

Ontologías y organización del conocimiento: retos y oportunidades para el profesional de la información

Por Francisco-Javier García-Marco

Resumen: Se analiza la convergencia que se está produciendo en el campo de las ontologías entre ingeniería del conocimiento y organización del conocimiento en el marco del proyecto de la web semántica. Se estudia el desarrollo de la investigación sobre ontologías en las ciencias de la documentación y en el conjunto de las disciplinas que se interesan por los problemas ontológicos. Se contextualiza el actual frente de investigación en el campo de las ontologías en el marco del desarrollo de Internet y especialmente de la web semántica. Finalmente, se analizan las implicaciones de futuro para el profesional de la información: integración en un campo transdisciplinar más amplio y con un gran porvenir; clarificar su posición en él, y asegurar una formación adecuada en los nuevos estándares y tecnologías.

Palabras clave: Ontologías, Organización del conocimiento, Ingeniería del conocimiento, Internet, Web semántica.

Title: Ontologies and knowledge organization: challenges and opportunities for information professionals

Abstract: The emerging convergence in the field of ontologies between knowledge organization and knowledge engineering is examined in the context of the semantic web project. We describe the development of research on ontologies in library and information sciences and in other disciplines interested in ontological problems. The emergence of a research agenda on ontologies is discussed in the context of the development of the internet and, specifically, in relation to the semantic web project. Finally, some implications of this convergence for the information professional are discussed: the integration of knowledge organization into a broader transdisciplinary research arena with a very promising future, the need to clarify the role and potential contributions of the information professional, and the urgency of adequate education and training in the new standards and technologies.

Keywords: Ontologies, Knowledge organization, Knowledge engineering, Internet, Semantic web.

NOTA: Este artículo ha sido evaluado y aprobado por pares.

García-Marco, Francisco-Javier. "Ontologías y organización del conocimiento: retos y oportunidades para el profesional de la información". En: *El profesional de la información*, 2007, noviembre-diciembre, v. 16, n. 6, pp. 541-550.

DOI: 10.3145/epi.2007.nov.01



Francisco Javier García Marco es doctor en filosofía y letras desde 1994 y profesor titular de universidad del área de biblioteconomía y documentación de la Universidad de Zaragoza desde 1996. Ha participado en la organización de numerosos eventos científicos, entre los que destacan los Encuentros Internacionales sobre Sistemas de Información y Documentación (Ibersid), de los que es director. Dirige las revistas *Scire: representación y organización del conocimiento* e *Ibersid*.

1. Las ontologías: un campo de convergencia transdisciplinar

EN LOS ÚLTIMOS AÑOS EXISTE UNA FUERTE tendencia en el campo de las ciencias de la documentación por asimilar los sistemas de organización del conocimiento tradicionales –clasificaciones bibliográficas y archivísticas, tesauros de distintos tipos, lenguajes de encabezamientos de materias y de autoridades, principalmente– y los que han sur-

gido en el nuevo entorno de internet –taxonomías corporativas y de usuario, términos libres asignados (*tags*), etc.– a un concepto nuevo procedente de la inteligencia artificial: las ontologías (v. g. Vickery, 1997; Baeza; Ribeiro, 1999; Currás, 2005; Hersh, 2005).

Se trata de una situación semejante a lo que ocurre en el campo de la descripción bibliográfica y documental en general, que cada vez se percibe más claramente

integrado con el movimiento general de descripción de objetos informáticos mediante metadatos. Por otra parte existe la tendencia paralela de los informáticos que trabajan en aplicaciones de lenguajes documentales en internet así como de los gestores que llevan y promueven los proyectos a denominarlos también ontologías. Como resultado se está produciendo una rápida convergencia de ambos lenguajes.

¿Por qué sucede esto? Sin duda existen varias razones. Una de ellas es el hecho obvio de que una de las aplicaciones más poderosas de las ontologías –inicialmente desarrolladas para soportar la inferencia¹ lógica en sistemas expertos y de adquisición de conocimiento– es la recuperación de información. Otra condición importante es que los lenguajes de descripción de ontologías formales implementados sobre software –por ejemplo, *Protégé*– permiten fácilmente representar los objetos, las relaciones y restricciones existentes en los lenguajes documentales, y facilitar su interoperabilidad en el marco de los sistemas de información interconectados.

La familia de términos “ontolog-“ no es ajena a la tradición de los lenguajes documentales; antes al contrario, ha estado muy presente a través del problema de las categorías universales de la clasificación; hasta el punto en que autores como **Dahlberg** (1978) han hablado tempranamente de ellas como de “estructuras ónticas”. Además, la indudable modernidad y prestigio que rodea a los términos procedentes de la informática avanzada –y en general de los campos de investigación punteros– favorece su adopción. Por último, el acceso a la información en la Web ha topado finalmente con el problema conceptual, y existen cada vez más voces –entre ellas la de su fundador **Tim Berners-Lee** (1998, 2001)– que reclaman el abordaje del problema del significado –el problema *semántico*– para que la información de la Red se pueda relacionar y utilizar de forma más precisa y potente. Siendo como es internet la tecnología de nuestros tiempos, la asunción del problema semántico dentro de ella sitúa la organización del conocimiento en un punto privilegiado, aunque subsidiario. Todo ello ha favorecido un proceso de rápida asimilación conceptual y terminológica, y el término ontologías se utiliza cada vez con más profusión.

Sin embargo, se olvida frecuentemente que dichas ontologías “informáticas” no son lenguajes documentales al modo tradicional, sino implementaciones conjuntas de sistemas de términos sobre diversas lógicas –generalmente, la de predicados de primer orden–, siendo lo auténticamente distintivo lo segundo. Esas herramientas sirven para la representación del conocimiento de cara a soportar las inferencias de sistemas expertos. Sus aplicaciones trascienden la recuperación de información, que constituirían dentro de ellas

“El problema semántico en internet sitúa la organización del conocimiento en algo privilegiado”

–nada más, pero nada menos– que una subdisciplina específica.

