, pero el producto resultante debe ser consecuente con los orígenes que lo propiciaron.

Al mismo tiempo, y de casi forma natural, surge la necesidad de evaluar lo realizado. La mayoría de metodologías empleadas para la creación de las ontologías también incluyen un apartado dedicado a la evaluación del resultado. Pero, en casi todos los casos, esta evaluación está excesivamente mediatizada por el sistema empleado en la formalización [por ejemplo, *methontology* (Fernández-López y otros, 1997)] o, directamente, no informa sobre cómo debe llevarse a cabo, pero sí indica la necesidad de ello [como sucede con Uschold (Uschold y King, 1995)], o es demasiado subjetiva [caso, por ejemplo, de la propuesta de Gruninger (Gruninger y Fox, 1995)].

Por otra parte, hay que recordar que una de las claves del éxito de las ontologías se encuentra en la posibilidad que ofrecen para su reutilización. Es decir, emplear total o parcialmente otra/s ontología/s para adaptarla/s a la nueva que se está creando. Si bien es cierto que cada ontología responde a una manera diferente de contemplar o comprender el estado de un conocimiento concreto, no es menos evidente que los conceptos, conceptos son, y siempre es posible emplear sólo elementos determinados en función de las necesidades (capturar la taxonomía principal o tratar la ontología origen únicamente como base terminológica, obviando otras propiedades, axiomas, instancias, etc., para la ontología destino son las dos soluciones más comunes). Sea cual sea la opción, siempre es necesario realizar una evaluación de las ontologías candidatas con el fin de averiguar cuál se adapta mejor a los requerimientos.

En la actualidad existen diversos mecanismos o métodos que permiten evaluar ontologías (Evermann y Fang, 2010; Staab, 2004). Los trabajos de Gómez-Pérez (Gómez Pérez, 1994) apuntan hacia la existencia de varias formas de evaluación de recursos ontológicos desde el punto de vista industrial, cognitivo y con el empleo de técnicas matemáticas generalmente conocidas como métricas. Varios proyectos de nivel ontológico (Yu y otros, 2009) han sido objeto de evaluación mediante el empleo de diversas técnicas y presupuestos teóricos, encaminados a evaluar esencialmente la expresión, la precisión, los problemas de diseño y las inconsistencias semánticas.

Una de las aplicaciones que se han desarrollado en el campo de la organización del conocimiento por medio de ontologías es Satcol (Leiva-Mederos y otros, 2009). Se trata de una aplicación que permite realizar procesos de minería de texto con los documentos de un corpus concreto, facilitando la creación de relaciones de acuerdo a la estructura gramatical y conceptual que presenten. Este programa emplea a un conjunto de autómatas que se encargan de la modelación, el análisis documental y la simulación del comportamiento de resumidores. Para llevar a cabo ese trabajo se emplea Onto-Satcol, ontología formada por una estrategia cognitiva y otra semántica que sirve para expresar las relaciones semánticas, interpersonales y de guía en el trabajo de los agentes.

Al ser Satcol una herramienta para representar información en el terreno de la Ingeniería de Puertos y Costas, provoca que la ontología resultante sea «diferente» al resto en cuanto a diseño y desarrollo. El hecho de no contar esta disciplina con diccionarios especializados y la ausencia de una teoría específica debido a la juventud de esta parcela del conocimiento, obligan a buscar métodos diferentes para la evaluación de la ontología. Onto-Satcol no posee los cánones clásicos (a nivel conceptual) de diseño ontológico como otros modelos de ontologías. Su estructura responde a la desambiguación del conocimiento generado en disímiles terrenos, entre los que se encuentran: medio ambiente, costas, puertos, ingeniería de caminos y canales.

Con Onto-Satcol se pretendió solucionar el problema de integrar los conocimientos que eran necesarios para la ingeniería de Puertos y Costas y contribuir así a la construcción de un sistema sumista que se apoyara en la semántica de la ontología para extraer, desambiguar y recuperar información.

Este artículo tiene como objetivo mostrar los pasos seguidos en la evaluación de la ontología Onto-Satcol, empleando para ello algunas métricas y principios específicos desarrollados pensando en las características concretas de este sistema. De esa forma, además de medir su eficacia, estaremos mostrando una metodología diferente para evaluar ontologías que se puede unir a las comentadas anteriormente.

2. Los métodos de evaluación de ontologías

Los modelos de evaluación de ontologías poseen diversas aplicaciones, de acuerdo a la ontología en cuestión que se esté analizando. Las formas de evaluar los sistemas ontológicos se han basado esencialmente en la semántica externa de las ontologías, la precisión de los modelos cognitivos, la estructura del conocimiento representado así como en medidas basadas en los criterios de exhaustividad y precisión (Kent, 1955) clásicos de la recuperación de información.

Los autores han decidido agrupar los métodos esenciales de evaluación en función a la filosofía principal que motiva dicha evaluación (básicamente recuperar información, costo que supone su implementación, uso de métricas matemáticas y análisis de la taxonomía creada) para, después, pasar a describir el modelo que servirá de base a la ontología Onto-Satcol. Teniendo en cuenta que el objetivo de este texto no es el de establecer una clasificación de sistemas y métodos, algo que por otra parte ya ha sido propuesto en otros artículos que tratan ese tema de manera monográfica, como los de Brank (Brank y otros, 2005), Obrst (Obrst y otros, 2007) o Hartmann (Hartmann y otros, 2005), no se seguirá un esquema lingüístico o semántico en su listado, ni por supuesto un recorrido exhaustivo, sino que se desarrollará sobre parámetros de practicidad, ya que no todos los métodos han sido igual de documentados ni implementados, por eso aquí mencionaremos únicamente aquellos que cuentan con información suficiente y clara, y que, además, hayan sido empleados con éxito en proyectos concretos.

2.1. Empleo de criterios clásicos en la recuperación de información

La primera variante utiliza dos métricas estándar en la recuperación de información para valorar la ontología: precisión (proporción de material relevante recuperado —bien por medio de clases o subclases, bien por medio de objetos asociados a éstas, etc.— del total de objetos recuperados) y exhaustividad (capacidad de la ontología para recuperar objetos —clases, subclases, documentos, etc.— relevantes del total de objetos relevantes almacenados en la ontología). En la mayoría de trabajos que han aplicado estos criterios se ha contemplado tanto la dimensión semántica, inherente a cualquier ontología, como la más pura vertiente de recuperación de información.

La aportación más interesante en este tipo de trabajos la presenta Zhang (Zhang y otros, 2008) quien defiende una relación puramente semántica del concepto de precisión, dentro de la posición y evaluación de mapas ontológicos. Para evaluar dichos mapas él y sus colaboradores comparan los criterios netamente semánticos con otros valores normalizados y simétricos.

Otro aspecto que caracteriza las evaluaciones realizadas a partir de estos criterios es la multiplicidad de variantes que, tanto precisión como exhaustividad, pueden aportar, y que son comentados en diversas contribuciones, en la línea de la defendida por Ramos (Ramos y otros, 2009) que presenta una variante para calcular la exhaustividad y la precisión basada específicamente en las capacidad semántica de la ontología y su operatividad en determinado corpus. Sin embargo, las fórmulas con las que trabaja son las típicas de estos dos criterios, y los valores resultantes deberían ser completados con la medida de F, que sirve para corregir el error de la distancia que se produce en aquellos casos en los que la exhaustividad y la precisión se compensan.

