
Ontologías terminológicas en la organización del conocimiento: una aproximación bibliométrica

Terminological ontologies in knowledge organization: a bibliometric approach

Adriana SUÁREZ SÁNCHEZ

Instituto de Investigaciones Bibliotecológicas y de la Información, Universidad Nacional Autónoma de México, Circuito Interior de Ciudad Universitaria s/n Pisos 11, 12 y 13 de la Torre II de Humanidades, Ciudad Universitaria, Alcaldía Coyoacán, Ciudad de México, México. asuarez@iibi.unam.mx

Resumen

Las ontologías terminológicas son un sistema de organización del conocimiento que en los últimos años ha ganado gran popularidad debido a sus ventajas en la representación y estructuración de dominios en el contexto digital. En la literatura se han tratado sus aspectos teóricos y sus aplicaciones, pero no se cuenta con suficientes estudios bibliométricos. En el presente artículo se analiza la producción científica sobre ontologías terminológicas desde una aproximación bibliométrica. Se estudia la producción científica sobre el tema registrada en Web of Science, Scopus, WorldCat, Dialnet, Redalyc y Google Académico. Se analiza: el rango temporal de la literatura, su crecimiento diacrónico, la tipología de los documentos publicados, los idiomas de publicación, las publicaciones periódicas con mayores trabajos, la productividad de los autores, el grado de colaboración y el rango de citación. Se encontraron 216 documentos producidos por 420 autores en un rango temporal de 1997 a 2022. Los documentos publicados corresponden a libros, capítulos de libros, artículos en revistas académicas, trabajos en actas de congresos y tesis. El idioma de publicación preponderante es inglés. El núcleo de productores se forma de 13 autores. La producción sobre el tema muestra un pico de crecimiento entre 2007 y 2016, actualmente registra una tendencia alta de producción y en la literatura prevalece un alto grado de colaboración.

Palabras clave: Ontologías terminológicas. Organización del conocimiento. Bibliometría. Producción científica.

1. Introducción

Las ontologías son “una explicitación de una conceptualización compartida” (Gruber, 1995, p. 908) o bien “un marco conceptual particular en el nivel del conocimiento” (Guarino, 1995, p. 633) que se concreta en símbolos y es utilizado con fines prácticos de gestión del conocimiento.

Por mucho tiempo, tales sistemas fueron asunto exclusivo de las ciencias computacionales, en tareas de modelado de conocimiento en agentes computacionales; sin embargo, al paso del

Abstract

Terminological ontologies are a knowledge organization system that has gained great popularity in recent years due to its advantages in the representation and structuring of domains in the digital context. Theoretical aspects and applications have been discussed in the literature, but there are not enough bibliometric studies. This article analyses the scientific production on terminological ontologies from a bibliometric approach. The scientific production on the subject registered in Web of Science, Scopus, WorldCat, Dialnet, Redalyc and Google Scholar is studied. The following aspects are analysed: the temporal range of the literature, its diachronic growth, the typology of the published documents, the languages of publication, the journals with the largest number of papers, the productivity of the authors, the degree of collaboration and the citation rank. We found 216 publications by 420 authors over a period from 1997 to 2022. The published documents correspond to books, book chapters, articles in academic journals, papers in conference proceedings and dissertations. The predominant language of publication is English. The core group of producers consists of 13 authors. The production on the subject shows a growth peak between 2007 and 2016, currently shows a high production trend and a high degree of collaboration prevails in the literature.

Keywords: Terminological ontologies. Knowledge organization. Bibliometrics. Scientific research.

tiempo se expandieron hacia diversos campos de saber que organizan el conocimiento, la información y los recursos de información, tal es el caso de la bibliotecología.

García Marco (2008) señala que la expansión del término hacia la bibliotecología constituye la punta del iceberg de un proceso de reconfiguración disciplinar, como resultado de la integración de diversas ciencias en el marco del desarrollo tecnológico info-comunicacional. En el ámbito de la representación, organización y recuperación

de la información y el conocimiento, las ontologías son un sistema de organización que establece el vocabulario de un campo de saber, mediante un conjunto de conceptos, atributos y relaciones entre tales términos. A grandes rasgos, las ontologías pueden considerarse un sistema de organización del conocimiento (Hodge, 2000) con una sintaxis comprensible para los ordenadores (Pedraza, Codina y Rovira, 2007).

Como sistemas de organización del conocimiento, las ontologías cuentan con un gran abanico de funciones (vocabulario controlado, modelado de procesos, soporte de algoritmos para agentes artificiales, etc.) que se asocian con una amplia tipología de los instrumentos. Como resultado, entre los diversos tipos de ontologías persiste una amplia categorización que engloba: ontologías de dominio, ontologías de tarea, ontologías terminológicas, ontologías ligeras, ontologías pesadas, entre otras.

Las ontologías terminológicas, de especial interés en esta investigación, derivan de la tipología propuesta por Van Heijst, Schreiber y Wielinga (1997) quienes hablan de ontologías terminológicas y ontologías modeladoras. Las ontologías terminológicas especifican los términos que son usados en un dominio mientras que las modeladoras se centran en las simbolizaciones del conocimiento y son más ricas en relaciones axiomatizadas. Lacasta, Nogueras-Iso y Zarazaga-Soria (2010, p. 24) señalan que: “dos grandes tipos de modelos de ontologías han sido estudiadas: por un lado, ontologías terminológicas [...] por otro lado, ontologías axiomatizadas o formales”.

Las ontologías modeladoras o axiomáticas son ideales en la representación de fragmentos de conocimiento en agentes computacionales, mientras que las ontologías terminológicas, con funciones de control, explicitación, direccionamiento o vinculación léxica son empleadas con fines terminológicos y como vocabularios controlados en la organización del conocimiento y la indexación de recursos de información. Tenemos entonces que las ontologías con fines terminológicos son teorías de dominios que, mediante el uso del lenguaje, explicitan las entidades de un área de saber, las organizan en un esquema jerárquico y determinan sus propiedades y relaciones, proporcionando un vocabulario y vínculos entre elementos denominativos.

En las últimas dos décadas, las ontologías terminológicas han acaparado la atención de los bibliotecólogos, en vista de sus características como un vocabulario controlado y una estructura categorizadora que las agrega al amplio abanico de los sistemas de organización del conocimiento (SOC) (Hodge, 2000) entre los que es posible

mencionar encabezamientos de materia, tesauros, taxonomías, mapas tópicos, ontologías, entre otros. Aunado a ello, se vislumbran como un SOC con gran potencial para el modelado temático-terminológico de campos de conocimiento (endometriosis, tumores cerebrales, plantas suculentas, canciones folclóricas búlgaras, Colección Obras Maestras del Museo del Prado, etc.) y la organización de recursos de información relacionados.

