

D.5. El nacimiento y el despegue de la investigación moderna sobre ontologías

Por **Francisco-Javier García-Marco**

García-Marco, Francisco-Javier. «El nacimiento y el despegue de la investigación moderna sobre ontologías». En: *Anuario ThinkEPI*, 2008, pp. 122-125.



Resumen: 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 mediante un análisis bibliométrico. 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. La investigación sobre las ontologías formales para la representación del conocimiento constituye un frente de investigación que está cambiando el curso de la investigación en el campo de la organización y la representación del conocimiento, y que incluso está modificando la terminología en el campo de las ciencias de la documentación, como ha ocurrido también con el movimiento de los metadatos en catalogación.

Palabras clave: Ontologías, Organización del conocimiento, Ingeniería del conocimiento, Internet, Web semántica, Análisis bibliométrico, Tendencias de investigación.

Title: The birth and growth of modern research on ontologies

Abstract: Recent developments in ontology research in the fields of Library and Information Science (LIS), Computer Sciences, and other related disciplines are examined through a bibliometric survey. Results confirm that the current research surge must be contextualized in terms of Internet developments and, especially, the Semantic Web Project. Research on formal ontologies for knowledge representation is changing the course of research on knowledge organization and representation, and is even causing changes in LIS terminology, as occurred previously with the metadata movement in the field of cataloguing.

Keywords: Ontologies, Knowledge organization, Knowledge engineering, Internet, Semantic Web, Bibliometric analysis, Research trends.

DURANTE MUCHOS AÑOS, el término ontología ha constituido un término oscuro incluso para la mayoría de las personas cultas. No era completamente desconocido, pues se estudiaba en los cursos de filosofía en las enseñanzas medias; pero tampoco era bien comprendido –se trata de un concepto de la máxima abstracción que se estudia en un momento en que la capacidad de abstracción formal está en desarrollo– y raramente se aterriza en problemas prácticos de interés de los alumnos. Se convertía en un concepto para iniciados que, por estudios o interés, realizaban cursos y lecturas de carácter normalmente filosófico.

Sin embargo, en los últimos años ha cobrado importancia creciente en diversas dis-

ciplinas y, particularmente en el campo que nos interesa, esto es, en las Ciencias de la Documentación. La experiencia de que esto está sucediendo es clara para los especialistas (Currás, 2005), pero parece necesario explicitar con datos esa evidencia e intentar abordar de esta manera qué es lo que está sucediendo, las tendencias y sus posibles explicaciones. También es importante caracterizar y contextualizar las nuevas tendencias dentro de la preocupación ontológica que ha sido constante –aunque de forma equilibrada con otras materias– en las Ciencias de

Los artículos sobre ontologías son muy raros hasta 1992, año en que el tema comienza a despegar

la Documentación, especialmente dentro de la corriente conocida como “organización del conocimiento”, que ha terminado por aglutinar lo que en el ámbito latino se conoce como “lenguajes documentales” y en el anglosajón como *indexing languages*.

Una búsqueda en la base de datos *Library and Information Science Abstracts (Lisa)* permite hacerse una idea bastante clara de la evolución del término y de la producción científica que se produce entorno a él en el campo de las ciencias de la documentación y su entorno (Figura 1)¹.

Como se puede observar, los artículos sobre ontologías son muy raros hasta 1992, año en que el tema comienza a despegar. Interesa destacar que los trabajos “ontológicos” de los años 1977-1991 están relacionados o bien con aspectos epistemológicos de la *Information Science* (**Brookes**, 1977, 1979; **Ivori**, 1991) o bien con la implicación de la ontología en el diseño de sistemas de organización del conocimiento, en concreto de las clasificaciones bibliográficas (**Dahlberg**, 1978, 1992a y 1992b; **Biagetti**, 1987). Es decir, se desenvuelven en el entorno de la aplicación de la filosofía y sus sistemas a la ciencia de la información en general y la organización del conocimiento en particular.

El primer trabajo relacionado con las ontologías en sistemas de información –esto es, con las ontologías en el campo de la computación– aparece en el año 1993 (**Vet y Mars**, 1993), coincidiendo significativamente con el despegue del área de investigación.