Un sistema que emplea un mecanismo parecido, y que presenta ese mismo inconveniente es Evalaxon (Spyns, 2005), que trabaja de manera conjunta con la ontología y el corpus sobre el que ésta se ha generado. Emplea diversas técnicas de minería de texto para realizar una evaluación más conceptual que lingüística, ya que se preocupa más de cómo de apropiados son los términos escogidos para representar conceptos, en función de la frecuencia con la que esos términos aparecen tanto en la ontología como en el corpus.

El principal problema que presentan estos métodos de evaluación es la necesidad de un entrenamiento previo de los observadores, para que puedan valorar las conclusiones de manera objetiva. En muchas ocasiones, los resultados no son los más factibles, debido a que tanto precisión como exhaustividad son dos cuestiones que se deben interpretar desde el punto de vista de la persona que interactúa con el sistema. El no contar con un criterio formal (expresado en cantidades) o una visión semántica clara (expresada por medio de la posición que existe entre el objeto que se obtiene y lo que el usuario necesita) son los principales hándicaps del empleo exclusivo de este método.

La segunda variante emplea las posibilidades que ofrece el modelo de espacio vectorial (Salton y otros, 1975) dentro de una ontología y un corpus de tex-

to. En este modelo de evaluación, que se denomina ontología de ajuste (ontology fit) y que fue propuesto por Brewster (Brewster y otros, 2004), se emplea una ontología base, que es comparada con la ontología que se desea generar. La idea sobre la que se construye ese sistema es el de evaluar el corpus de los textos empleados en la creación de la ontología en relación con las instancias de dicha ontología, construyendo un cluster que permita valorar el nivel de consistencia semántica alcanzado.

Los resultados deberían revelar si la metodología empleada para el diseño de la ontología ha sido la idónea, o si el corpus con el que se ha trabajado cumple con los estándares correctos, pero no se centra demasiado en valorar el nivel de recuperación de información obtenido, o las inconsistencias semánticas (taxonomías mal formadas, errores en la ubicación de las clases, axiomas mal declarados, descripciones simplistas de las clases por el uso incorrecto de anotaciones, etc.). Un ejemplo claro de dificultad a la hora de aplicar esta técnica se produjo en el análisis de Wordnet (<http://wordnet.princeton.edu/>), donde existen instancias declaradas en diferentes partes de la ontología con evidente relación semántica entre ellas, pero sin que dicha relación esté explicitada de forma alguna (Fellbaum, 1998). Según este sistema de evaluación, esa ontología tenía problemas metodológicos de construcción debido a ese tipo de relaciones, ya que era detectado como un problema estructural, cuando precisamente era uno de los fuertes de esta herramienta terminológica.

2.2. Sistemas centrados en el costo

A nadie se le escapa que el proceso de construcción de una ontología, independientemente del modelo escogido (desde cero, con herramientas para la obtención automática, o semi-automática desde información textual, desde modelos entidad/relación, aprendizaje, etc.), es posiblemente el más costoso, desde todos los puntos de vista, dentro de la representación del conocimiento. Tanto es así, que existen numerosos sistemas que hacen especial hincapié en este apartado cuando se evalúa una ontología. En ocasiones, demasiado.

Teniendo en cuenta que existe una estrecha relación entre la ingeniería de software y la construcción de ontologías (dependiendo del modelo escogido para la creación se pueden llegar a solapar diversas etapas), no son pocos los autores que han defendido la idea de adoptar sistemas híbridos basados en ambos. Así por ejemplo, Ontocom (Paslaru Bontas y Mochol, 2005) usa el sistema Cocomo (Constructive Cost Model) como principio para construir un mecanismo de evaluación de ontologías. Emplea para ello la unión de tres metodologías propias de la ingeniería de procesos, como son top-down, parametric y expert-based (Boehm, 1981) para analizar el costo de construcción, mantenimiento y reutilización de la ontología teniendo en cuenta valores tales como el número de personas involucradas en la construcción, cantidad de horas dedicadas, tamaño de la ontología, etc.

El principal problema que presentan este tipo de sistemas se localiza en la total ausencia de análisis de la estructura taxonómica creada, la semántica aportada, las relaciones entre las propiedades, los axiomas, etc.

2.3. Basados en métricas

Aquí recopilamos aquellos sistemas que utilizan diferentes cálculos, generalmente con base matemática, para valorar diversos aspectos de la ontología. Éste es el caso, por ejemplo, del modelo propuesto por Yao (Yao y otros, 2005), donde se emplean un determinado número de métricas de cohesión específicamente creadas para su uso en ontologías. Están basadas en diferentes teorías matemáticas y analizan, fundamentalmente, la profundidad de la descripción en función del número de clases principales, subclases, notaciones formales, profundidad media del árbol taxonómico, etc. Para llevar a cabo su propuesta se evaluaron varias ontologías con licencia creative commons por medio de un parser xml denominado OMP (Ontology Metrics Parser). Para la fase final del experimento se empleó a 18 evaluadores con dilatada experiencia en el proceso de creación de ontologías, no tanto en la materia que era representada en dichas herramientas.

Otros sistemas de evaluación también emplean métricas diferentes, y, además, con objetivos distintos. Así, por ejemplo, OntoMetric (Lozano-Tello y Gómez-Pérez, 2004), se centra más en analizar la posible reutilización de la ontología que la calidad de la misma. Para ello utiliza la jerarquía de procesos analíticos (Analytic Hierachi Process, AHP) como criterio para adoptar procesos de negocio a los ontológicos, analizando el contenido representado en la ontología, el lenguaje con que la ontología es implementada, la metodología con la que se llevó a cabo la construcción de la ontología, el software empleado y el costo de uso del sistema.

En total, este sistema analiza 160 características diferentes, convirtiéndose en una de las metodologías más exhaustivas. Quizá el principal problema que tenga es que, en ocasiones, está demasiado centrada en la reutilización de la ontología, y no en la evaluación como herramienta para la recuperación de información. En ese sentido, cabe destacar el trabajo de Sabou (Sabou y otros, 2006), que emplea un conjunto de variables muy parecidas a las de OntoMetric, pero orientadas hacia ese aspecto. Destaca el análisis que realizó sobre el conocido motor de búsqueda Swoogle (<http://swoogle.umbc.edu/>).

2.4. Sistemas basados en el análisis de jerarquías

Mediante este análisis, basado en la comparación, se facilita el estudio de la estructura conceptual de una ontología concreta para observar los posibles errores de exhaustividad y precisión, además de otro tipo de restricciones. Emplea

como medida de evaluación el cálculo de distancias y la similaridad entre clases de las ontologías analizadas.

En esencia, compara una ontología con otra —que se considera como la «ontología ideal» o perfecta, que no necesariamente de ser de la misma temática, pero sí debe contar con similares características en cuanto a estructura y profundidad en la representación— asignando valores de 0 a 1 a cada una de las clases e instancias de ambas. Estos valores son los que se emplean para delimitar la distancia entre nodos, determinando el nivel léxico que existe entre las dos representaciones. Cuando más se aproximen a cero más iguales serán las ontologías y, por tanto, más cercana a la ontología ideal estará la creada.

Es un mecanismo parecido al empleado en OntoClean (Guarino y Welty, 2002) que, además de establecer criterios de selección de ontologías a partir de una taxonomía base, analiza la posible existencia de problemas de inconsistencia semántica y/o estructural. Se centra especialmente en las propiedades definidas en la ontología, sobre todo las relaciones «is a» y «part of», lo que en ocasiones conlleva evaluaciones demasiado simplistas si la ontología es conceptualmente muy amplia.