A partir de su relevancia en la organización del conocimiento, la producción científica sobre ontologías terminológicas ha sido un asunto relevante en las últimas dos décadas; no obstante, los estudios bibliométricos sobre el tema son escasos.

García Marco (2008) establece que los artículos sobre ontologías eran muy raros hasta 1992. Luego, se observó un aumento en la producción científica durante el periodo de 1990 a 1998. Por último, hacia principios de milenio hubo una consolidación del tema, especialmente en el contexto de la web semántica y el problema terminológico-conceptual en ambientes informatizados.

Un segundo trabajo que aportó información respecto a la producción científica sobre ontologías fue desarrollado por Zhu, Kong, Hong, Li y He (2015). En su artículo estudiaron la literatura registrada en Web of Science y determinaron que la investigación ontológica ha entrado en una fase de auge, caracterizada por un alto grado de colaboración entre autores.

En 2017, Netto y Lima estudiaron la producción bibliométrica sobre modelos y tecnologías para la visualización de las ontologías. Detectaron un nicho de investigación reciente, con un rango de publicación entre 2007 a 2017. Las autoras concluyeron que muchas aristas de las ontologías se encuentran en un estado de auge de investigación y que la producción científica crece día a día, por lo que es necesario llevar a cabo estudios bibliométricos que permitan analizar el dominio, los autores, las fuentes, etc.

A partir del estado del arte del tema, revisado en los párrafos anteriores, es posible recapitular dos aspectos: las múltiples aristas de estudio de las ontologías (epistemología, estructura, tecnologías, formas de visualización, entre otros) y los pocos estudios bibliométricos, que derivan en la necesidad de analizar su producción científica.

Consecuentemente, surgen algunas preguntas de investigación:

- ¿Cuál es la tendencia de la producción científica sobre ontologías terminológicas en funciones de organización del conocimiento?
- ¿A través de qué tipo de documentos se distribuye dicha producción científica?

- ¿En qué idioma se produce la literatura sobre ontologías terminológicas?
- ¿Quiénes son los autores con mayor producción científica?
- ¿La colaboración se observa como una característica de dicha literatura?
- ¿Qué prospectiva de crecimiento presenta el tema en los tiempos futuros?

En aras de dar respuesta a las preguntas antes esbozadas, el presente artículo tiene por objetivo: analizar desde una perspectiva bibliométrica la producción científica sobre ontologías terminológicas en funciones de organización del conocimiento. Se enfatiza la relevancia del trabajo a partir de la trascendencia que las ontologías terminológicas poseen en la organización terminológica-temática de dominios de conocimiento, su uso como lenguajes de indización y su potencial en el arreglo de recursos de información albergados en el espacio web.

2. Metodología

Para lograr el objetivo establecido, se siguió una metodología descriptiva-cuantitativa, fundamentada en la técnica de análisis bibliométrico. Dicha técnica de investigación implica “el uso de métodos matemáticos y estadísticos para analizar el desarrollo histórico de un cuerpo específico de literatura, especialmente su autoría, publicación y uso” (Reitz, 2007).

Los estudios bibliométricos son un método nuclear en las ciencias de la información y consideran dos enfoques: el mapeo gráfico y el análisis de funcionamiento. Entre ellos, la presente investigación hace uso del análisis de funcionamiento, un método que analiza aspectos generales de la producción científica (idioma, autores, tipología de la publicación) así como indicadores de productividad y citación (Aparicio, Iturralde y Maceda, 2019).

Por producción científica se entiende el conjunto de obras (libros, artículos, capítulos de artículo, tesis y trabajos compendiados en actas de congresos) resultantes del trabajo intelectual de la investigación desarrollada en un área de conocimiento (Maletta, 2009). Se optó por medir la producción científica a la luz del análisis bibliométrico porque es claro indicio de la cantidad de investigación producida por un científico, una institución o un país (Spinak, 1996). Piedra Salomón y Martínez Rodríguez (2007, p. 34) agregan que “la productividad científica refleja el resultado de investigaciones científicas traducidas en nuevo conocimiento.”

Una vez establecida la perspectiva de estudio, la investigación comprendió tres etapas: obtención de los registros bibliográficos, sistematización y limpieza de registros bibliográficos, y procesamiento de los datos.

2.1. Obtención de los registros bibliográficos

Para localizar la producción científica sobre ontologías terminológicas en funciones de organización del conocimiento se revisaron las siguientes bases de datos (Tabla I):

Base de datos	Descripción
Scopus	Es una base de datos especializada en información científica, técnica, ciencias médicas y sociales. Incluye artículos de diversos campos de conocimiento. Contiene más de 29 millones de resúmenes de más de 18,500 revistas científicas tanto comerciales como de acceso abierto.
Web of Science	Es una base de datos que contiene información sobre investigación multidisciplinaria, proveniente de más de 3,300 revistas líderes mundiales en las ciencias puras, ciencias sociales, artes y humanidades. Es un referente sobre la producción científica internacional.
Dialnet	Es un portal de difusión de la producción científica hispana. Su base de datos fue creada por la Universidad de la Rioja (España). Incluye índices de revistas científicas y humanísticas de España, Portugal y Latinoamérica. Incorpora también libros, tesis doctorales, congresos, homenajes y otro tipo de documentos.
WorldCat	Es una base de datos sobre recursos albergados en las colecciones de numerosas bibliotecas del mundo. Es gestionada por la cooperativa OCLC.
Google Académico	Es un motor de búsqueda de Google especializado en la búsqueda de bibliografía científico-académica. Indexa bibliotecas, repositorios, bases de datos bibliográficas, etc. Contiene registros de libros, artículos, congresos, informes, tesis. Tesinas y archivos depositados en repositorios.

Tabla I. Bases de datos consultadas

Luego de establecer las bases de datos que se consultarían, se emprendió la búsqueda de los

documentos a partir de un conjunto de descriptores en tres idiomas (inglés, español y portugués). La recuperación de los registros consideró toda la producción científica existente en cada una de las bases de datos hasta diciembre del año 2022.