Por otra parte, y aunque se interesen en ellos, la especialidad de las ontologías formales “informáticas” no aborda sistemáticamente ni agota la totalidad de los problemas ontológicos ni en documentación ni en otras disciplinas adyacentes, especialmente el de las categorías universales que son del interés de la ontología filosófica, la semántica y la organización del conocimiento; o las leyes de la estructuración del significado que estudia la semántica. Las ontologías aportan una formalización de las meta-relaciones ontológicas que permite las operaciones lógicas entre ellas, pero la especificación de los entes, las restricciones y las relaciones específicas queda como un problema dentro de cada dominio y aun a nivel universal en buena parte. Esto es, las ontologías “informáticas” han implementado los aspectos ontológicos que contiene la lógica de primer orden, pero los aspectos categoriales quedan en gran parte fuera de la formalización (**Poli**, 2002).

Se vislumbra en ocasiones el peligro de confundir los diferentes instrumentos de organización conceptual y terminológica, haciendo sinónimos conceptos próximos pero que son claramente diferentes –sistemas de organización del conocimiento, tesauros, taxonomías y ontologías-. Esta posición es comprensible desde la perspectiva gerencial, general y lejana de los responsables máximos de las servicios de información, y de las personas interesadas en el tema. Sin embargo es inaceptable entre los expertos y, especialmente, dentro del campo de investigación, donde la asimilación conceptual y terminológica sólo contribuye a crear confusión. Este problema afortunadamente es uno de los que queda meridiamente desentrañado en la nueva norma británica *BS 8723 (British Standards Institute, 2006-)*. En este aspecto cabe ser optimistas: está claro que, superada la confusión inicial, los diferentes conceptos ocuparán el lugar que les corresponde.

Por tanto, parece que estamos ante la aparición de uno de esos nuevos campos de investigación eminentemente transdisciplinares como las ciencias cognitivas, la neurociencia o las ciencias ambientales en los que ninguna disciplina termina por organizar a las demás, ni tampoco es posible obviar los avances que se producen en los diferentes campos. Ciencias ontológicas es, posiblemente, un término presuntuoso, porque, aunque variado e interdisciplinar, el número de investigadores

activos y trabajos publicados en el área, como se verá seguidamente, es muy reducido frente a las grandes macrodisciplinas que se acaban de mencionar. En cualquier caso, la expansión de los términos metadatos y ontologías en las ciencias de la documentación constituye la punta del iceberg de un proceso de reconfiguración disciplinar, como resultado de la integración de diversas ciencias en el marco de desarrollo tecnológico info-comunicacional (figura 2).

2. Advenimiento de la tecnología ontológica y el desarrollo de internet

2.1. Caracterizando el campo de las ontologías técnicas

Ahora bien, ¿qué son esas ontologías que han irrumpido en la investigación moderna en las ciencias de la computación?

Uno de los más famosos investigadores y pionero en el tema, **Tom Gruber** (1993) las define simplemente como una “especificación de una conceptualización”; mientras que, más recientemente, la *International DOI Foundation* (2005) define una ontología estructurada como “an explicit formal specification of how to represent the entities that are assumed to exist in some area of interest and the relationships that hold among them”, es decir, como una especificación formal explícita –una declaración– sobre la manera de representar las entidades existentes en un área de interés y las relaciones que mantienen entre ellas. Ésta y otras definiciones, sin embargo, son muy amplias para el no experto y pueden servir para describir realidades de un nivel de complejidad muy diferente: listas de palabras (vocabularios simples), jerarquías de términos (taxonomías) o complejas formalizaciones léxicas como *WordNet 2.0*, que es capaz de representar relaciones de hiperonimia², hiponimia³, instanciación⁴, así como otras relaciones léxicas, incluyendo la representación morfológica. Por otra parte, no diferencian adecuadamente lo que aportan de nuevo las ontologías frente a las clasificaciones, los tesauros y otros lenguajes documentales.

Lo primero que conviene establecer es que las ontologías son un campo de investigación de la inteligencia artificial y más específicamente de la rama relacionada con la representación del conocimiento, la ingeniería del conocimiento, que se ocupa de la construcción de sistemas expertos. Se trata de un área de investigación que en su época emergente –y todavía hoy– ha tenido un carácter marcadamente interdisciplinar con aportaciones de la filosofía, la lingüística y las ciencias cognitivas en general (**Nicles; Pease; Schalley; Zaefferer**, 2007).

El objetivo de la ingeniería del conocimiento es constituir grandes bases de conocimientos sobre un

“Una ontología es un sistema de términos que sirve para describir y representar un área de conocimiento y que expresa las relaciones entre ellos por medio de un lenguaje formal que puede ser entendido por un ordenador”

tema en forma de declaraciones, reglas de inferencia y mecanismos de razonamiento (motor de inferencia) para resolver automáticamente problemas del dominio en cuestión. Si se quiere, y dicho más sencillo, para responder automáticamente preguntas sobre el dominio de representación, ya sea a un agente humano, a otro automático, con el fin de ayudarle en su proceso de toma de decisiones y, eventualmente, de ejecución de una tarea. Las ontologías son un procedimiento basado en la lógica de primer orden desarrollado para codificar adecuadamente el sistema de términos utilizados en dichas declaraciones, de forma que se exprese adecuadamente la relación entre los términos; esto es, son herramientas para construir sistemas conceptuales o, por utilizar una terminología común, vocabularios estructurados, en los que se explicitan todas las relaciones entre los términos que se utilizan y otras restricciones de significado.

McGuinness (2002) realiza una descripción de sus características: deben poseer un vocabulario controlado limitado –aunque extensible–, con una interpretación estricta de sus clases y relaciones entre términos –sin ambigüedades– y una relación jerárquica estricta entre sus clases; como propiedades típicas aunque no obligatorias, deben permitir la especificación de propiedades para cada clase, la inclusión de individuos –ejemplares– en la ontología y la especificación de restricciones de valor a nivel de cada clase; y, como propiedades deseables –ni obligatorias ni típicas–, se recomienda que permitan la especificación de clases disjuntas, la especificación de relaciones arbitrarias lógicas entre los términos y la distinción de relaciones, como inversas y parte-todo.

El objetivo de las ontologías es constituir un almacén de información semántica donde sea posible consultar el significado de un término a través de los mecanismos proposicionales propuestos ya por Aristóteles en su teoría de la definición (1965, ed.). Los significados de los términos son resueltos por los expertos en ontologías a partir de las relaciones entre los mismos: “an ontology is a document or file that formally defines the relations among terms” (**Berners-Lee; Hendler; Lassila**, 2001).