2.5. Otros sistemas de evaluación

En alguna de las metodologías comentadas anteriormente se hablaba de la existencia de observadores que, generalmente al final del proceso, se encargaban de aplicar los criterios de calidad. Existen otras opciones, como por ejemplo, que sobre estos observadores recaiga un peso mayor. Aunque no suele ser muy habitual, por el elevado nivel de subjetividad que implica, no es menos cierto que existen experiencias basadas en esta idea.

En el trabajo de Fernández-Breis (Fernández-Breis y otros, 2009) se utiliza un sistema basado en la ISO 9126 que, aunque está pensada para la evaluación de la calidad del software, los autores logran adaptar al mundo de las ontologías, reconvirtiendo los 10 criterios generales de esta ISO en aspectos específicos del análisis de aspectos internos, externos y de uso de las ontologías. Una vez confeccionada la plantilla, 8 alumnos de doctorado realizan la evaluación de dos versiones diferentes de una ontología especializada en biología.

Resulta evidente que, en este tipo de métodos, el peso que debe soportar la evaluación se encuentra sobre un terreno excesivamente subjetivo. Además de mencionar que criterios relacionados con la calidad y la cantidad de clases y sus relaciones no son analizados con la profundidad que se requiere. Algo que, por otra parte, no sucede con el sistema propuesto por Gangemi (Gangemi y otros, 2005), quien emplea una meta-ontología (se trata de una propuesta meta-teórica sobre cómo se debe realizar la evaluación tanto de la ontología como de las anotaciones que incluya), denominada O y una ontología específica para la evaluación y validación, llamada oQual, que es la que incluye los parámetros a aplicar.

Hemos dejado para el final los métodos experimentales llevados a cabo por Evermann (Evermann y Fang, 2010), quien evalúa la ontología desde un paradigma cognitivo, muy apegado a los principios del desarrollo de aplicaciones.

3. Errores frecuentes en el diseño de ontologías

Independientemente del esquema escogido para evaluar la ontología, lo normal es que se presente al final un listado de errores, que pueden ser achacados a su creación, implementación en el sistema de información, disonancia entre el objetivo y lo realizado, etc. La importancia de éste determinará al final el nivel de operatividad alcanzado con la misma. La tipología de errores en el diseño de ontologías va desde los fallos en la taxonomía y diseño, hasta equivocaciones en la semántica conceptual, en la postura estructural y en los axiomas que se declaran. Todas estas cuestiones son las causantes de los posibles bajos resultados cuando se realizan operaciones contra estos sistemas.

Son múltiples los errores que comprometen el desarrollo exitoso de una ontología. En el trabajo de Fahad (Fahad y Abdul Qadir, 2008) se describe un sinnúmero de anomalías en el diseño ontológico. Dichos errores de diseño y concepción pueden clasificarse como simples errores de construcción, pues evidencian problemas en el diseño de las instancias y en la construcción de las clases del sistema ontológico. Otros autores, Staab, por ejemplo (Staab, 2004), analizan los problemas de construcción de las ontologías y los ponen en relación con los resultados que acarrearán estos fallos. Recogemos a continuación la relación de errores más comunes, a partir de lo expuesto por Fahad (Fahad y Abdul Qadir, 2008), Staab (Staab, 2004), Ning (Ning y Shihan, 2006), Maedche (Maedche y otros, 2003) y nuestra propia experiencia:

- Errores de ubicación: Se producen cuando no se tienen en cuenta la disposición correcta de las clases y se solapan estas, principalmente por la existencia de cruces entre nodos y estructuras conceptuales (Ning y Shihan, 2006). Suele suceder cuando una misma clase es definida como una subclase y como una superclase al mismo tiempo en diferentes niveles de jerarquía de la ontología. Este tipo de errores se producen cuando, a la hora de construir las relaciones semánticas, no se realizan análisis lingüísticos y/o terminológicos de los conceptos, provocando una estructuración del conocimiento débil o poco coherente.
- Errores de distribución. Suele aparecer cuando, a la hora de estructurar el conocimiento, los procesos creados se realizan desde una base netamente clasificatoria. Es en ese momento cuando se puede producir una excesiva dependencia (genérico/específico o «tipo de» y «subtipo de») entre las clases y las subclases (Ning y Shihan, 2006). Como resultado de este error aparecen clases que están mutuamente atadas a subclases disjuntas (Qadir y Noshairwan, 2007). Otra inconsistencia de este problema está en la frag-

mentación de una clase en muchas subclases, sin tener en cuenta que las instancias de las subclases no necesariamente tienen que pertenecer a las subclases que se han declarado.

- Errores de inconsistencia semántica. Se produce al desarrollar una jerarquía de nodos para un concepto erróneo. Es decir, cuando en la jerarquía aparecen conceptos que no pertenecen a la clase principal. La razón principal de este error es el poco conocimiento semántico y terminológico del área del conocimiento con la que se está trabajando.
- Clases y clasificaciones incompletas. En ocasiones se presta muy poca atención a elementos muy importante en la descripción, en la anotación y en la conceptualización de la ontología, cuestión que favorece la ambigüedad y entorpece la construcción de herramientas de razonamiento. Esto repercute en clases o clasificaciones poco documentadas.
- Omisión de conocimiento disjunto. Errores que ocurren cuando existe un gran número de clases y subclases, y en su creación se obvia la inclusión de axiomas, que tienen determinado grado de disjunción. Según Qadir (Qadir y Noshairwan, 2007) estos errores aparecen en la fase de diseño debido a la falta de conexión entre el desarrollador y los usuarios, lo que hace que se obtengan resultados catastróficos.
- Omisión de conocimiento por falta de exhaustividad. Se produce cuando los constructores de la ontología no crean restricciones de integridad, declarando de forma arbitraria particiones y subclases de un mismo concepto. Con esto se deja en el aire la capacidad de exhaustividad del sistema, conduciendo a la redundancia.
- Errores de redundancia. La repetición de conceptos, lo que evidencia la falta de un plan en la confección de las taxonomías de base del sistema, como apunta Fahad (Fahad y Abdul Qadir, 2008).
- Errores en la poca especificación o delimitación de las propiedades de los componentes del sistema. Esto produce que el razonamiento se vea poco desarrollado, ya que se da una misma definición a todo, desde el punto de vista formal y conceptual.
- Errores por falta de exhaustividad en la declaración de etiquetas. Este error está muy asociado a la construcción de inferencias en las ontologías que generan su estructura usando OWL. Las etiquetas que deben ser desarrolladas o declaradas tienen que tener suficientes elementos y atributos para detallar correctamente un dominio, y contribuir a que los conceptos puedan estar asociados a determinadas propiedades. En algunos sistemas ontológicos se pasa por alto esta propiedad, lo que hace que los modelos de razonamiento se vean desarticulados. Los problemas esenciales que se describen por este error son del tipo IFPO (Omisión de las Propiedades), es decir, que si no se declaran estos valores el sistema sólo identifica un único valor para un concepto y una sola clave en la base de datos.
- Errores por no describir correctamente el conocimiento. Lo que acarrea que no se reconozca exactamente qué elemento se está declarando en

cada concepto, ya que no hay definiciones exactas para subclases y sus relaciones con los conceptos (Maedche y Staab, 2002). Esto obliga a una necesaria descripción de reglas que proporcionen criterios esenciales para la interacción con nuevos conceptos (Kuhn, 1979) que se creen en las subclases y sus relaciones, es decir, obliga a generar modelos de axiomas, complementaciones y restricciones para el trabajo con la ontología.

- Errores de redundancia en las extensiones. Aparece cuando se describe un concepto como disjunto con otro concepto de la misma jerarquía (Kuhn, 1979; Maedche y Staab, 2002).