En la detección de los documentos, se emplearon los siguientes descriptores:

- “ontologías terminológicas” [español]
- “terminological ontologies” [inglés]
- “ontologias terminológicas” [portugués]
- “ontologías” and “términos” [español]
- “ontologies” and “terms” [inglés]
- “ontologias” and “termos” [portugués]
- “ontologías” and “términos” and “organización de conocimiento” [español]
- “ontologies” and “terms” and “knowledge organization” [inglés]
- “ontologias” and “termos” and “organizaçãodo conhecimento” [portugués]

La consulta de tales descriptores se efectuó en los campos de título, temas, resumen y palabras clave, según los campos de búsqueda disponibles en cada una de las bases de datos en las que se detectaron los trabajos (1). A partir del trabajo señalado, se recuperaron 312 registros sobre ontologías terminológicas en funciones de organización del conocimiento que incluyeron tanto trabajos teóricos como aplicados.

2.2. Sistematización y limpieza de registros bibliográficos

Las referencias localizadas en las bases bibliográficas antes enunciadas se exportaron a una base de datos en EndNote X8. Cada documento se revisó cuidadosamente para verificar que tratara sobre ontologías terminológicas en tareas de organización del conocimiento. Se verificó que no hubiese registros duplicados y se normalizaron los nombres de los autores, los títulos de los documentos, los nombres de las revistas y los nombres de los congresos. Tras dichas actividades, de los 312 registros obtenidos se eliminaron 96. Luego de tal acción, el corpus de análisis, que sirvió de base al presente estudio, se constituyó de 216 trabajos sobre ontologías terminológicas publicados desde 1997 hasta 2022 por 420 autores.

2.3. Procesamiento de los datos

En el análisis bibliométrico del corpus sobre ontologías terminológicas se usó el programa SPSS 25.0.

La productividad de los autores se analizó con la aplicación de la ley de Lotka (Urbizagástegui Alvarado, 1999), según el modelo de mínimos cuadrados:

$$Y_x = C \left(\frac{1}{x^n} \right)$$

Donde:

- C = número total de autores con un único documento entre número total de autores
- n = parámetro que se estima de los datos
- x = número de autores con 1, 2, 3... n documentos producidos

Para medir el grado de colaboración, se estableció la teoría del coeficiente de colaboración (Aji-feruke, Burell y Tague, 1988). El CC está basado en la idea de atribución fraccional de productividad en la generación de documentos y puede ser expresado mediante la siguiente ecuación:

$$CC = 1 - \frac{\sum_{j=1}^k \left(\frac{1}{j} \right) f_j}{N}$$

Donde:

- f_j = el número de artículos de investigación con j -autores múltiples publicados en una disciplina durante determinado periodo.
- N = número total de artículos de investigación publicados en una disciplina durante determinado periodo.
- K = el mayor número de autores por artículo en una disciplina.

Para medir el crecimiento de la literatura se empleó la función:

$$C(t) = C(0)^{at}$$

que según la teoría de Egghe y Ravichandra Rao (1992), puede ser desplegada como:

$$C(t) = cg^t \text{ donde } c > 0, g > 1, \text{ y } t \geq 0$$

También en el rubro crecimiento de la producción científica se detectó el periodo de duplicación de la literatura, según la siguiente ecuación en regresión no lineal:

$$n(\log g) = \log 2,0$$

$$n = \frac{\log 2,0}{\log g}$$

Por último, en lo tocante a los datos para determinar los trabajos más citados se consideró el número de referencias a los documentos en Google Académico. Si bien algunos registros bibliográficos fueron recuperados en bases de datos como Scopus o Web of Science, que incluían el

número de citas según el conteo de la base, se optó por el número de citas registrado en Google Académico con fines de homologar la base a partir de las citas directas en una misma fuente de información.

3. Resultados

A partir de la investigación realizada fue posible obtener un panorama general sobre el estado de las ontologías terminológicas desde una perspectiva bibliométrica. En los apartados siguientes se presentan los resultados del estudio en diversos rubros que contemplan el rango temporal de los trabajos, los idiomas de publicación, rangos de citación y otros aspectos bibliométricos.

3.1. Rango temporal de la producción científica

A partir del estudio realizado, se encontró que las ontologías terminológicas en funciones de sistemas de organización del conocimiento son un tema reciente (Gráfico 1).

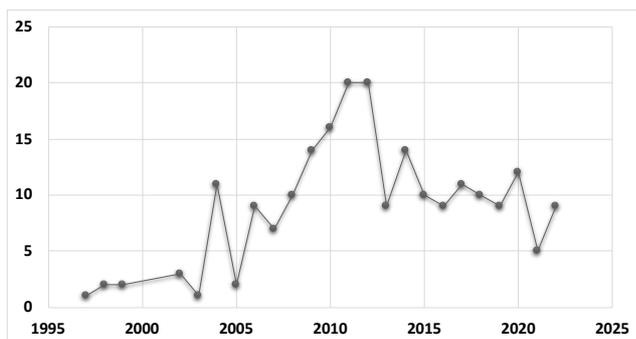


Gráfico 1. Producción científica 1997-2022

En el corpus examinado registran presencia literaria desde 1997. El primer trabajo detectado corresponde a D.M. Pisanelli, A. Gangemi y G. Steve (1997), titulado *WWW-available conceptual integration of medical terminologies: the ONIONS experience*; en tal estudio, los autores examinan la terminología médica a la luz de una arquitectura ontológica.

Las ontologías terminológicas emergieron con gran fuerza en las ciencias computacionales a mediados de la década de 1980, campo en el que fueron empleadas para modelar el mundo en sistemas computacionales, comunicar explicitaciones de conocimiento consensuadas y transferir conocimiento entre agentes mecánicos y humanos, así como entre diversos agentes mecánicos.

Su integración en la bibliotecología, como un sistema de organización del conocimiento se registró hasta una década después. Como señala Suárez Sánchez (2022, p. 39):

Las ontologías, como un sistema de organización del conocimiento, empezaron a tener presencia en la bibliotecología a partir de la década de los noventa. Su introducción se registró en el área de la organización del conocimiento y la información, campo en el que aconteció un traslape teórico-práctico entre las teorías y herramientas de las ciencias computacionales y los postulados de las ciencias documentales.

Según la línea temporal de producción científica (Gráfico 1), el estudio coincide con lo señalado por López Alonso (2000): algunos modelos conceptuales para la representación del conocimiento y la información (tesauros, encabezamientos de materia) cuentan con larga presencia en el ámbito bibliotecológico y son conocidos de antaño; sin embargo, las ontologías, como instrumentos hipermedia y asociadas a ambientes virtuales, son nuevas formas, con apenas unas cuantas décadas de inserción y estudio en la disciplina.