¿Por qué hacerlo? Ciertamente, dichos significados se pueden dar por obvios en un sistema experto cerra-

do; pero es absolutamente necesario declararlos explícitamente para intercambiar información entre diferentes sistemas expertos que se comunican entre sí. Como dice **Gruber** (1993), “lo importante es para qué sirve una ontología. Mis colegas y yo hemos estado diseñando ontologías con el objetivo de facilitar el intercambio y la reutilización de conocimientos”.

Imaginemos los sistemas de información de dos empresas de trabajo temporal de dos países distintos –por ejemplo, Reino Unido y España– intercambiando datos de ofertas y demandas de trabajo para equilibrar sus respectivos mercados. El sistema de una empresa inglesa puede solicitar a un servidor español empleados con un determinado perfil si existe un vocabulario que exprese la relación interlingüística entre los respectivos términos. Ese vocabulario puede residir en cualquiera de los dos sistemas o, aún mejor, ser independiente de los mismos, de manera que otros que requieran una información semejante puedan aprovechar esa información.

En definitiva, una ontología es un sistema de términos que sirve para describir y representar un área de conocimiento, y que expresa las relaciones entre ellos por medio de un lenguaje formal (lógico) que puede ser entendido por un ordenador. Comparte el vocabulario y su estructuración con los lenguajes de descriptores, los tesauros y las taxonomías que se utilizan para la organización de la información, que utilizan conjuntos de términos relacionados para describir un dominio o área de conocimiento. Sin embargo, implica una descripción formal exigente de esas relaciones que pueda ser interpretada por un ordenador. Por ello, gran parte del esfuerzo en el campo de las ontologías ha sido dedicado a elaborar lenguajes capaces de ello.

Existen varios lenguajes disponibles. Algunos han sido desarrollados en el ámbito de la inteligencia artificial como *CycL*, el lenguaje del proyecto *Cyc* (**Stephen; Lenat**, 2002), *Ontolingua* (2005) o el *Knowledge Interchange Format (KIF)* y su sucesor el *Simplified common logic*, o la *Knowledge representation system specification (KRSS)* (**Patel-Schneider; Swartout**), orientados a la representación de enunciados en la lógica de primer orden. Sin embargo, el más importante ahora es el *Web ontology language (OWL)* desarrollado sobre *rdf* y *rdfs*, y deudor a su vez de *OIL*, *DAML* y *DAML+OIL*.

2.2. Las ontologías en la www

Las ontologías han recibido un impulso decisivo de la estrategia de web semántica del *World Wide Web Consortium (W3C)* (1994-2004). Como es bien conocido, el objetivo es llevar la web a su máximo potencial, es decir, al máximo nivel de automatización en los procesos de transferencia de la información y el cono-

cimiento. Hasta ahora la *www* ha facilitado enormemente el proceso de compartir información entre personas gracias a sus eficaces estándares de comunicación –*http*– y de normalización de los documentos –*html*–. Sin embargo no permite la recuperación y procesamiento de la información a nivel de dato y combinaciones de ellos (información), sino tan sólo de documento, lo que es imprescindible para los procesos de automatización que soportan los diferentes tipos de lenguajes de programación. Los líderes del programa lo expresan así: “To date, the web has developed most rapidly as a medium of documents for people rather than for data and information that can be processed automatically. The semantic web aims to make up for this” (**Berners-Lee; Hendler; Lassila**, 2001).

Así por ejemplo, en la actualidad es posible consultar las páginas web con ofertas de viaje a un destino dado, pero es imposible saber sin leerlas cuáles son más convenientes. Actualmente esto sólo podría hacerse después de un cuidadoso análisis y de su introducción en una base de datos, pero eso requiere una operación centralizada o una cuidadosa concertación entre una red de bases de datos.

¿Cuál es la alternativa que ofrece la web semántica? Propone etiquetar esos documentos de ofertas de manera que ciertas informaciones como los lugares de salida, destino, medio, clase y precio pudieran ser fácilmente reconocibles, y que luego un programa pudiera recopilar esos datos, procesarlos y presentar resultados al usuario. Más aún, dichas etiquetas podrían incluso ser distintas siempre que en algún lugar de la red estuviera disponible una equivalencia de las mismas y de sus relaciones. Por fin, sus contenidos podrían ser procesados si se almacenasen en algún lugar común las relaciones entre ellos, que es de lo que se ocupan las ontologías (que también tienen un gran potencial para ocuparse del problema anterior). Así, por ejemplo, si se produce una oferta genérica de vuelos a Brasil con una tarifa común a todos los destinos internos llegando a Río de Janeiro, sería posible que el sistema valorara esa oferta automáticamente en el caso de que quisiéramos ir a Brasilia, si un fichero especificara que esta ciudad existe y es una subclase de Brasil.

En definitiva, como dice la web oficial (*World Wide Web Consortium*, 1994-2004): la web semántica trata de dos cosas. Se ocupa de que existan formatos comunes de intercambio de datos, mientras que en la web original sólo teníamos intercambio de documentos. Además tiene que ver con la elaboración de un lenguaje para codificar cómo los datos se relacionan con los objetos del mundo real. Eso permitiría a una persona o a una máquina comenzar a trabajar en una base de datos, y luego moverse a través de otras que no están conectadas por cables, sino por tratar del mismo asunto.