4. Evaluación de Ontol-Satcol

Onto-Satcol es una ontología desarrollada para formalizar parte del conocimiento asociado al dominio de la Ingeniería de Puertos y Costas. Su nacimiento viene provocado por dos motivos: como herramienta para un sistema de desambiguación de textos científicos, y para dar respuesta a la necesidad de paliar el desigual aumento de información en lo que se refiere a la existencia de herramientas terminológicas que faciliten la recuperación de información (taxonomías, tesauros y ontologías) dentro de esta área.

La Ingeniería de Puertos y Costas es un ejemplo de este último problema. Aunque se trata de una disciplina relativamente nueva, mezcla de la clásica Ingeniería de Caminos, Canales y Puertos y las ingenierías oceánica y naval, la notable evolución científica que ha experimentado debida al desarrollo de nuevas técnicas de construcción, explotación y gestión de estructuras y recursos marinos no ha sido suficiente como para que exista una formalización de las estructuras de conocimiento que representen los procedimientos, definiciones, objetos y metaconocimiento que en ella se llevan a cabo.

Esto es fácilmente comprobable si se realiza una consulta simple sobre dos tipos de herramientas de ontologías, buscando por los términos —en español y su correspondiente en inglés— más representativos relacionados con cuatro disciplinas escogidas de forma aleatoria dentro de la Ingeniería (Agrícola, Geológica, Aeronáutica y Civil) a la que sumamos la Ingeniería de Puertos y Costas. El resultado será, por tanto, el número total de clases dentro de las ontologías que tratan, de una manera u otra, aspectos relacionados con dichas disciplinas y, por tanto, que permitan representar el conocimiento inherente a ellas. Para esta sencilla demostración se emplean diversas herramientas:

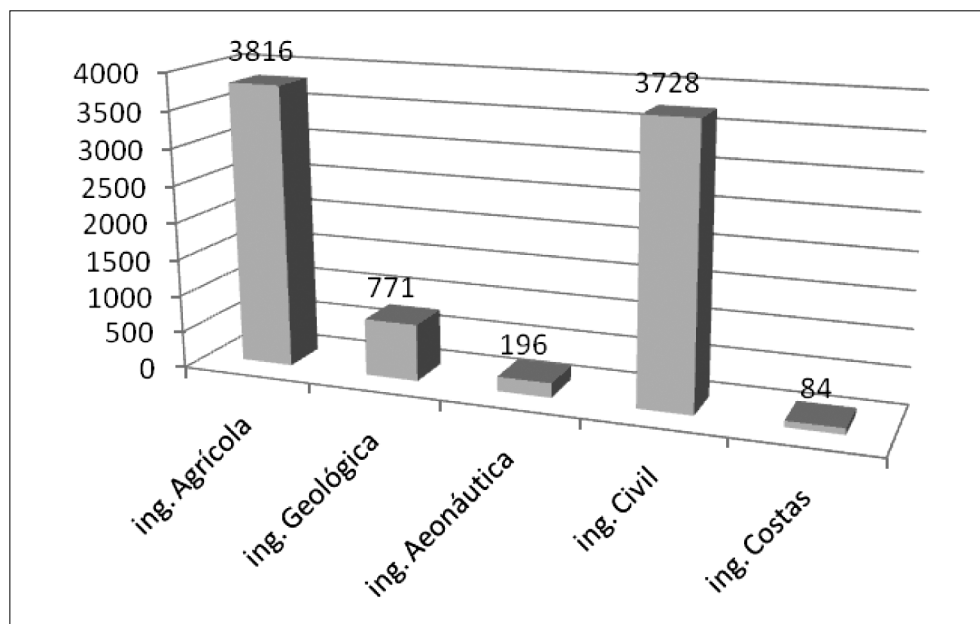
- Buscadores de ficheros OWL y, en general, cualquier tipo de lenguaje para ontologías:
 - Swoogle (<http://swoogle.umbc.edu/>).
 - Falcons (<http://iws.seu.edu.cn/services/falcons/conceptsearch/index.jsp>).

- Intellidimension (<http://www.semanticwebsearch.com/query/?st=yss=architecture&go=search>).
- Dr Watson (http://watson.kmi.open.ac.uk/editor_plugins.html).
- Repositorios de ontologías:
 - OwlSight Ontology Repository (<http://pellet.owldl.com/ontology-browse>).
 - OpenOntology Repository (<http://ontolog.cim3.net/cgi-bin/wiki.pl?OpenOntologyRepository>).
 - Tones (<http://owl.cs.manchester.ac.uk/repository/>).

Los resultados, plasmados en la figura 1, demuestran la poca formalización que tiene la disciplina de Ingeniería de Puertos y Costas, al menos en «formato ontología», en relación con el resto dentro de su misma área de actuación. Como es bien sabido, no existen herramientas similares que permitan realizar este tipo de búsquedas sobre tesauros.

FIGURA 1

Número de clases y subclases en ontologías sobre ingeniería



Esta búsqueda demuestra que nuestra ontología es única en su tipo, al menos para los buscadores especializados. Con el fin de evaluar Onto-Sactol se utilizaron diversas métricas, sobre todo aquellas que facilitarían la evaluación de su diseño y aplicación en los procesos extractivos y semánticos.

Se trata de una ontología de elevado nivel de complejidad, tanto por la temática como por su desarrollo. Posee 6.284 conceptos en español y 23.546 instancias, formalizadas en OWL, al igual que los 89 axiomas que facilitan la estructuración del conocimiento. Esta ontología posee, además, imágenes que están asociadas a los conceptos, así como una estructura conceptual que deriva relaciones paradigmáticas con un elevado nivel de complejidad, parecido a Wordnet y EuroWordnet.

Los términos de la ontología están asociados a subclases que enuncian su estructura semántica (término, sustantivo, adjetivo, homónimo, merónimo, hipónimo, hiperónimo, etc.). Cada término definido como clase principal tiene asociada una propiedad que se denomina *translate*, es decir, un conjunto de equivalencias para ese término en inglés y alemán. Se prefirió este camino al uso de las anotaciones con el fin de enriquecer las futuras búsquedas en otros idiomas. A la vez, cada término tiene asociada una imagen, almacenada en el sistema junto a su anotación, definida por las siguientes etiquetas:

- **Connotativo:** información sobre lo que connota el objeto.
- **Denotativo:** elementos que facilitan declarar aquellas características que se declaran en la imagen: dimensiones, estado físico, técnica de reproducción y obtención, etc.
- **Palabras clave.**
- **URL:** ubicación donde se encuentra ese archivo.
- **Autor:** responsable del documento y de su información.
- **Datos onomásticos:** personalidades, etc., incluidas en la imagen.
- **Datos topográficos:** lugar físico de donde se ha tomado la imagen.