3.2. Crecimiento de la producción científica

Si desglosamos la línea temporal de la producción científica sobre ontologías terminológicas en quinquenios es posible identificar cinco grandes periodos:

- 1997-2001, etapa incipiente
- 2002-2006, etapa de desarrollo
- 2007-2011, etapa de masificación
- 2012-2016, etapa de auge
- 2017-2022, etapa de estabilización.

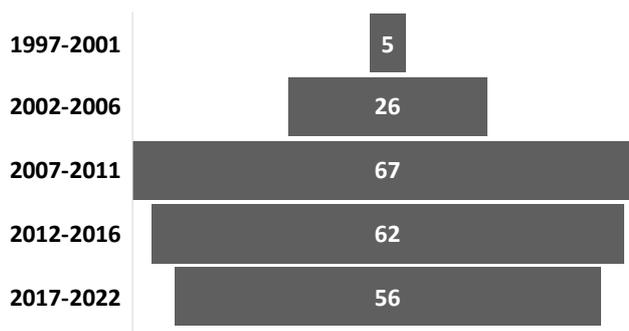


Gráfico 2. Producción científica por periodo

Observamos (Gráfico 2) que antes del año 2000 los estudios sobre ontologías terminológicas eran escasos. El crecimiento de la producción científica empieza a desarrollarse después del año 2002. Luego, en 2007 la literatura sobre el tema registra un momento de masificación que Sánchez-Cuadrado, Morato-Lara, Palacios-Madrid, Llorens-Morillo y Moreiro-González (2007)

mencionan en su artículo *De repente: ¿todos hablamos de ontologías?* En el quinquenio siguiente (2012-2016), la producción científica mantiene una etapa de auge.

Por último, después del año 2017 se advierte un ligero descenso en la producción literaria, tal hecho entraña relación con un efervescente interés por el tema en el rango 2007-2011 y la consolidación del tema en el periodo 2012-2016.

El pico más alto de crecimiento de la literatura se registra en el periodo 2007-2011 y de esos años a la actualidad el tema se ha mantenido como asunto de reflexión y análisis.

Para continuar con la tendencia de la producción científica, se analizó el crecimiento de la literatura, según los preceptos de Egghe y Revichandra Rao (1992).

Si retomamos la fórmula de crecimiento:

$$n(\log g) = \log 2,0$$

$$n = \frac{\log 2,0}{\log g}$$

Los valores con los que contamos para determinar el rango de crecimiento de la producción científica son:

$$c = 5,094$$

$$g = 1,038$$

$$R^2 = 0,230$$

Tomando en cuenta el modelo $C(t) = cg^t$, establecido en la metodología, los valores quedaron como sigue:

$$C(t) = 5,094 * 1,038^t \quad \text{donde } t = 0,038$$

Se determinó que la tasa de crecimiento de la producción científica anual es del 3,8 %. Luego, para estimar la tasa de duplicación se empleó la siguiente ecuación:

$$g * n = 2,0$$

Al reemplazar los valores, se obtuvo:

$$(1,038) n = 2$$

A partir de los logaritmos de ambos lados de la ecuación, se sintetizó que:

$$n(\log 1,038) = \log 2,0:$$

$$n = \frac{\log 2,0}{\log 1,038}$$

$$n = \frac{0,693}{0,037}$$

$$n = 18,7$$

Dado que la unidad de medida de crecimiento son años, se dedujo que la literatura científica sobre ontologías terminológicas en funciones de organización del conocimiento se duplica cada 18 años (2).

3.3. Idiomas de publicación

Respecto al idioma en el que la literatura sobre ontologías terminológicas se publica, el estudio sólo consideró tres lenguas (inglés, portugués y español). El total de 216 documentos, según idioma, se distribuyó del modo siguiente (Gráfico 3).

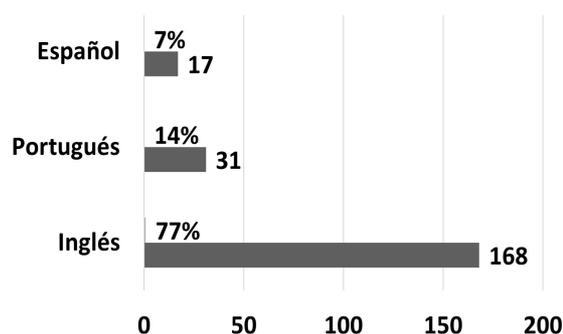


Gráfico 3. Idioma de las publicaciones

El idioma de publicación mayormente empleado es inglés con el 77 %, le siguió portugués con el 14 % y finalmente el idioma español con el 7 %. Tal hecho reafirma la idea de que (Drubin y Kellogg, 2017, p. 1)

[...] el inglés se usa actualmente como el lenguaje de la ciencia [...] esta adopción de un idioma universal o lengua franca ha tenido un efecto extraordinario en la comunicación científica, permitiendo a los investigadores de todo el mundo tener acceso a la vasta producción de artículos y haciendo posible la comunicación entre científicos en cualquier país.

Si bien un número considerable de estudios son presentados en el idioma local del investigador y de su comunidad de lectores más cercana, es recurrente que muchos expertos decidan publicar en idiomas extranjeros. Aunado a ello, en el caso de las ontologías terminológicas fue posible identificar que el idioma dominante guarda relación con el tipo de recurso dominante: actas de congresos, que suelen adherirse a un idioma establecido (inglés) con miras a la comunicación presencial entre los expertos que asisten a la convención y la publicación de las actas del encuentro. En el caso de los congresos, frecuentemente, los expertos se ven obligados a presentar sus comunicaciones orales y escritas en lengua inglesa que, consecuentemente, registra un mayor número de publicaciones.