En 1994, cuando el número de referencia prácticamente se dobla, se produce la inflexión: 7 de las 10 referencias son del campo de la inteligencia artificial y la informática en general. Sin embargo, el número de documentos permanece todavía bajo. Los tres picos de una tendencia al crecimiento durante el resto de la década de los noventa provienen de números monográficos o

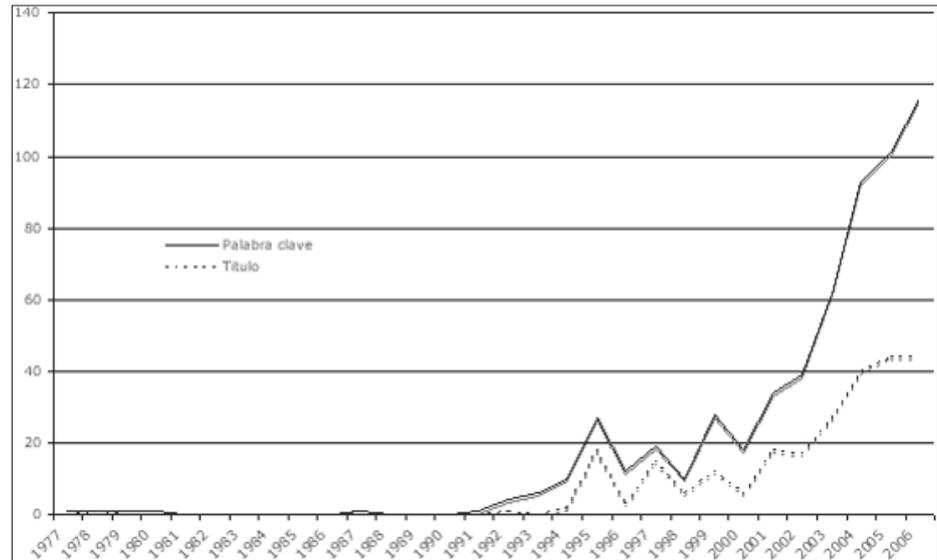


Figura 1. Evolución de los artículos sobre ontologías en la base de datos Lisa.

congresos en diversos campos de las ciencias de la computación: el *International Journal of Human-Computer Studies* dedicó sendos monográficos en 1995 y 1997 –el primero dedicado a las ontologías formales y el segundo a las explícitas–, y en 1999 se celebró una sesión especial del *Ieee Intelligent Systems*. En 1999 la tendencia al crecimiento parece consolidada, pues aparecen ya muchos artículos en diversas fuentes, casi todos relacionados con la aplicación de las ontologías en las tecnologías de la información y la comunicación. A partir de 2001 las publicaciones no dejan de crecer y en los cuatro años que van desde el 2003 a 2006 los artículos se duplican en número.

Para situar esta evolución dentro de un contexto más global, se contrastaron las tendencias obtenidas de la base de datos *Lisa* con las de las tres que componen el *Web of Science* del *Institute for Scientific Information*². Esto es, se buscó la aparición del concepto ontología –a partir de la raíz “ontolog-”– en todas las disciplinas. Se observa que el interés por el campo crece en dos tirones a principios y mediados de los años setenta, se mantiene suavemente al alza, y crece de nuevo visiblemente entorno a 1994, con un pico en el 2001 y un despegue impresionante desde el 2002 (Figura 2).

En los años 1956-57 todos los artículos eran filosóficos o de ciencias sociales con un fuerte componente filosófico; en 1969-72 la tendencia era la misma con un 73,78 % de

los artículos de filosofía; en 1975-1978 el porcentaje de artículos de filosofía se reduce al 66,66 %, aumentando los de ciencias sociales y humanas. En 1989 **Tim Berners Lee** inventa la *World Wide Web*; en 1990, aparece en internet; y en abril de 1993 el *Cern* hace público que será un servicio gratuito. La comunidad de investigadores en Inteligencia Artificial y otros campos de la computación ve pronto el potencial del nuevo medio. En 1993 aparece el famoso artículo de **Gruber** (1993) y en 1994 aparecen otros tres trabajos (un 2,5 % del total). Por fin, en los años 1995-1996 aparecen los primeros artículos muy citados procedentes de revistas de informática, fundamentalmente del campo de la inteligencia artificial (**Tu et al**,