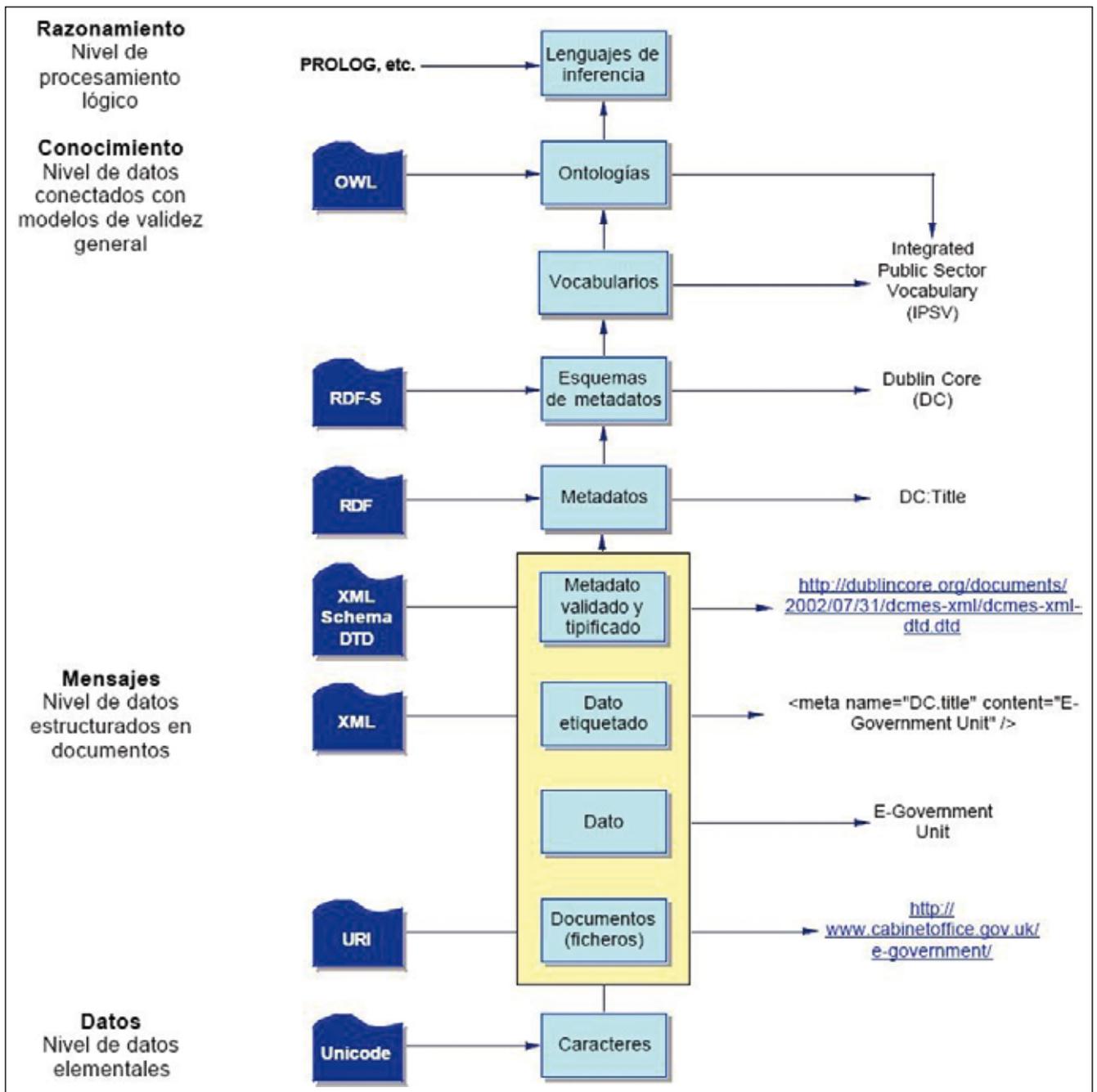


Figura 1. La arquitectura de la web semántica y los procesos cognitivos (adaptación de García-Marco, 2006)

La figura 1 resume el conjunto de estándares que van a convertir la www en un entorno de intercambio no sólo de documentos, sino también de datos y de relaciones entre los mismos (García-Marco, 2006)

Como se puede apreciar, la última capa de estándares está formada por el OWL, que constituye el estándar propuesto por el *World Wide Web Consortium* (2004) para la descripción formal de ontologías. Por fin, existe un estándar ISO que se puede utilizar para la presentación de ontologías de cara a la recuperación de la información y a la presentación de las relaciones para la navegación: se trata de los mapas temáticos o *topic maps* (*International Organization for Standardization*, 1998).

3. Mirando al futuro

Es fácil caracterizar el mundo actual desde el punto de vista informacional: internet se ha convertido en el nuevo entorno de distribución, almacenamiento, publicación y acceso a la información. El abaratamiento que ha conllevado en los procesos de información en todas sus fases ha sido enorme, y como consecuencia, tanto la información disponible como la accedida crecen a un ritmo vertiginoso. El proceso de digitalización que se está realizando con el objetivo de trasladar todos los contenidos a la Red es imparable y, como resultado, los que antes se distribuían por los diferentes canales de comunicación existentes –prensa y publicaciones periódicas, libro, cine, televisión, radio, dvd, audio cd,

etc.– se están trasladando al nuevo medio masivamente y cada vez con más rapidez. Por todo ello parece claro que los esfuerzos de la organización del conocimiento como disciplina –y de la documentación en general (v. g. **López-Yepes**, 1998, p. 17)– deben concentrarse en el nuevo medio, sin que eso deba suponer, lógicamente, descuido de los demás.

Por otra parte, la descomunal sobreabundancia de información que conlleva internet está produciendo un viraje decisivo en la manera de entender la gestión de la información: el acento que tradicionalmente se ponía en la conservación –del lado de los custodios– y en la velocidad y potencia de la recuperación –desde el punto de vista de los tecnólogos– ha cedido a un nuevo énfasis en la selección y el filtrado (**García-Marco** 2002). Esta nueva corriente no debería suponer en ningún caso descuido ni desprecio de las funciones de conservación y recuperación masiva, pero resulta absolutamente necesario en el nuevo contexto.

“La web semántica es a la www lo que los tesauros fueron a la recuperación en texto libre en los albores de la documentación automatizada”

Todo ello redunda en una renovada vigencia de las herramientas de control conceptual y terminológico

–que, como sabemos, en el campo de la biblioteconomía y documentación se han venido denominando lenguajes documentales-. La primera impresión es que esa situación privilegiada se refuerza al converger con el movimiento de las ontologías. En cierta manera, y en el campo específico de la recuperación de la información, la web semántica es a la www lo que los tesauros fueron a la recuperación en texto libre en los albores de la documentación automatizada. También entonces la fascinación inicial por la búsqueda mediante palabras extraídas cedió enseguida ante la necesidad del control terminológico, lo que llevó a su vez a la invención y a la generalización de los tesauros (**Moreiro**, 2007), que finalmente convergieron con las clasificaciones bajo el paraguas del movimiento analítico. De forma semejante, el proyecto de la web semántica pretende añadir a la www lo que le falta en el ámbito de la recuperación y el procesamiento de la información: granularidad⁵, precisión, exhaustividad e inferenciabilidad en el manejo a la información. Con estos nuevos objetivos, tanto los ingenieros del conocimiento como los profesionales de la información y la documentación ven multiplicada su relevancia y sus posibilidades de acción.