3.4. Tipología de la producción científica

En cuanto al tipo de recursos en los que la literatura sobre ontologías terminológicas se difunde, se detectó lo siguiente (Gráfico 4):

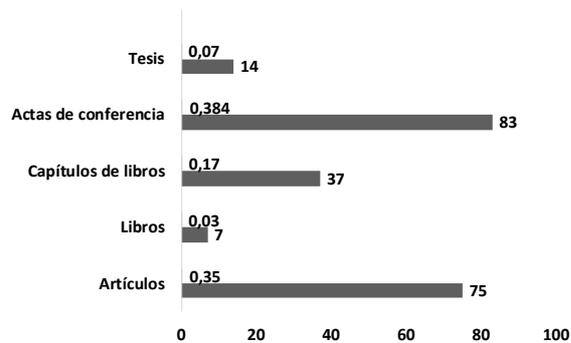


Gráfico 4. Tipo de recursos

Como puede apreciarse, las actas de conferencia, derivadas de los encuentros de la especialidad, generan la mayor producción científica sobre el tema, con 83 casos que corresponden al 38% de los registros analizados. En dicho rubro es importante destacar la relevancia de eventos internacionales como *Proceedings of the Iberoamerican Meeting of Ontological Research*, *CEUR Workshop Proceedings*, *Proceedings of The Semantic Web in Libraries Conference*, *Proceedings of The International Conference on Knowledge Engineering and Ontology Development* y *Advances in Knowledge Organization*.

Tras la fuerte tendencia de publicación sobre ontologías terminológicas en encuentros de la especialidad, las revistas son el segundo recurso con mayor literatura, con el 35% de la publicación. Los capítulos de libros son el tercer tipo de recurso con el 17%, le siguen las tesis con el 7% y los libros con el 3%.

En lo relativo a la presencia de los artículos académicos en títulos de revistas, los 75 artículos detectados se registran en 62 títulos de revistas. Siete títulos de revistas concentran el 25% de los artículos y el 75% de los artículos restantes se distribuyen en 55 títulos de revistas que sólo registran un artículo sobre el tema. Tal fenómeno confirma la idea de que, frecuentemente, un número reducido de publicaciones altamente especializadas reúne un número considerable de la producción científica sobre un tópico, mientras que el resto de las revistas son de frontera, generales o de dispersión (Urbizagástegui Alvarado, 2021).

Las revistas con mayor presencia de artículos sobre ontologías terminológicas son (Tabla 2):

Revista	Artículos
Applied ontology: an interdisciplinary journal of ontological analysis and conceptual modeling	4
Ciência da informação	3
IEEE Transactions on knowledge and data engineering	3
Onomázein: revistas de filología, lingüística y traducción	3
Perspectivas em ciência da informação	2
Revista signos	2
Scire: representación y organización del conocimiento	2

Tabla II. Revistas con mayor producción científica

Entre las publicaciones periódicas destaca la revista *Applied ontology: an interdisciplinary journal of ontological analysis and conceptual modeling* que es la principal fuente de diseminación de investigaciones. Tal publicación, comprende dos grandes rubros: el análisis ontológico y el modelado conceptual. Según el análisis ontológico, integra trabajos sobre epistemología, modelado de dominios y lógica para la estructuración de redes semánticas de información. Según el modelado conceptual, incluye estudios sobre las estructuras cognitivas, categoriales y lingüísticas para modelar el mundo en agente computacionales, así como la terminología que es la base de diversos sistemas de organización del conocimiento (tesauros, taxonomías, ontologías, entre otros).

Entre las revistas con mayor producción sobre ontologías terminológicas destacan dos revistas brasileñas que publican en portugués sobre el tema: *Ciência da informação* y *Perspectivas em ciência da informação*. Aunado a ello, el trabajo más citado en el rubro también está escrito en portugués. A partir de ello, es posible señalar una contribución significativa de la producción científica en lengua portuguesa en la investigación sobre ontologías.

3.5. Rango de citación

Los trabajos más citados fueron otro elemento de análisis en la producción científica del tema. Se encontró que los diez trabajos más citados con sus respectivos autores son los siguientes (Tabla III, en la página siguiente).

El trabajo más citado *Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção* (Almeida y Bax, 2003) es una revisión general sobre las ontologías que sólo abunda sobre las ontologías terminológicas como uno de sus tipos. Al

respecto señala que la relevancia de las ontologías terminológicas en el ámbito documental reside en su alcance como sistemas para la organización temática de dominios de conocimiento, la indización de recursos de información en bibliotecas digitales y su utilidad como plantillas conceptuales para la recuperación de la información en la web.

Autores	Título	Citas
Almeida, M. B. y M. P. Bax	Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção	352
Borge-Holthoefer, J. y A. Arenas	Semantic networks: structure and dynamics	226
Wang W., P.M. Barnaghi y A. Bargiela	Probabilistic topic models for learning terminological ontologies	142
García Jiménez, A.	Instrumentos de representación del conocimiento: tesauros versus ontologías	123
Nielsen, L.P.	Knowledge dissemination based on terminological ontologies	82
Rani, M., A.K. Dhar y O.P. Vyas	Semi-automatic terminology ontology learning based on topic modeling.	60
Lacasta, J., J. Noguera-Iso y F.J. Zarazaga-Soria	Terminological ontologies: design, management and practical applications	38
Madsen B.N., H.E. Thomsen y C. Vikner	Principles of a system for terminological concept modelling	36
Jiménez-Briones, R. y A. Luzondo Oyon	Building ontological meaning in a lexico-conceptual knowledge base	35
Soler Monreal, C. e I. Gil Leiva	Posibilidades y límites de los tesauros frente a otros sistemas de organización del conocimiento: folksonomías, taxonomías y ontologías	34

Tabla III. Trabajos con mayor número de citas

3.6. Productividad de los autores

En lo tocante a la productividad de los autores, en la investigación se siguieron los preceptos de la Ley de Lotka. Se hizo una comparación de la productividad de los autores considerando un rango estimado y un rango observado. También se realizó un conteo completo de todos los autores y se obtuvo un valor para $c = 0,838$, $n = 3,07$ con un error estándar de 0,081 y con un R^2 de 0.95, es decir, una confiabilidad de 95%.

En tal rubro se obtuvieron los siguientes resultados:

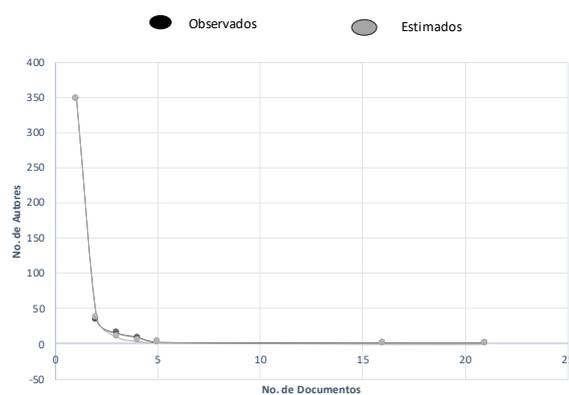


Gráfico 5. Productividad estimada vs observada

La productividad estimada se mantuvo en un rango bastante cercano con la productividad observada (Gráfico 5) mientras que los 216 documentos publicados sobre ontologías terminológicas fueron producidos por 420 autores.