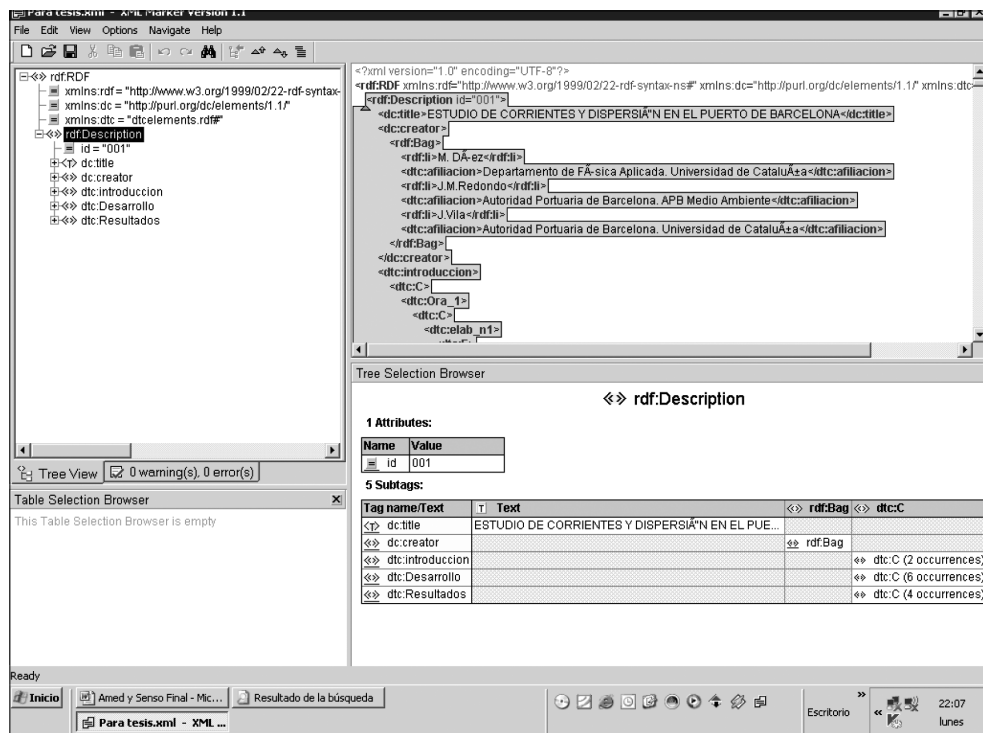
El sistema ontológico trabaja con textos estructurados en RDF Schema, y con etiquetas que están en consonancia con la ontología, con lo que se facilita que un sistema automático (generalmente llevado a cabo por autómatas) realice acciones concretas en el documento (localizar nuevos documentos, conectar enlaces a clases, etc.) (véase figura 2).

La ontología cuenta también con un modelo de desambiguación de textos asociado a partir de las palabras que se marcan en el texto. Dicho modelo, empleado también para la recuperación de la información, se implementó con un sistema de anotación constituido por el conjunto de metadatos Dublin Core junto a otro grupo de etiquetas creadas ad-hoc que tienen como finalidad el facilitar los procesos de recuperación, extracción, desambiguación y equivalencia semántica de los textos.

Este elevado nivel de complejidad obliga a que los procesos de evaluación de la aplicación sean desarrollados de forma diferente a los propuestos hasta la fecha, y ya comentados en el apartado 2 de este artículo. Se ha desarrollado un conjunto de medidas o métricas construidas a partir de las necesidades de los disímiles procesos que asume esta gran ontología. Los indicadores de análisis de la ontología se han desarrollado a partir de las mismas métricas que se conocen, pero se les ha dado una connotación especial para adaptarlas a la aplicación que se desarrolla.

FIGURA 2

Estructura sintáctica de los documentos de la ontología en RDF visualizada a través del editor XMLSpy



5. Indicadores para la evaluación: análisis

Los indicadores para la evaluación de esta ontología están pensados para evaluar los objetivos de la misma dentro de un sistema de extracción y desambiguación de textos, con el fin de introducir mejoras para futuras aplicaciones y sistemas. Para evaluar la ontología se han declarado indicadores léxicos, de recuperación de la información, de la estructura sintáctica y una fase de constatación.

5.1. Indicadores léxicos

Miden la calidad del léxico en lo referente a la cobertura temática, exhaustividad de la anotación, redundancias semánticas, taxonomías ambiguas, capacidad de desambiguación y capacidad de traducción. A continuación se declaran las ecuaciones y la explicación de cada indicador:

Cobertura temática: una ontología tiene cobertura temática suficiente si formaliza la mayoría de los conceptos de la esfera del conocimiento que representa. En el caso de los sistemas que simbolizan una materia o conocimiento nuevo habría que verificar cuál es el nivel de alcance de la ontología en cuanto a conceptos dentro de esta ciencia nueva. Se desarrolla con la fórmula $Ct = CCO/CCM$, donde CCM representa las temáticas y CCO es la cantidad de conceptos que posee la ontología.

Exhaustividad de la anotación: valora el nivel de las anotaciones y su valor semántico en el momento de representar la información. La fórmula empleada es: $EA = CNC/CN$, donde EA es la exhaustividad de la anotación, CNC es la cantidad de notas correctas y CN es la cantidad de notas de la ontología. Se entiende que una anotación es correcta cuando describe correctamente el objeto al que está asociado o cuando cumple de manera eficaz la función para la que se creó, ya que hay anotaciones que no necesariamente deben servir para definir conceptos. Dado que esta comprobación es un trabajo arduo y de difícil automatización, se puede emplear una muestra representativa en lugar del análisis de la ontología completa. Cuanto más se aproxime el resultado a 1 mejor descrita estará la ontología.

Redundancias semánticas: la fórmula usada es: $RSE = CCA/CCO$, donde RSE son las redundancia semánticas, CCA la cantidad de conceptos ambiguos y CCO representa a la cantidad de conceptos de la ontología. Se entiende que un objeto es ambiguo cuando no está claramente definido su uso, no está bien ubicado en la categoría que le corresponde, o cuenta con una descripción vaga. El objetivo es que haya cero redundancias semánticas.

Capacidad de desambiguación: se calcula con la fórmula $CDA = NTDA/NTD$, donde CA es capacidad de desambiguación, $NTDA$ es el número de etiquetas desambiguadas en un documento y NTD responde al número de etiquetas marcadas en el documento. Para su aplicación se calcula la sumatoria de cada resultado de cada etiqueta desambiguada, entre las etiquetas de cada uno de los documentos de la colección de estudio declarado.

Capacidad de traducción: es la capacidad que tiene el sistema de realizar equivalencia en el texto con aquellos términos que poseen una etiqueta como marca (en nuestro caso, esa propiedad es «translate»). Se denota aquí como $CT' = CTT/CTE$, donde CTT es la cantidad de términos traducidos dividido entre la cantidad de términos equivalentes (CTE). No se trata de una capacidad excesivamente desarrollada en las ontologías, pero que se consideró importante en función de los objetivos del proyecto dentro del cual se enmarca Onto-Satcol.

5.2. Indicadores de recuperación de información

Estos indicadores miden la efectividad de la ontología para recuperar información relevante para el usuario, a partir de las medidas clásicas de recuperación de la información: precisión y exhaustividad. Dado que estas medidas están pen-

sadas para evaluar la recuperación de información, hemos desarrollado una variación con el fin de que se pueda adaptar específicamente a la ontología que estamos analizando.

Precisión en la recuperación de documentos: esta medida estudia la capacidad del sistema para recuperar documentos relevantes a una petición de los usuarios. La fórmula es $PrecD = CDR/CDO$, donde CDR es la cantidad de documentos recuperados y CDO la cantidad de documentos en la ontología.

Precisión en la recuperación de documentos de tipo 2: es la capacidad de la ontología de devolver imágenes relativas a algún elemento lingüístico, y se calcula con $PrecD2 = CIR/CIO$, donde CIR es la cantidad de imágenes recuperadas y CIO es la cantidad de imágenes que han sido descritas en la ontología.

Exhaustividad: es la proporción de material relevante recuperado, del total de objetos relevantes de la ontología, y se calcula con la fórmula $ExausD = DR/CDOst$, donde DR es la cantidad de objetos relevantes a la solicitud y $CDOst$, es la cantidad de objetos que existen sobre ese tema en la ontología.