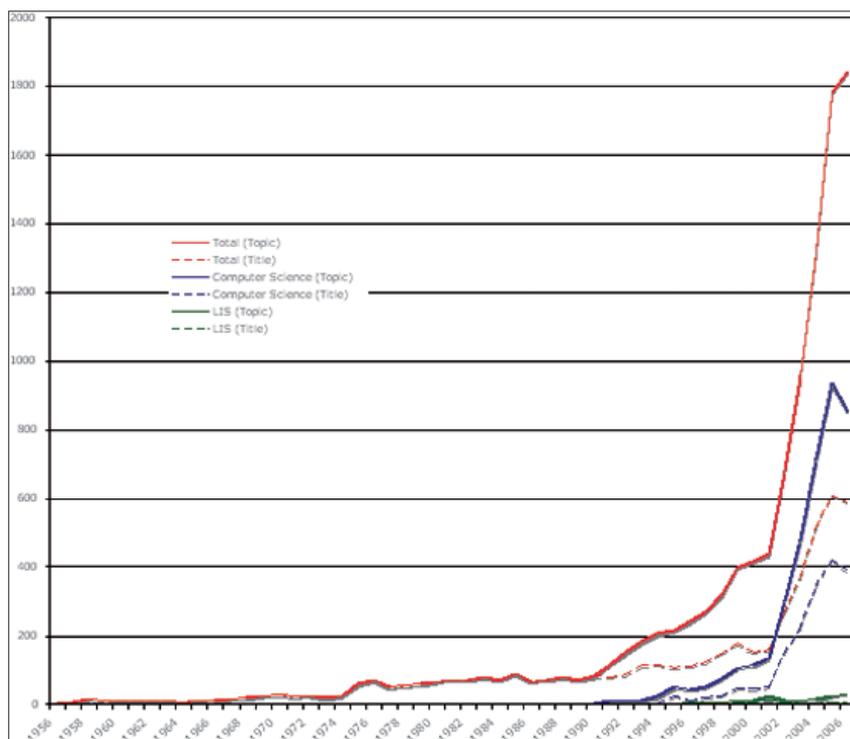


Figura 2. Evolución de los artículos sobre ontologías en la base de datos Web of Science (2006 incompleto).

1995; **Uschold, Gruninger**, 1996; etc.), pero el grueso de los trabajos empieza a aparecer a partir de 2002, incrementándose a un ritmo muy rápido a partir de entonces (Figura 2). El hecho fundamental que dispara el crecimiento de la Red es la consolidación del proyecto de la Web semántica, que **Tim Berners-Lee** y sus colaboradores difunden en 2001 entre un público amplio en un famoso artículo en la revista *Nature* (**Berners-Lee, Hendler y Lassila**, 2001), pero que ya llevaba unos años en curso (**Berners-Lee**, 1998).

La tendencia, por lo tanto, se puede considerar interdisciplinar; y se puede afirmar, sin lugar a dudas, que el interés por el tema en el campo de las Ciencias de la Documentación forma parte de un movimiento más amplio, liderado por las tecnologías de la información y la comunicación, que cataliza con la puesta en marcha del proyecto de la Web semántica. Hay que tener en cuenta, además, que el campo de investigación de la *Information Science* no es disjuncto con la computación; antes al contrario, como se puede rápidamente inferir de los títulos de revistas incluidos en las bases de datos que hemos estudiado.

Que el despegue de la investigación en ontologías se debe al interés por sus aplicaciones informáticas se constata también al ver la disciplina de asignación de los trabajos en el *Web of Science*. A pesar de los datos que hemos dado para el principio del período –en que la mayoría eran estudios procedentes de la filosofía con una tendencia al aumento de los que se producen en otras ciencias sociales y humanas—, los artículos procedentes de las áreas técnicas son al menos un tercio de las referencias recuperadas, más aún si se consideran las aplicaciones en otras áreas científicas.

En los años de la explosión de trabajos (2002-2007), los artículos de Filosofía se habían reducido ya tan sólo el 10 %, pasando al tercer lugar, y los de Informática e Inteligencia Artificial en sus distintas especialidades alcanzan entorno al 40 % de los trabajos, dependiendo de las categorías que se incluyan. Sin embargo, los trabajos dentro de las áreas de la computación no absorben la totalidad de los trabajos, sino que son la punta de lanza de un interés por el tema en un conjunto de disciplinas muy amplio.