En cuanto al número de artículos que los expertos publicaron, la relación número de artículos/autores se distribuyó del modo siguiente (Tabla IV):

Artículos	Autores esperados	Autores observados
1	354	354
2	42	38
3	12	15
4	5	9
5	2	2
16	0	1
21	0	1

Tabla IV. Número de artículos por autor

Según los datos, 354 autores (84 %) publicaron solo un trabajo sobre ontologías terminológicas; 64 autores (15,5 %) registran entre 2 y 5 estudios y sólo 2 autores (0,5 %) cuentan con más de cinco investigaciones sobre el tema. Tales resultados coinciden con la teoría de Ramsden (1994) quien afirma que la producción científica de los investigadores en una línea de trabajo suele ser baja, debido a la naturaleza de la investigación científica, los intereses mutables de estudio y las exigencias de tratamiento de temas o líneas de investigación a los que, frecuentemente, los académicos están sujetos en sus instituciones de afiliación.

Los autores con mayor producción científica sobre ontologías terminológicas son (Gráfico 6):

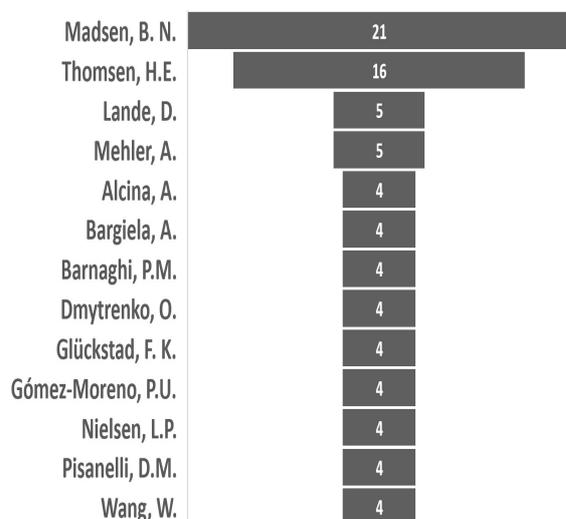


Gráfico 6. Número de trabajos por autor

La autora más productiva es Bodil Nistrup Madsen, reconocida investigadora de la Escuela de Negocios de Copenhagen. Sus líneas de estudio engloban la terminología, los sistemas de gestión terminológica, las bases de datos, la terminología para el modelado de datos, las ontologías y la representación y organización del conocimiento. En vida, Bodil Nistrup Madsen (1948-2022) recibió el reconocimiento de profesora emérita en su centro de trabajo y obtuvo dos renombrados premios internacionales: en 2002 recibió el Premio de la Fundación Hedorf a la Investigación en Lenguaje para Aplicaciones Específicas, y en 2006 recibió el premio Eugen Wüster en reconocimiento a sus investigaciones y proyectos sobre terminología aplicada.

La segunda autora con mayor productividad sobre ontologías terminológicas es Hanne Erdman Thomsen. Sus líneas de investigación son ciencias computacionales y lingüística aplicada. Ha laborado en la Universidad de Copenhagen como asistente de investigación y en la Escuela de Negocios de Copenhagen en el departamento de lingüística computacional. Actualmente se encuentra adscrita al proyecto Danish Taxation en tareas de arquitectura de la información.

Si consideramos la teoría de producción científica establecida por Braun, Glänzel y Schubert (2001) y adaptada por Walter, Lanza, Sato, Silva y Bach (2010) la relación autores/número de trabajos puede ser agrupada en categorías:

- *Los de única vez*: autores que en un periodo temporal establecido sólo han publicado un documento.

- *Retirantes*: autores con al menos dos documentos publicados, pero ninguno en los últimos tres años.
- *Entrantes*: autores con al menos dos documentos publicados en últimos tres años del periodo establecido.
- *Transeúntes*: autores con al menos dos documentos publicados, pero al menos un documento publicado en los últimos tres años del periodo de estudio.
- *Continuantes*: autores con al menos cinco documentos publicados y al menos uno de ellos publicado en los últimos tres años del lapso de estudio.

Con base en tal categorización, los autores que publican sobre ontologías terminológicas responden a la siguiente tipología (Tabla V):

Categoría	Autores	Porcentaje
Los de única vez	354	84 %
Retirantes	47	11 %
Entrantes	11	3 %
Transeúntes	6	1,5 %
Continuantes	2	0,5 %

Tabla V. Categorización de autores

El 84 % de los autores sólo ha publicado un documento sobre el tema, mientras que hay un 11 % de autores retirantes y un 3 % de entrantes. Tal hecho implica que el tema carece de un número considerable de autores cautivos, cuya línea de investigación se centre en ontologías terminológicas.

3.7. Colaboración

En la producción científica sobre ontologías terminológicas también se analizó la colaboración entre autores, según el coeficiente de colaboración propuesto por Ajiferuke, Burell y Tague (1988). Los datos evidenciaron trabajos producidos por 1 y hasta 13 autores (Tabla 6, en la página siguiente).

La colaboración entre académicos de una misma disciplina o de disciplinas diversas posibilita alcanzar un fin común y en los últimos años se ha convertido en una práctica que eficiente los recursos asignados a los investigadores y las instituciones. En adhesión, la colaboración es una práctica que enriquece a las comunidades científicas a través de la comunicación y el enriquecimiento entre pares (González Alcaide y Gómez Ferri, 2014, p. 4):

Ciencia y colaboración se han constituido en un binomio indisoluble para posibilitar el progreso y el avance del conocimiento [...] la colaboración constituye uno de los rasgos definitorios de la evolución de la Ciencia a lo largo de las últimas décadas.

Quinquenio	Número de publicaciones con 1, 2, 3, n autores											CC
	1	2	3	4	5	6	7	8	9	12	13	
1997-2001	0	1	2	2	0	0	0	0	0	0	0	0.67
2002-2006	7	5	10	3	0	1	0	0	0	0	0	0.47
2007-2011	24	18	16	4	4	0	0	0	1	0	0	0.40
2012-2016	17	20	9	6	5	2	1	0	0	1	1	0.47
2017-2022*	14	15	11	10	4	2	0	0	0	0	0	0.49
Total	62	59	48	25	13	5	1	0	1	1	1	0.46

* Dado que la investigación abarca hasta 2022, el último grupo se formó de seis años.