5.3. Indicadores para la evaluación de la estructura sintáctica

Para evaluar la ontología a nivel sintáctico utilizamos Protex, una herramienta desarrollada específicamente para este sistema de información. Protex (Domínguez, 2010) está desarrollado a partir de un algoritmo que genera un parser y un analizador sintáctico escrito en Python, permitiendo la visualización de los elementos de una taxonomía al igual que las herramientas Case (Computer Aided Software Engineering) desarrolladas a tales efectos. El objetivo de esta fase del análisis es la localización de incumplimientos en los estándares ontológicos como OWL (Ontology Web Language), RDF (Resource Description Framework) y DAML (Darpa Agent Markup Lenguaje). Protex fue creado con la finalidad de permitir localizar los siguientes errores:

- Identificación de solapamiento en el desarrollo de clases
- Localización de conceptos vacíos
- Omisión de conocimiento disjunto
- Omisión de conocimiento por falta de exhaustividad
- Conceptos mal formulados

5.4. Adecuación de los requerimientos

Como en todo sistema ontológico que se evalúa, siempre se hace imprescindible valorar los resultados del proceso. Para ello es obligatorio decidir si la ontología cumple claramente con los estándares internacionales y si los resultados que arroja la investigación empírica tiene la validez necesaria para ser tomados

en cuenta. De lo contrario habrá que comenzar desde el inicio para valorar las deficiencias.

6. Experimento

Para llevar a cabo la evaluación se han escogido a 6 expertos en el tema de la Ingeniería de Puertos y Costas. En otros métodos, como los comentados anteriormente en el apartado 2, se suele elegir principalmente a expertos en ontologías antes que a expertos del área de conocimiento. Entendemos que esa opción aporta una visión excesivamente sesgada de la evaluación, muy centrada en la ontología y poco en el grado de concreción de la organización del conocimiento plasmada, y por eso hemos optado por esta otra. Ellos serán los que analizarán los documentos y la estructura semántica de la ontología, y verificarán la precisión y la exhaustividad de cada una de las preguntas empleando los indicadores anteriormente descritos. Para llevar a cabo su trabajo emplearon Protex ya que, por un lado permite la navegación dentro de la ontología, y por otro tiene implementados gran parte de los indicadores propuestos en este modelo, por lo que puede generar los informes necesarios para proceder a la evaluación.

7. Resultados

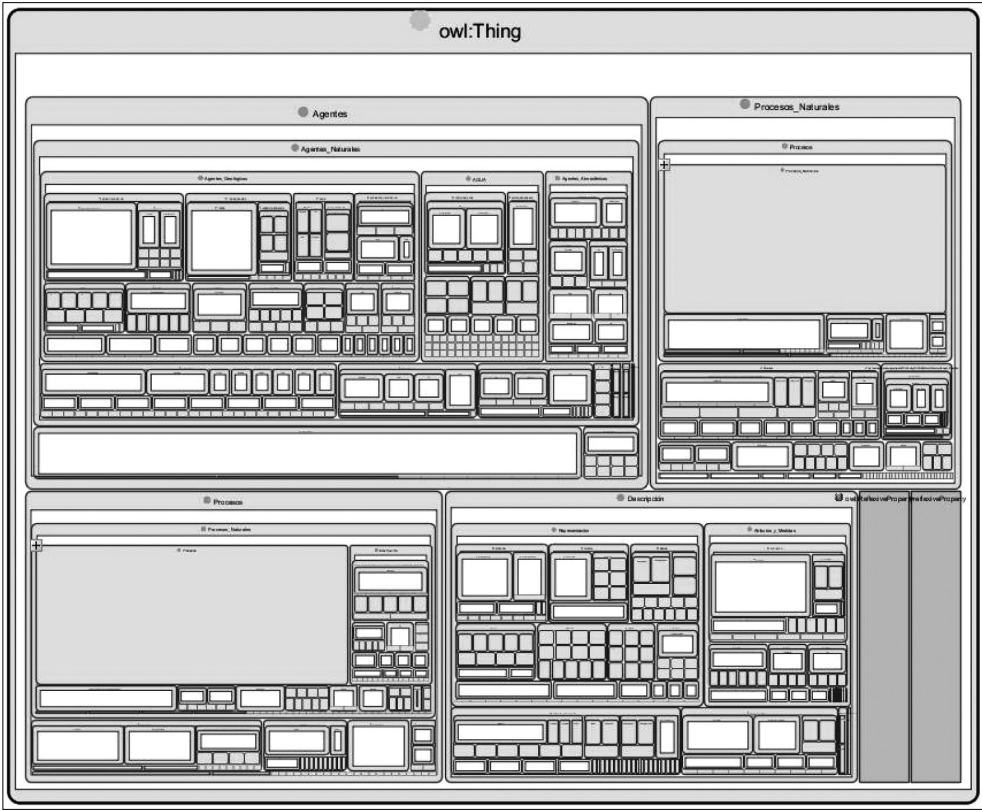
7.1. Indicadores léxicos

Se valoran a partir de los datos aportados por los observadores, analizando el nivel de cobertura temática alcanzado en los nodos principales de la ontología. Cada experto anota el número de nodos temáticos que cree se corresponden con el dominio que representa. Los resultados han sido calculados de forma individual para cada uno de los procesos, añadiendo una variación en la fórmula básica para el cálculo final. El cálculo de la cobertura temática ($Ct = CCO/CCM$) será la media de resultados aportados por cada observador ($CtI = CCO_{i1}/CCM$). En nuestro caso, el resultado final ha sido 0,7 lo que indica que sólo un 70% de los conceptos relativos a ese dominio están representados conceptualmente en esa ontología (véase tabla I). En parte es lógico este resultado, ya que la ontología está todavía en proceso de creación, y aún no se ha terminado de vaciar una parte del corpus. La inclusión de los conceptos que faltan implica, además, el establecimiento de relaciones con los que ya existen, la creación de axiomas para facilitar su comportamiento y añadir la correspondiente anotación para la descripción y posterior recuperación. Por ejemplo, se constató con este indicador que una de las clases menos desarrolladas era «Procesos naturales» (figura 3), que aglutina todos aquellos procesos que lleva a cabo la propia naturaleza, y que no ha sido estudiada del todo en esta fase de creación de la ontología. Esto sucede especialmente con los «sucesos bióticos», la «abrasión», los «procesos hidrodinámicos» y los «eventos hidrológicos».

TABLA I
Indicadores de Evaluación Léxica

Indicadores	Valor
Cobertura temática	0,7
Exhaustividad de la anotación	0,16
Redundancias semánticas	0,15
Capacidad de desambiguación	0,88
Capacidad de traducción	0,78

FIGURA 3
Representación gráfica de los nodos principales de la ontología por medio de la herramienta Jambalaya



Aquellas clases que cuentan con muchas subclases ofrecen una representación formada por muchos cuadros pequeños. Por el contrario, los cuadros grandes —caso de «procesos naturales», que aparece abajo a la derecha y ampliado abajo a la izquierda— se muestran como un gran cuadrado vacío, lo que implica una mala representación de la clase.

Para el cálculo de la exhaustividad en la anotación los expertos dividieron la ontología en 6 secciones, y tenían como objetivo valorar todas las notas que no estaban correctas. De las 23.546 instancias sólo 376 estaban correctas, lo que otorga un valor de 0,16. Esto indica que las anotaciones no son eficientes, lo que puede impedir una buena recuperación de la información y provocar una baja relevancia. Una parte de esas instancias son ejemplos de clases (por ejemplo, los diferentes grados con los que se mide un terremoto en la escala Richter, o las denominaciones para el oleaje —marejadilla, marejada, gruesa, arbolada, montañosa— en la escala Douglas) y otra parte pueden ser clases que, con el fin de agilizar las consultas, en ocasiones se convierten en instancias, aunque su naturaleza siga siendo la de clase. Esta técnica se emplea sólo en aquellas consultas en las que no sea necesario inferir conocimiento, puesto que cuando una clase se convierte en instancia pierde la mayoría de las relaciones que la conectan con el resto de clases.