Por último, interesa señalar que, frente a los estudios de la época anterior, los más

El despegue de las ontologías en Documentación aparece muy conectada al movimiento general para abordar el problema terminológico-conceptual dentro de la informática y, concretamente, de internet a través de la propuesta de la Web semántica

recientes suelen ser de carácter aplicado y orientados a la solución de problemas en un gran número de ámbitos del conocimiento.

Así pues, el despegue de las ontologías en el campo de las ciencias de la documentación aparece como intensamente conectada con el movimiento general para abordar el problema terminológico-conceptual dentro de la informática y, concretamente, dentro de internet a través de la propuesta de la Web semántica.

De hecho, la impresión a la vista de los datos es que la investigación sobre las ontologías formales para la representación del conocimiento constituyen una riada que está cambiando el curso de la investigación en el campo de la organización y la representación del conocimiento, y que incluso está modificando la terminología en el campo de las ciencias de la documentación, como ha ocurrido también con el movimiento de los metadatos.

Notas

1. La búsqueda se realizó el día 9 de julio de 2007, recuperando las raíces "ontolog*" y "ontic*" en la base de datos *Library and Information Science Abstracts (LISA)* en los campos de cualquiera o palabras clave, título, descriptores y clasificación. La raíz *ontic** sólo produjo un resultado suplementario (Marsden, 1979). Los campos de descriptores y clasificación no ofrecieron resultados, lo que es entendible, pues el bajo número de documentos producidos no exige su codificación con un descriptor –y mucho menos un clasificador– específico.

2. La búsqueda se realizó en el *Web of Science* recuperando la raíz "ontolog*" en el campo del título (TI) y de tópico (TS) el 10 de julio de 2007, y filtrando en dos fases los resultados por las categorías temáticas correspondientes a *Computer Science e Information Science & Library Science*.

Referencias

Berners-Lee, T. *Semantic web road map: an attempt to give a high-level plan of the architecture of the Semantic WWW*. W3C, September 1998/10/14. Consultado en: 12-07-2007.
<http://www.w3.org/DesignIssues/Semantic.html>

Berners-Lee, T.; Hendler, J.; Lassila, O. "The Semantic Web". En: *Scientific American*, 2001, v. 284, n. 5, pp. 34-43.

Brookes, B. C. "The developing cognitive viewpoint in information science". En: *Journal of informatics*, 1977, v. 1, n. 2, pp. 55-61.

Brookes, B. C. "The foundation of information science, part 1: Philosophical aspects". En: *Journal of information science*, v. 2, n. 3/4, pp. 125-133.

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

Dahlberg, I. "Knowledge, organisation and terminology: philosophical and linguistic bases". En: *Information services and use*, 1992, v. 12, n. 1, pp. 65-71.

Dahlberg, I. *Ontical structures and universal classification*. Bangalore, Sarada Ranganathan Endowment for Library Science, 1978, 64s.

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

International Journal of Human-Computer Studies, Feb.-March 1997, v. 46, n. 2/3. A special issue devoted to using explicit ontologies in knowledge-based system development.

International Journal of Human-Computer Studies, Nov.-Dec. 1995, v. 43, n. 5/6. Special issue devoted to the role of formal ontology in information technology, and in particular knowledge representation.

Ivari, J. "A paradigmatic analysis of contemporary schools of IS development". *European journal of information systems*, 1991, v. 1, n. 4, pp. 249-272.

Tu, S. W.; Eriksson, H.; Gennari, J. H.; et al. "Ontology-based configuration of problem-solving methods and generation of knowledge-acquisition tools - application of Protege-II to protocol-based decision-support". En: *Artificial intelligence in Medicine*, v. 7, n. 3, June 1995, pp. 257-289.

Uschold, M.; Grüninger, M. "Ontologies: principles, methods and applications". En: *Knowledge engineering review*, v. 11, n. 2, June 1996, pp. 93-155.

Vet, P. E. v. d.; Mars, N. J. I. "Structured system of concepts for storing, retrieving and manipulating chemical information". En: *Journal of chemical information and computer sciences*, 1993, v. 33, n. 4, pp. 564-568.