Tabla 6. Colaboración entre autores

En el corpus analizado, se detectó la siguiente tendencia de coeficiente de colaboración por quinquenios (Gráfico 8).

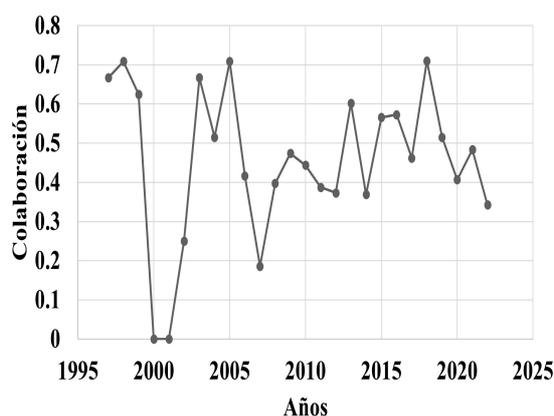


Gráfico 7. Colaboración entre autores por años

El estado de la colaboración en la literatura sobre ontologías terminológicas ha pasado por varios momentos. Al principio (1997-2005) el índice de colaboración se mantuvo con tendencia alta; sin embargo, en los últimos años (2005-2021) ha decrecido ligeramente.

Tal hecho se debe a que, cuando las ontologías se introdujeron en la bibliotecología y otras ciencias de la información, se les observó como sistemas altamente complejos que para su estudio e implementación requerían trabajo multidisciplinario y conocimiento experto de lingüistas, documentalistas e informáticos. Luego, en las últimas décadas el perfil de los bibliotecólogos y otros profesionales de la información se ha enriquecido. Hoy en día tales expertos cuentan con mayor formación sobre terminología, especialmente enfocada a la construcción de lenguajes documentales y experiencia en el manejo de bases de

datos relacionales y programas para la codificación de sistemas de organización del conocimiento, lo que reduce el número de autores en proyectos vinculados con ontologías terminológicas.

Pese a lo señalado anteriormente, la colaboración puede asumirse como una característica esencial de la producción científica sobre ontologías terminológicas. Al respecto es importante retomar a Almeida y Costa (2021) quienes señalan que los retos que plantea el desarrollo de los recursos terminológicos para la web semántica han propiciado el estrechamiento de vínculos entre la terminología, la ingeniería y las ciencias documentales, en una búsqueda por insertar los vocabularios de diversos campos, desde las ciencias médicas hasta el patrimonio cultural, en el marco de la web semántica y una recuperación más eficaz en la búsqueda temática o por palabras clave.

4. Conclusiones

Las tecnologías de la web semántica han ganado gran atención en los últimos años y para apoyarlas han surgido recientes herramientas y técnicas. La utilización de ontologías es una de las principales características de las bibliotecas digitales semánticas y, en particular, las ontologías terminológicas son un soporte fundamental como instrumentos para definir la terminología que represente temáticamente recursos y campos de conocimiento mediante conceptos, clases, atributos y relaciones.

Considerando lo anterior, la producción científica sobre organización de la información y el conocimiento desde principios del siglo veinte ha registrado una creciente producción científica sobre ontologías terminológicas. De esos años a la fecha la literatura sobre el tema se ha mantenido como un tema de interés: las ontologías han mostrado ser una herramienta útil para la representación y organización terminológica-temática de dominios de conocimiento, eliminando los problemas de ambigüedad, proporcionando interrelación y semántica a la terminología del área.

En términos de investigación, se han revisado los fundamentos teóricos de las ontologías terminológicas; se han construido modelos terminológicos que cubren diversas áreas de conocimiento (medicina, humanidades, arte, etc.) y muchas ontologías han sido integradas en sistemas de recuperación. Pese a ello, aún persiste un cuestionamiento esencial sobre: ¿cuál será el futuro de las ontologías terminológicas en funciones de organización del conocimiento?

A partir de la investigación realizada es posible concluir que las ontologías son un tema de estudio con tendencia de crecimiento estable. Actualmente, el tópico mantiene el interés de un grupo de autores realmente comprometidos con el estudio del tema, quienes han descubierto que el uso de modelos axiomáticos proporciona una semántica adicional y una explicitación formal de relaciones de gran utilidad en la representación, organización y recuperación de la información.

En la producción literaria sobre el tema, predomina la publicación en lengua inglesa y las actas de conferencia como recurso dominante. Basado en ello, las tendencias futuras nos permiten proponer una ampliación de las ontologías terminológicas en contextos locales, en publicaciones en lengua española, revistas locales y encuentros regionales donde los expertos, que asisten a convenciones internacionales, puedan tener contacto no sólo con investigadores sino con bibliotecarios, que día a día laboran en unidades de información en las cuales las ontologías terminológicas puedan tener implementaciones propositivas.

Un segundo rasgo de la literatura sobre ontologías terminológicas es la colaboración, dicha característica se observa como elemento fundamental para ampliar al panorama de aplicación de las ontologías terminológicas en funciones de indexación de recursos, organización de dominios de conocimiento y recuperación de información. Los resultados del artículo coinciden con el trabajo previo de Zhu, Kong, Hong, Li y He (2015) en el que se argumentó sobre la colaboración y el trabajo interdisciplinario como aspecto clave en la investigación sobre ontologías.

Los resultados obtenidos en la presente investigación coinciden con lo señalado por García Marco (2008) cuando, hacia 2007, apuntaba una etapa de despegue de las ontologías. En adhesión también es posible mencionar que, de esos años a la actualidad, la producción científica sobre ontologías ha entrado en una etapa de estabilización que ya no registra las cifras efervescentes de un primer interés sobre el tema.

Finalmente, es posible afirmar que las ontologías han trastocado el mundo de la organización del conocimiento y son un tema interdisciplinar de gran interés para documentalistas, lingüistas, filósofos y expertos en ciencias computacionales. Como señalan Netto y Lima (2017) las ontologías aún tienen muchas aristas de estudio pendientes. Por tanto, se espera que su tendencia de investigación se mantenga constante tanto en la determinación de sus aspectos teóricos como en la definición de sus implementaciones en sistemas de información.