En este contexto, el despegue de la investigación sobre ontologías y el movimiento de convergencia de la investigación sobre organización del conocimiento con la que se realiza sobre ontologías aparecen llenos de oportunidades –y también de alguna amenaza-. La convergencia es tanto más interesante cuanto que, por un lado, se apoya en un sustrato común –la tradición ontológica y lógica, y el paradigma del procesamiento de

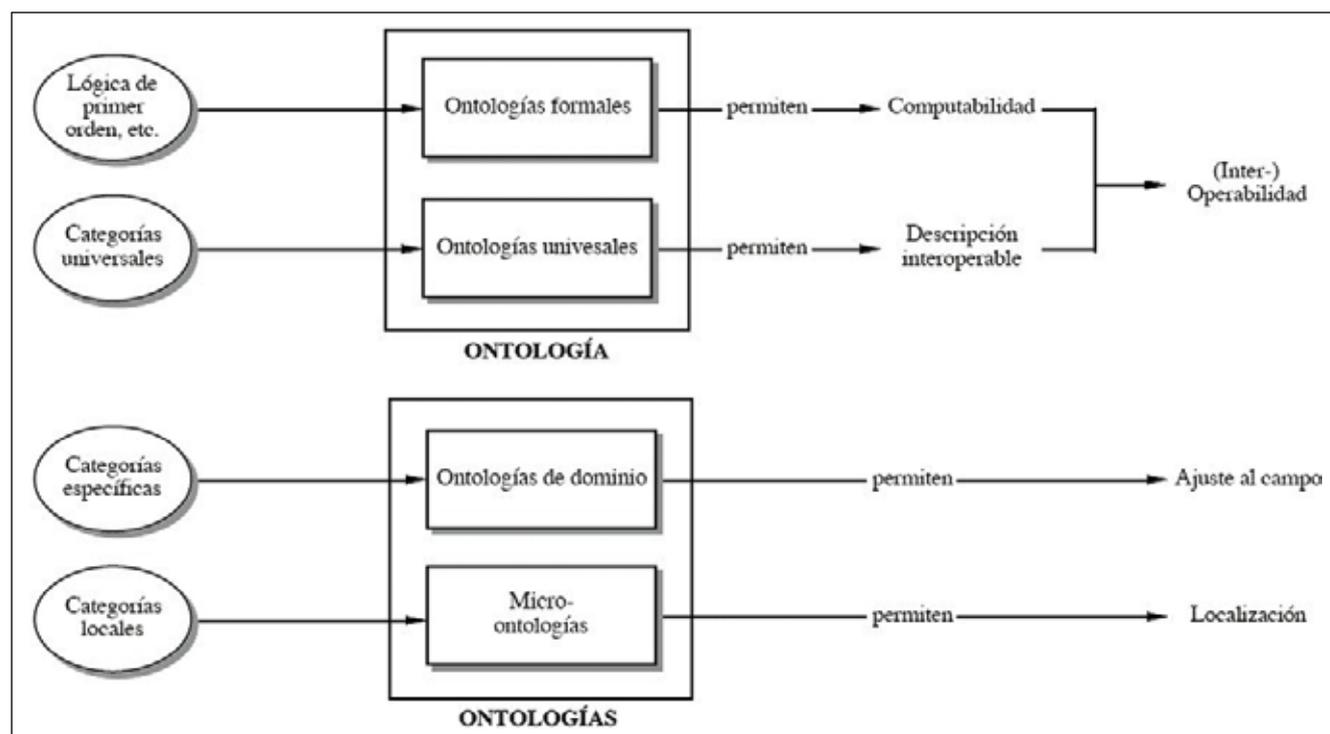


Figura 2. La ecología de las disciplinas ontológicas

la información—; y, por el otro, sucede de la mano de un conjunto de ciencias con las que ambas vienen históricamente manteniendo una relación estrecha (figura 2). De esta manera, la organización del conocimiento se inserta en un campo más amplio de investigación transdisciplinar —¿ciencias ontológicas?— que, por otra parte, cuenta entre sus fuerzas con un interesantísimo, potente y vital frente de investigación.

En el plano práctico, la convergencia entre sistemas de organización del conocimiento y ontologías promete dar abundantes frutos. La implementación de lenguajes documentales en forma de ontologías facilita la realización de operaciones lógicas sobre ellos: expansiones y restricciones de búsquedas, desambiguaciones, etc. Ya era posible en sistemas automatizados anteriores, pero de forma “cerrada”. La asunción de estándares abiertos en el entorno de la web semántica permite hacer todo lo anterior a una escala anteriormente impensable. Al compartir las diferentes ontologías es posible una “gramática” común, por ejemplo, detectar un término en diversas fuentes de información, comprobar sus términos genéricos y equivalencias, desambiguarlo, unir hipónimos de diferentes ontologías y presentarlos en un orden dado al usuario o resolver directamente los recursos a los que apuntan.

Los cambios raramente vienen tan sólo con oportunidades y en este caso, como en tantos otros, también hay que salir al paso de amenazas. Los peligros más importantes que plantea el proceso de convergencia a la organización del conocimiento y a la documentología son, por un lado, el ser arrastrado a la confusión inicial que rodea a todo este magma investigador, y, por el otro, el de ser absorbido y perder su entidad. La solución a todo ello pasa por clarificar campos de práctica, entender el lenguaje de las otras disciplinas, explicar el propio, interconectar —pero también deslindar— los programas de investigación, y determinar con precisión las aportaciones y tareas más importantes de unos y de otros.

En esta dirección, la figura 3 intenta bosquejar un deslinde de la I+D+I en el campo de las ontologías que lleva implícito una propuesta de división del trabajo. Se detectan cuatro ámbitos distintos, agrupados a su vez en dos bloques. El superior está constituido por “la ontología” propiamente dicha, en singular, que se ocuparía de estudiar y desarrollar los aspectos generales comunes a todas las implementaciones ontológicas con el fin de asegurar su interoperabilidad global, necesaria en internet. Dentro de ella, se identifican dos zonas diferentes: a) la superior, que se refiere a la formalización de las ontologías en lenguajes lógico-matemáticos de cara a asegurar su computabilidad y, en definitiva, la automatización de las tareas ontológicas de propósito general; y b) la inferior, que se refiere a la teoría de la

“Los cambios raramente vienen tan sólo con oportunidades y, en este caso, como en tantos otros, también hay que salir al paso de amenazas”

descripción de los entes en el nivel máximo de generalidad, y que tiene como tarea fundamentalmente el estudio de las categorías universales.