Para analizar las redundancias de significado se utilizó la estructuración Sense de la herramienta PROTEX, la cual brindó a los expertos la posibilidad de visualizar los nodos para determinar, mediante etiquetado, aquellas redundancias semánticas. El índice resultante fue de 0,15, lo que indica que el nivel de representación en una buena cantidad de términos es relativamente bajo. Ese nivel de ambigüedad conceptual es muy importante, y posiblemente esté debido a la extensión de la ontología, lo que nos indica que es necesario hacer cambios en la estructura de la misma con el fin de tener mejores resultados. No obstante, dependiendo del nivel de profundidad de los nodos se observaba una varianza evidente en cuando a la ambigüedad se refiere. Esto podía ser debido a que algunas subáreas de conocimiento estaban mal representadas, pero no todas (figura 3). Por ese motivo realizamos un test de desambiguación.

Con el fin de conocer si la ontología se podría emplear como herramienta para desambiguar textos especializados se han seleccionado 50 documentos de ese tipo debidamente etiquetados con RDF. En cada uno de ellos se han marcado palabras clave, así como verbos, sustantivos, adjetivos y frases semánticas. Los textos fueron sometidos a desambiguación, por medio de la opción correspondiente de la herramienta Protex. De las 890 palabras marcadas en cada texto se desambiguaron 789, con un nivel de desambiguación de 0,88, lo que indica que el sistema es bastante fiable en cuanto a desambiguación y transformación del corpus.

Para el estudio de la capacidad de traducción se analizaron 50 textos, en los cuales aparecían 150 marcas de equivalencia, que permitieron la traducción de 118 términos, lo que evidencia que se consigue un índice elevado de traducción, 0,78, en función de la muestra.

7.2. Precisión en la recuperación de documentos

Para analizar este indicador se pidió a los expertos que realizaran 15 búsquedas en el sistema y que midieran la relevancia de los documentos recuperados. Con el fin de dificultar más el proceso, y aumentar el riesgo de ruido, las consultas estaban

formadas por un único término. El resultado fue que de los 2.890 documentos de media recuperados, 2.345 de media eran relevantes, evidenciando un elevado nivel de relevancia, un valor medio de 0.81. En cuanto a la recuperación de documentos de tipo dos, es decir, material no textual, se recuperaron todas las imágenes del sistema para las 15 peticiones, aportando un valor de 1. Esto indica que el sistema es confiable en lo que respecta a precisión cuando recupera material no textual.

7.3. Indicadores de evaluación de la estructura sintáctica

La herramienta empleada para estos cálculos, Protex, descubrió la existencia de 26 clases que poseen un nombre diferente, pero un contenido (expresado en la anotación) similar, por tanto, se solapan. Se han visualizado 69 conceptos vacíos, es decir, con errores de significado, eso crea problemas semánticos y hace más compleja la taxonomía. En la ontología se han encontrado 89 errores al tratar conocimiento disjunto y 67 conceptos macro, que poseen igual número de sub-conceptos con igual estructura jerárquica. De igual forma aparecen 67 problemas por conceptos mal descritos por falta de calidad en la exhaustividad en la descripción de los documentos (tabla II). Para solucionar de manera automatizada estos problemas se está analizando la posibilidad de integrar en el sistema ORE-Tool (Ontology Repair and Enrichment), de la Universidad de Leipzig, que permite depurar inconsistencias y crear axiomas de manera automática.

TABLA II

Problemas de sintaxis y semántica. Se muestra el número de clases que han presentado problemas después del cálculo realizado con Protex y el porcentaje que supone esas clases erróneas con respecto al número total de clases de la ontología

Conceptos vacíos	Conocimiento disjunto	Conceptos macro	Descripciones erróneas	Clases solapadas
0,11	0,14	0,11	0,11	0,04
69	89	67	67	26

Este experimento no puede terminar hasta que no exista una verificación de la exactitud de los datos y la calidad de los mismos. Para ello se decidió formular algunas preguntas que facilitarán la detección de los conceptos, la traducción y la desambiguación. Estas preguntas fueron las siguientes:

1. ¿Cuáles son los efectos que produce la marea en la costa de Cantabria?
2. ¿Cuáles son las centrales azucareras más contaminantes de Cuba?
3. ¿Con qué sustancia química se hace la espectrografía de gases?

4. ¿Cuáles son las zonas geográficas de menor desarrollo portuario?
5. ¿Cuál es la traducción del término «rompiente de derrame» en alemán?
6. ¿Cuál es su contexto de aplicación?
7. ¿Cuántas imágenes de derrame de petróleo existen en la ontología?

Los resultados de estas preguntas demostraron valores muy similares a los observados durante los experimentos, con lo que queda demostrada la veracidad de la información proporcionada por el experimento. Las respuestas a estas interrogantes fueron contestadas por la ontología con los siguientes niveles (tabla III).

TABLA III
Nivel de respuesta del sistema en la fase de adecuación

	Exhaustividad	Precisión
Pregunta 1	0,5	0,832
Pregunta 2	0,3	0,901
Pregunta 3	0,5	0,700
Pregunta 4	0,7	0,892
Pregunta 5	0,1	0,823
Pregunta 6	0,1	0,992
Pregunta 7	0,2	0,657

8. Conclusiones

Estamos ante un modelo de evaluación de ontología que, si bien emplea conceptos clásicos en el desarrollo de las ontologías, también aporta nuevos parámetros para el quehacer de las evaluaciones de estos sistemas. Máxime si representan el conocimiento de dominios noveles, contruidos a partir de otras disciplinas.

Los dos principales problemas detectados en los sistemas de evaluación empleados en la actualidad se centran en un exceso de peso por parte de las aportaciones de los observadores, y en contar con una visión centrada únicamente en una o dos dimensiones del problema ontológico. El modelo propuesto aquí emplea diferentes criterios para evaluar parámetros relacionados con cuestiones léxicas (midiendo la cobertura temática, la anotación, las redundancias semánticas, la capacidad de desambiguación y la de traducción), de recuperación de información y de la estructura sintáctica. Es decir, que se aporta una visión que contempla las tres dimensiones de una ontología.

Este modelo de evaluación ha permitido desarrollar nuevas formas de reconocer en la ontología patrones de relevancia semántica y establecer equivalencias para la traducción y desambiguación.

Aunque el proceso de evaluación de la ontología no se suele contemplar en muchas de las metodologías para la construcción de ontologías, los resultados de la evaluación permiten conocer los errores del sistema antes de su puesta en marcha, por lo que entendemos que debería de formar parte de ésta, favoreciendo la construcción por medio de la retroalimentación.

Por otra parte, pensamos que queda patente la necesidad de contar con herramientas que faciliten el proceso de evaluación. Si bien es cierto que en la mayoría de métodos propuestos este apartado no se suele tocar demasiado, es impensable realizar el proceso de la evaluación de manera manual. En nuestro caso la aplicación Protex, creada específicamente para este proyecto, se ha perfilado como un instrumento fundamental en el marcado automático de errores en la ontología y la visualización de los nodos y los conceptos. Esto ha facilitado la eliminación de inconsistencias semánticas, estructurales y ambigüedades en los conceptos de la ontología.