Notas

- (1) En Scopus y Web of Science fue posible llevar a cabo la búsqueda de los descriptores en los campos de título, temas y resumen. En Redalyc, Dialnet y WordCat fue posible buscar en los campos de título y palabras clave. Mientras que en Google Académico sólo se llevó a cabo la búsqueda en palabras clave.
- (2) La estimación del periodo de duplicación de la literatura es precisa en términos estadísticos; sin embargo, se observa algo incierta, debido a que los años 2021 y 2022 registraron baja producción sobre el tema (5 y 6 documentos, respectivamente) e incidieron, de manera negativa, en la tendencia de crecimiento que registraba 12 documentos anuales en promedio. Tal hecho, unido a que esos años, se caracterizaron por una pandemia mundial derivada del COVID-19 y la suspensión de encuentros de la especialidad, que son la mayor fuente de producción literaria, son un aspecto negativo que, probablemente, no esté ofreciendo datos completamente confiables. De este modo, se asume que: es probable que el tiempo de duplicación de la muestra sea algo menor al estimado en el artículo.

Referencias

- Ajiferuke, Isola; Burell, Q.; Tague, Jean (1988). Collaborative coefficient: a single measure of the degree of collaboration in research. // *Scientometrics*. 14:5/6, 421-433.
- Almeida, Bruno; Costa, Rute (2021). OntoAndalus: an ontology of Islamic artefacts for terminological purposes. // *Semantic Web*. 12:2, 295-311.
- Almeida, Mauricio; Bax, Marcello (2003). Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. // *Ciência da informação* 32, 7-20.
- Aparicio, Gloria; Txomin Iturralde; Maseda, Amaia (2019). Conceptual structure and perspectives on entrepreneurship education research: a bibliometric review. // *European research on management and business economics*. 25:3,105-113.
- Braun, Tibor; Glänzel, Wolfgang; Schubert, András (2001). Publication and cooperation patterns of the authors of neuroscience journals. // *Scientometrics*. 51:3, 499-510.
- Drubin, David; Kellogg, Douglas (2012). English as the universal language of science: opportunities and challenges. // *Molecular biology of the cell*. 23:8, 1399.
- Egghe, Leo; Ravichandra Rao, I.K (1992). Classification of growth models based on growth rates and its applications. // *Scientometrics*. 25:1, 5-46.
- García Marco, Francisco Javier (2008). El nacimiento y el despegue de la investigación moderna sobre ontologías. // *Anuario ThinkEPI*. 2:122-125.
- González Alcaide, Gregorio; Gómez Ferri, Javier (2014). La colaboración científica: principales líneas de investigación y retos de futuro. // *Revista Española de Documentación Científica*. 37:4, 1-15.
- Gruber, Thomas (1995). Toward principles for the design of ontologies used for knowledge sharing? // *International Journal of Human and Computer Studies*. 43:5-6, 907-928.
- Guarino, Nicola (1995). Formal ontology, conceptual analysis and knowledge representation. // *International Journal of Human-Computer Studies*. 43:5-6, 625-640.
- Hodge, Gail (2000). Systems of knowledge organization for digital libraries: beyond traditional authority files. Washington: The Digital Library Federation Council on Library and Information Resources.

- Lacasta, Javier; Nogueras-Iso, Javier; Zarazaga-Soria, Francisco Javier (2010). Terminological ontologies: design, management and practical applications. Nueva York: Springer.
- López Alonso, Miguel Ángel. (2000). Las estructuras conceptuales de representación del conocimiento en Internet. // Scire: representación y organización del conocimiento. 6:1, 107-123.
- Maletta, Héctor (2009). Metodología y técnica de la producción científica. Lima: CEPES-Universidad del Pacífico.
- Netto, Cristiane; Lima, Gercina (2017). Visualização de ontologias: estudos e perspectivas. // Informação & Sociedade. 27:3, 59-72.
- Pedraza, Rafael; Codina Lluís; Rovira Cristòfol (2007). Web semántica y ontologías en el procesamiento de la información documental. // El profesional de la información. 16:6, 569-78.
- Piedra Salomón, Yelina; Martínez Rodríguez, Ailín (2007). Producción científica. // INFO: ciencias de la información. 38:3, 33-38.
- Pisanelli, Domenico; Gangemi, Aldo; Steve, Geri (1997). WWW-available conceptual integration of medical terminologies: the ONIONS experience. // Masys, Daniel (ed.) Proceedings of the AMIA Annual Fall Symposium. Nashville, Tennessee: American Medical Informatics Association. 575-579.
- Ramsden, Paul (1994). Describing and explaining research productivity. // Higher education. 28:2, 207-226.
- Reitz, Joan (ed.) (2004). Dictionary for library and information science. Westport: Libraries Unlimited.
- Sánchez-Cuadrado, Sonia; Morato-Lara, Jorge; Palacios-Madrid, Vicente; Llorens-Morillo, Juan; Moreira-González, José-Antonio (2007). De repente, ¿todos hablamos de ontologías? // El profesional de la información. 16:6, 562-568.
- Suárez Sánchez, Adriana (2022). Ontologías: principios y aplicaciones en la bibliotecología. Ciudad de México: Instituto de Investigaciones Bibliotecológicas y de la Información, Universidad Nacional Autónoma de México.
- Spinak, Ernesto (1996). Diccionario enciclopédico de bibliometría, cienciometría e informetría. Caracas, Venezuela: UNESCO.
- Urbizagástegui Alvarado, Rubén (1999). La ley de Lotka y la literatura de bibliometría. // Investigación Bibliotecológica. 13:27,125-141.
- Urbizagástegui Alvarado, Rubén (2021). Análisis bibliométrico de la literatura sobre fósiles en el Perú, 1840-2020. // Revista del Instituto de Investigación de la Facultad de minas, metalurgia y ciencias geográficas. 24:48, 153-165.
- Van Heijst, G.; Schreiber, A.; Wielinga, B. (1997). Using explicit ontologies in KBS development. // International Journal of Human-Computer Studies. 45, 183-292.
- Walter, Silvana Anita; Lanza, Beatriz Barreto Brasileiro; Sato, Kawana Harue, Silva, Eduardo Damião; Bach, Tatiana Marceda (2010). Análise da produção científica de 1997 a 2009 na área de estratégia: produção e continuidade de atores e cooperação entre instituições brasileiras e estrangeiras. // Anais do Encontro da Associação Nacional dos Programas de Pós-graduação em Administração.34. São Paulo: ANPAD.
- Zhu, Qiaoli; Kong, Xuesong; Hong, Song; Li, Junli; He, Zongyi (2015). Global ontology research progress: a bibliometric analysis. // Aslib Journal of Information Management. 67:1, 27-54.

Enviado: 2023-03-02. Segunda versión: 2023-06-30.
Aceptado: 2023-11-02.