El bloque inferior se dedica a “las ontologías” en plural; esto es, a las concreciones de los modelos generales en los diferentes dominios específicos (disciplinas, subdisciplinas, etc.). Este bloque se concibe también dividido en dos capas: a) una referente a la disciplina o área disciplinar específica (las ontologías de dominio); y b) otra que se ocupa de los aspectos idiosincrásicos de una determinada comunidad o entidad (micro-ontologías). En esta arquitectura existen muchos nichos para las diferentes disciplinas de la figura 2, notablemente, la investigación sobre las categorías universales de organización del conocimiento, el lenguaje y los documentos; las propiedades formales de las ontologías; o las implementaciones tecnológicas. Sin embargo, los resultados en cualquiera de estas áreas serán mejores en la medida en que se reconozca la unidad del área y se realice un auténtico trabajo transdisciplinar.

Además del posicionamiento dentro de este esquema, el investigador y profesional de la información tiene ante sí la tarea de abordar tres frentes diferentes: el de la teoría, el de la práctica y el de la formación. En el primero urge integrar la reflexión ontológica realizada en la organización del conocimiento con la que se ha realizado en las tecnologías informáticas y la inteligencia artificial, más cercana a la filosofía analítica dentro de la inteligencia artificial y sus aplicaciones; y seguir con atención y cuidado los desarrollos que se producen en ese campo y en las otras disciplinas relacionadas para no perder el paso y quedar descolgados. En una segunda fase habrá que comunicar en el nuevo lenguaje aquellos aspectos propios en los que hay mucho que aportar, ahorrando a los colegas esfuerzos de “reinventar la rueda”.

En el frente de la práctica, es necesario aprender a expresar los diferentes lenguajes documentales en los nuevos formalismos y tecnologías, y desde esa experiencia y esa práctica colaborar en la tarea común: el desarrollo de vocabularios estructurados para la recuperación de información en el contexto de internet y sus aplicaciones corporativas (García-Marco et al, 2007).

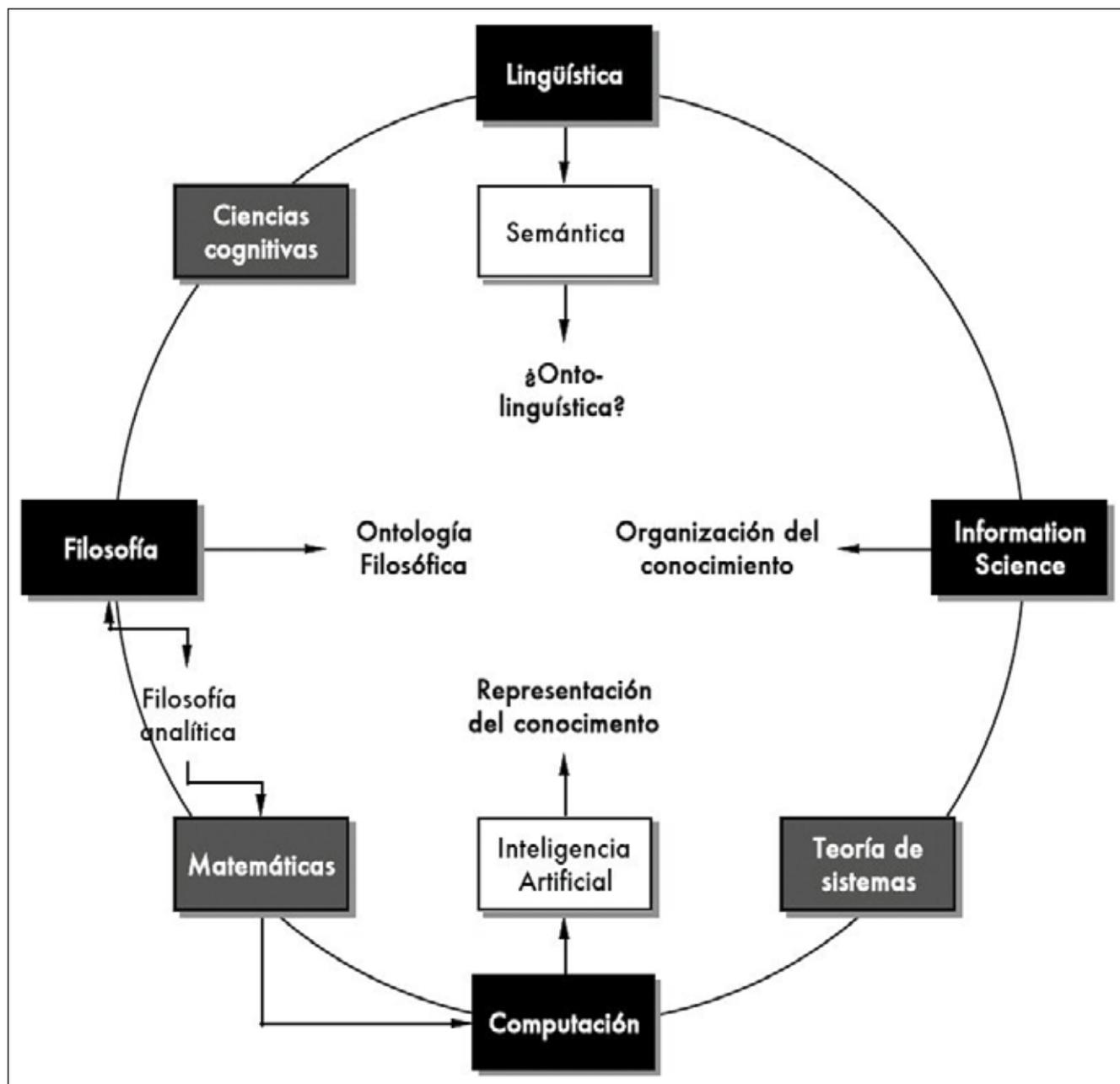


Figura 3. Ontología y ontologías

“Ante las ontologías, el investigador y profesional de la información debe abordar tres frentes diferentes: teoría, práctica y formación”

Por último, es necesario insistir en que para hacer posibles los objetivos anteriores, se debe realizar un importante esfuerzo en el campo de la formación inicial y continua, pues el éxito requiere que los profesionales conozcan y sean capaces de utilizar los estándares de representación y navegación ontológica, y especialmente las nuevas herramientas de gestión de ontologías y de utilización de las mismas en entornos

de recuperación de información; o al menos, que sean capaces de trabajar eficaz y eficientemente con las personas que se encarguen de estos menesteres. Esto no será posible sin reforzar, –siquiera mínimamente y con carácter instrumental– la formación lógico-matemática de los futuros profesionales de la información, una preocupación sentida y abordada ya por **Sayers** (1926), en la que ha insistido en nuestro país **Izquierdo** (1990) y que ha sido recientemente revisada por **Bufrem, Breda** y **Sorbías** (2007).