Como trabajo futuro, y no necesariamente relacionado con el proceso de evaluación, será imprescindible seguir desarrollando la ontología Onto-Satcol, ya que ha quedado patente que cuenta con numerosas inconsistencias semánticas que ponen el peligro su puesta en marcha dentro de un sistema de mayor calado. Por otra parte, se hace necesario minimizar su complejidad, ya que puede ser causante de un hecho que no ha sido contemplado en esta evaluación, pero que se planteará en una revisión del modelo propuesto. Se trata del tiempo que el sistema tarda en dar respuesta a las preguntas de los usuarios.

Es evidente que existen y existirán muchas formas para evaluar las ontologías. Sin embargo será necesario que estos mecanismos de evaluación vayan en busca de análisis integrales de los sistemas ontológicos, de manera que los mismos puedan ser valorados en su totalidad. Las ontologías de dominios de conocimiento nuevo tienen que buscar mejores formas de presentación para que su representación cognitiva sea cada vez más apegada a las necesidades de los futuros usuarios. Es por ello que en los nuevos métodos de evaluación el usuario final debe tener una marcada interacción en los procesos evaluativos.

9. Agradecimientos

Este artículo forma parte del proyecto de investigación FFI2008-06080-C03-01, subvencionado por el Ministerio de Ciencia e Innovación, España.

10. Bibliografía

- Boehm, B. W. (1981): *Software Engineering Economics*. New York; Prentice-Hall.
- Brank, J.; Grobelnik, M., y Mladenic, D. (2005): A survey of ontology evaluation techniques. Paper presented at the *SIKDD 2005 at multiconference IS 2005*, Ljubljana, Slovenia. <http://eprints.pascal-network.org/archive/00001198/01/BrankEvaluationSiKDD2005.pdf>.

- Brewster, C.; Alani, H.; Dasmahapatra, S., y Wilks, Y. (2004): Data Driven Ontology Evaluation. Paper presented at the Proceedings of the *Language Resources and Evaluation Conference (LREC 2004)*, Lisboa, Portugal.
- Domínguez, S. (2010): PROTEX (Version beta) [Lematizador]. Santa Clara, Cuba: Universidad Central de las Villas, Departamento de automática.
- Evermann, J. E., y Fang, J. (2010): Evaluating ontologies: Towards a cognitive measure of quality. *Information Systems*, vol. 35, 391-403.
- Fahad, M., y Abdul Qadir, M. (2008): A Framework for Ontology Evaluation. Paper presented at the Supplementary Proceedings of the *16th International Conference on Conceptual Structures (ICCS'08)*, Toulouse, France. <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-354/p59r.pdf>.
- Fellbaum, C. (1998): WordNet – An Electronic Lexical Database. Massachusetts: MIT-Press.
- Fernández-Breis, J. T.; Egaña Aranguren, M., y Stevens, R. (2009): *Quality evaluation framework for bio-ontologies*. Paper presented at the *ICBO: International Conference on Biomedical Ontology*, Buffalo: New York.
- Fernández-López, M.; Gómez-Pérez, A., y Juristo, N. (1997): Methontology: from ontological art towards ontological engineering. Paper presented at the *Spring Symposium on Ontological. Engineering of AAAI*, California.
- Gangemi, A.; Catenacci, C.; Ciarmita, M., y Lehmann, J. (2005): *A Theoretical Framework for Ontology Evaluation and Validation*.
- Gómez Pérez, A. (1994): Some Ideas and Examples to Evaluate Ontologies. Conference paper at an *IEEE Event*.
- Gruninger, M., y Fox, M. S. (1995): Methodology for the Design and Evaluation of Ontologies. Paper presented at the *Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95*, Montreal.
- Guarino, N., y Welty, C. (2002): Evaluating Ontological Decisions with OntoClean. *Communications of the ACM*, vol. 45 (2), 61-65.
- Hartmann, J.; Spyns, P.; Giboin, A.; Maynard, D.; Cuel, R.; Suárez-Figueroa, M. C., y otros. (2005): *D1.2.3 Methods for ontology evaluation*, Knowledgeweb.
- Kent, A. (1955): Machine literature searching. VIII. Operational Criteria for Designing Information Retrieval Systems. *American Documentation*, vol. 6 (2), 93-101.
- Kuhn, T. (1979): *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Leiva-Mederos, A.; Senso, J. A.; Domínguez-Velasco, S., e Hípola, P. (2009): An Automat for the Semantic Processing of Structured Information. Paper presented at the *International Conference on Intelligent Systems Design and Applications*.
- Lozano-Tello, A., y Gómez-Pérez, A. (2004): ONTOMETRIC: A method to choose the appropriate ontology. *Journal of Database Management*, vol. 15 (2).
- Maedche, A.; Motik, B.; Stojanovic, L.; Studer, R., y Volz, R. (2003): Ontologies for Enterprise Knowledge Management. *EEE Intelligent System*, 26-34.
- Maedche, A., y Staab, S. (2002): Measuring Similarity between Ontologies. Paper presented at the In: Proc. of the *European Conference on Knowledge Acquisition and Management – EKAW-2002*, Madrid: Spain.
- Maiga, G., y Ddembe, W. (2008): A Systems Approach to User Evaluation of Biomedical Ontologies Paper presented at the *First International Workshop on Ontologies in Interactive Systems*.

- Ning, H., y Shihan, D. (2006): Structure-Based Ontology Evaluation. Paper presented at the *International Conference on e-business on Engineering*.
- Obrst, L.; Werner, C.; Inderjeet, M.; Ray, S., y Smith, B. (2007): *The Evaluation of Ontologies: Toward Improved Semantic Interoperability*. En C. J. O. Baker & K.-H. Cheung (Eds.), *Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences*: Springer.
- Paslaru Bontas, E., y Mochol, M. (2005): A Cost Model for Ontology Engineering. En F. U. Berlin (Ed.).
- Qadir, M., y Noshairwan, W. (2007): Warnings for Disjoint Knowledge Omission in Ontologies. Paper presented at the *Second International Conference on internet and Web Applications and Services (ICIW07)*.
- Ramos, E.; Núñez, H., y Casañas, R. (2009): Esquemas para evaluar ontologías únicas para un dominio de conocimiento. *Revista Venezolana de Información, Tecnología y Cocimiento*, vol. 6 (1), 57-71.
- Sabou, M.; López, V.; Motta, E., y Uren, V. (2006): Ontology Selection: Ontology Evaluation on the Real Semantic Web. Paper presented at the *WWW2006*, Edimburgo.
- Salton, G.; Wong, A., y Yang, C. S. (1975): A Vector Space Model for Automatic Indexing. *Communications of the ACM*, vol. 18 (11), 613-620.
- Spyns, P. (2005): *EvaLexon: Assessing triples mined from texts*. Bruselas.
- Staab, S. (2004): Why Evaluate Ontology Technologies? Because It Works! *IEEE Intelligent Systems*.
- Uschold, M., y King, M. (1995): Towards a Methodology for Building Ontologies. *Workshop on Basic Ontological Issues in Knowledge Sharing*.
- Yao, H.; Mark, A.; Orme, M., y Etzkorn, L. (2005): Cohesion Metrics for Ontology Design and Application. *Journal of Computer Science*, vol. 1 (1), 107-113.
- Yu, J.; James, A. T., y Tam, A. (2009): Requirements-oriented methodology for evaluating ontologies. *Information Systems*, vol. 34, 766-791.
- Zhang, T.; Xu, D., y Chen, J. (2008): Application-oriented purely semantic precision and recall for ontology mapping evaluation. *Knowledge-Based Systems*, vol. 21, 794-799.