Teoría, práctica, formación y apertura interdisciplinar son, en definitiva, los indicadores que van a marcar el éxito en la nueva oportunidad que las ontologías y la web semántica abren al profesional e investigador de la información y la documentación.

Notas de la Redacción

1. Inferencia es el acto o proceso de derivar una conclusión basada solamente en lo que uno ya conoce.

Se usa en varios campos:

La inferencia humana (o sea, cómo los humanos sacan conclusiones) se estudia tradicionalmente en psicología cognitiva. La Lógica estudia las leyes de la inferencia válida.

Los estadísticos han desarrollado reglas formales de inferencia de datos cuantitativos. Los investigadores en inteligencia artificial realizan sistemas automáticos de inferencia.

Fuente: Wikipedia inglesa

2. En semántica lingüística se denomina hiperónimo a aquel término general que puede ser utilizado para referirse a la realidad nombrada por un término más particular.

Semánticamente, un hiperónimo no posee ningún rasgo semántico, o *sema*, que no comparta su hipónimo, mientras que éste sí posee rasgos semánticos que lo diferencian de aquél.

Por ejemplo, coche posee sólo los semas [+vehículo], [+con motor] y [+pequeño tamaño], que comparte con descapotable, mientras que descapotable posee además el rasgo [+sin capota], que lo diferencia de coche.

Al redactar un texto conviene utilizar hiperónimos para evitar la repetición de palabras ya empleadas anteriormente, como se hace en el siguiente ejemplo: “De repente, un *descapotable* rojo paró frente al banco. Del *automóvil* salieron dos individuos encapuchados, mientras otro esperaba en el *vehículo*”.

Fuente: Wikipedia española

3. En lingüística, un hipónimo es una palabra o frase cuya gama semántica está incluida dentro de otra palabra. Por ejemplo, escarlata, bermellón, carmín y carmesí son hipónimos de rojo (su hiperónimo).

Según **Victoria Fromkin** y **Robert Rodman** (“Introducción al lenguaje”, ISBN 0-03-018682-X), hipónimos son un conjunto de palabras relacionadas cuyo significado son los casos específicos de una palabra más general (así, por ejemplo, rojo, blanco, azul, etc., son hipónimos de color). La hiponimia es pues la relación entre un término general tal como el *polígono* y los casos específicos, como *triángulo*.

Fuente: Wikipedia inglesa

4. Instanciación. Veamos unos conceptos:

Clase. Define los aspectos abstractos de una cosa (objeto), incluyendo sus características (atributos, campos de aplicación o propiedades) y sus conductas (las que puede hacer, sus métodos o su comportamiento). Por ejemplo, la clase “perro” consistiría en rasgos compartidos por todos perros, como la especie (características), y la habilidad de ladrar (la conducta). Objeto. Es una *instancia* (un caso) particular de una clase. La clase “perro” define todos los perros posibles listando sus características y sus conductas; el objeto “Lassie” es un perro concreto, con versiones particulares de las características. Un “perro” tiene piel; “Lassie” tiene piel marrón y blanca. El objeto “Lassie” es una *instancia* (o sea, un caso particular) de la clase “perro”. El conjunto de valores de los atributos de un objeto particular se llama su estado. El objeto consiste del estado y del comportamiento que se ha definido en la clase del objeto.

Fuente: Wikipedia inglesa

5. La granularidad es una medida del tamaño de los componentes de un sistema. La granularidad es el tamaño relativo, la escala, el nivel de detalle o profundidad de la penetración que caracteriza un objeto o una actividad. Es “el grado en que se subdivide una entidad mayor”. Por ejemplo, un metro desmenuzado en milímetros tiene la granularidad más fina que un metro desmenuzado en centímetros”.

Fuente: Wikipedia inglesa

Referencias bibliográficas

Aristóteles. *Categorías*. Traducción del griego y prólogo de **Francisco de P. Samaranch**. 2a ed. Buenos Aires: Aguilar, 1965.

Baeza Yates R.; **Ribeiro Neto, B.** *Modern information retrieval*. Harlow, England: Addison Wesley, 1999.

Berners-Lee, T.; **Hendler, J.**; **Lassila, O.** “The semantic web”. En: *Scientific American*, 2001, v. 284, n. 5, pp. 34-43.

Berners-Lee, Tim. *Semantic web road map: an attempt to give a high-level plan of the architecture of the semantic www*. W3C, September 14-10-94. Consultado en: 12-07-07.

<http://www.w3.org/DesignIssues/Semantic.html>

British Standards Institute. *BS 8723, Structured vocabularies for information retrieval*. London: British Standards Institute, 2006-.

Bufrem, Liela Santiago; **Breda, Sônia Maria**; **Sorbías, Tidra Viana**. “The presence of logic in the domain of knowledge organization: interdisciplinary aspects of college curricula”. En: **Rodríguez Bravo, Blanca**; **Alvite Díez, María Luisa** (eds.). *La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico: Actas del VIII Congreso ISKO-España, 2007*. León: Universidad de León, Secretariado de Publicaciones, pp. 179-185.

Currás, Emilia. *Ontologías, taxonomía y tesauros: manual de construcción y uso*. 3ª ed., act. y amp. Gijón: Trea, 2005.

Dahlberg, I. *Ontical structures and universal classification*. Bangalore, Sarada Ranganathan Endowment for Library Science, 1978.

García-Marco, Francisco-Javier. “La literatura científica sobre lenguajes poscoordinados en España: de la divulgación del concepto a la internet”. En: *Documentación de las ciencias de la información*, 2002, v. 25, pp. 291-319.

García-Marco, Francisco-Javier. “Ontologías y documentación electrónica en las actividades públicas”. En: **Galindo, Fernando** (ed.). *Gobierno, derecho y tecnología: las actividades de los poderes públicos*. Madrid: Thomson-Civitas, 2006, pp. 173-225.

García-Marco, Francisco-Javier (coord.); **Agustín Lacruz, Carmen**; **Caro Castro, Carmen**; **Martínez Usero, José Ángel**; **San Segundo, Rosa**. “Proyectos internacionales de reforma y ampliación de las normas sobre tesauros para su adaptación a los nuevos contextos de integración e interoperabilidad en el entorno digital”. En: **Rodríguez Bravo, Blanca**; **Alvite Díez, María Luisa** (eds.). *La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico: Actas del VIII Congreso ISKO-España, 2007*. León: Universidad de León, Secretariado de Publicaciones, pp. 27-40.

Gruber, T. R. “A translation approach to portable ontologies”. En: *Knowledge acquisition*, 1993, June, v. 5, n. 2, pp. 199-220.

Hersh, William Richard. “Ontologies for information retrieval”. En: **Jorde, L. B.**; **Little, P.**; **Dunn M.**; **Subramaniam, S.** (eds.). *Encyclopedia of genetics, genomics, proteomics and bioinformatics*. London: John Wiley & Sons, Part 4, 2005.

International DOI Foundation. *The DOI Handbook, Version 4.2.0, released February 2005*. Oxford: International DOI Foundation (IDF), 2005. Consultado en: 12-07-07.

http://www.doi.org/handbook_2000/glossary.html

International Organization for Standardization. *ISO/IEC 13250, Information Technology-SGML Applications-Topic Maps*. Geneva: ISO, 1998.

Izquierdo Arroyo, José María. *Esquemas de lingüística documental*. Barcelona: PPU, 1990.

López Yepes, José. “Hombre y documento: del homo sapiens al homo documentator”. En: *Scire*, 1998, julio-diciembre, v. 4, n. 2, pp. 11-22.

McGuinness, D. L. “Ontologies come of age”. En: **Fensel, D.**; **Hendler, J.**; **Lieberman, H.**; **Wahlster, W.** (eds.). *Spinning the semantic web: bringing the world wide web to its full potential*. Cambridge: MIT Press, 2002.

Moreiro, José Antonio. “Evolución paralela de los lenguajes documentales y la terminología”. En: **Rodríguez Bravo, Blanca**; **Alvite Díez, María Luisa** (eds.). *La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico: Actas del VIII Congreso ISKO-España, 2007*. León: Universidad de León, Secretariado de Publicaciones, pp. 27-40.

Nicles, Matthias; **Pease, Adam**; **Schalley, Andrea C.**; **Zaefferer, Dietmar**. “Ontologies across disciplines”. En: **Schalley, Andrea C.**; **Zaefferer, Dietmar**. *Ontolinguistics: how ontological status shapes the linguistic coding of concepts*. Berlin, etc.: Mouton de Gruyter, 2007, pp. 23-67.

Ontolingua. Standford: Stanford University, 2005. Consultado en: 12-07-07.

<http://www.ksl.stanford.edu/software/ontolingua/>

Patel-Schneider P. F.; **Swartout, B.** “Description-logic knowledge representation system specification”. En: *KRSS group of the ARPA knowledge sharing effort*. Consultado en: 12-07-07.

<http://www-db.research.bell-labs.com/user/pfjps/papers/krss-spec.ps>

Poli, R. "Glanzing at the problems of contemporary ontology". En: *Scire*, 2002, enero-junio, v. 8, n. 1, pp. 17-40.

Sayers, W. C. Berwick. *A manual of classification for librarians & bibliographers*. London: Grafton, 1926.

Schalley, Andrea C.; Zaefferer, Diezmar. *Ontolinguistics: how ontological status shapes the linguistic coding of concepts*. Berlin, etc.: Mouton de Gruyter, 2007.

Stephen, R.; Lenat, D. "Mapping ontologies into Cyc". En: *AAAI 2002 conference workshop on ontologies for the semantic web*, 2002.

Vickery, B. C. "Ontologies". En: *Journal of information science*, 1997, v. 23, n. 4, pp. 277-286.

World Wide Web Consortium (1994-2004). *Semantic web activity*. Cambridge, Keio, Paris: W3C, 1994-2004. Consultado en: 12-07-07. <http://www.w3.org/2001/sw/>

World Wide Web Consortium (2004). *OWL. Web Ontology Language Overview. W3C Recommendation 10 February 2004*. Cambridge, Keio, Paris: World Wide Web Consortium, 2004. Consultado en: 12-07-2007. <http://www.w3.org/TR/2004/REC-owl-features-20040210/>

Francisco-Javier García-Marco, Área de bibliotecología y documentación, Universidad de Zaragoza. jgarcia@unizar.es

El profesional de la

información

Precios 2008

<http://www.elprofesionaldelainformacion.com>

REVISTA IMPRESA + ACCESO ONLINE (ISSN 1386 6710 + ISSN-e 1699-2407)

Suscripción normal: 147,2 € + IVA = 153 €

Tarifa reducida para personas individuales*: 75 € + IVA = 78 €

* **exclusivamente a domicilios particulares**

Coste adicional de correo aéreo:

- Europa (menos España) 30 €
- Américas y resto del mundo 45 €

Acceso online (incluye 1 clave de acceso: username-password)

Claves de acceso adicionales: 85 € + IVA = 88,4 € / cada una

NÚMERO SUELTO ACTUAL O ANTIGUO

Número suelto actual o antiguo: 25 € + IVA = 26 €

Coste adicional de correo aéreo:

- Europa (menos España) 8 €
- Américas y resto del mundo 14 €

SÓLO ACCESO ONLINE **¡¡ NUEVO !!**

Únicamente acceso a versión electrónica 85 € + IVA = 88,4 €

NOTAS:

1. De cara a las tarifas de 2009, la editorial está estudiando la introducción de suplementos de precio según el número de bibliotecas existentes en cada institución.
2. El IVA aplicado es España es del 4%
3. El período 1992-2006 es de acceso libre y gratuito desde nuestra web: <http://www.elprofesionaldelainformacion.com>