

CAPÍTULO 2

EL CAMPO DE LA CURACIÓN DIGITAL CON PERSPECTIVA DE GÉNERO*

MARÍA-ANTONIA OVALLE-PERANDONES
Universidad Complutense de Madrid

MIRELYS PUERTA-DÍAZ
Universidade Estadual Paulista

DANIEL MARTÍNEZ-ÁVILA
Universidad de León

RESUMEN

En este capítulo se analiza el campo de la curación digital con una perspectiva de género. Se realiza un análisis del dominio hasta 2020 haciendo uso, principalmente, de técnicas bibliométricas. Se trabaja con la principal fuente de referencia en el área según consenso, *International Journal of Digital Curation* (Miguel *et al.*, 2013), y bases de datos internacionales para identificar otras publicaciones que citan esta importante fuente. De estas fuentes se obtuvo una relación exhaustiva de todos los datos de autoría para después identificar el género (masculino o femenino) y trabajar con el concepto de liderazgo a partir del autor de correspondencia.

Palabras clave: Curación digital, análisis de dominio, liderazgo femenino.

ABSTRACT

In this text we analyze the field of digital curation from a gender perspective. We conducted an analysis of the domain until 2020 using mainly bibliometric techniques. We worked with the main reference source in the area, according to the scientific

* Este capítulo se incluye dentro del Proyecto AI-GenBias financiado por la Agencia Estatal de Investigación (PDI2019-106695RB-100/AEI/10.13039/501100011033).

consensus, that is the International Journal of Digital Curation (Miguel *et al.*, 2013), as well as international databases to identify other publications that cited this important source. We obtained from these sources an exhaustive list of all the authorship data to identify the gender of the authors (male or female) and the concept of leadership based on the corresponding author.

Keywords: Digital curation, domain analysis; female leadership.

INTRODUCCIÓN A LA CURACIÓN DIGITAL

LA CURACIÓN DIGITAL, o curación de contenidos, es un campo profesional emergente, surgido en el contexto de la explosión de contenidos web y la Web 2.0. El *Digital Curation Centre* (DCC) define el concepto como aquel que tiene la cualidad de aglutinar las diferentes tareas para la gestión, clasificación y organización de los activos digitales a lo largo de toda su vida útil, desde el momento de su conceptualización, pasando por su uso como elementos activos, y cuidando su conservación para que puedan ser presentados y utilizados a largo plazo desde su lugar de preservación (Rusbridge *et al.*, 2005). En el manifiesto profesional de la curación de contenidos (Bhargava, 2009), se define el objeto como la búsqueda, agrupación, organización y compartición de los contenidos más relevantes sobre un aspecto específico.

Como campo profesional ligado cada vez más a la investigación, debido a la propia naturaleza del uso de los ambientes electrónicos para la difusión de la investigación, el DCC brinda asesoramiento de expertos y ayuda práctica sobre cómo almacenar, administrar, proteger y compartir datos de investigación digital; ofrece una amplia gama de recursos que incluyen herramientas, orientación y formación en línea; y brinda servicios de consultoría en temas como el desarrollo de políticas y la planificación de la gestión de datos. Como dominio de conocimiento, la curación digital ha sido disputada por las áreas del marketing digital, el periodismo y la comunicación, la ingeniería y la informática, la información y la documentación o la educación entre otros (Guallar *et al.*, 2020). En cualquier caso, al estar el objeto de estudio ligado a las tecnologías, existe un peligro de exclusión de las mujeres en su esfera científica, tradicionalmente arraigada en las posiciones racionalistas y positivistas a las que han tenido acceso las clases dominantes.

En este sentido, el objetivo de este capítulo es analizar el liderazgo científico de las mujeres en el campo emergente de la curación digital. Para ello, se toma como fuente principal la revista científica electrónica *International Journal of Digital Curation* (IJDC), dedicada íntegramente a publicar trabajos, artículos y noticias sobre la conservación de objetos digitales y otros temas afines. Aunque existen trabajos que han realizado estudios bibliométricos de la producción académica sobre curación digital (por ejemplo, Guallar *et al.*, 2020), ningún estudio anterior ha sido realizado con perspectiva de género.

En el día internacional de la mujer del 8 de marzo de 2021 se puso de relevancia el valor del liderazgo femenino. En el contexto de cualquier campo o dominio, el liderazgo científico mide la cantidad de producción de una determinada unidad de medida como contribuyente principal, es decir, la parte de la producción en la que el autor de correspondencia lo es en esa unidad de

medida (Moya-Anegón, 2012). El presente trabajo busca mostrar si ese valor trasciende en un campo tan emergente como el de la curación digital o si, al contrario, se mantienen las barreras que asociamos a una cierta tradición, bien sean estructurales o culturales (Meiksins *et al.*, 2019), impidiendo aproximarse a una cierta equidad científica entre ambos géneros (Palomba, 2006).

CUESTIONES METODOLÓGICAS

Metodológicamente el trabajo hace uso de técnicas bibliométricas para analizar el dominio de la curación digital, con especial énfasis en el rol, contribuciones y características de las mujeres dentro de la comunidad epistémica que conforma el dominio analizado. El análisis de dominio, como paradigma teórico-epistemológico, fue introducido en la ciencia de la información en 1995 (Hjørland & Albrechtsen, 1995). Posteriormente, Hjørland (2002; 2017) listó la bibliometría como una de las formas de aproximarse al análisis de dominio y ha sido trabajada de forma satisfactoria en numerosos estudios (Smiraglia, 2015)

El presente trabajo utiliza la bibliometría para analizar el dominio de la “curación digital” con base en la producción científica en el periodo 2007-2020. Como fuente de investigación primaria trabaja con la revista científica especializada en el campo de la curación digital *International Journal of Digital Curation*, consultada disponible en la dirección <http://www.ijdc.net/>. En el contexto de la ciencia de la información y el análisis de dominio son numerosos los estudios que han analizado un dominio a partir de la producción en una revista especializada (Smiraglia, 2012; Miguel *et al.*, 2013; Guimarães *et al.*, 2015; Oliveira *et al.*, 2017; Alves *et al.*, 2019; Martínez-Ávila *et al.*, 2020). La relación entre las revistas y los dominios está explicada por Smiraglia (2015, p. 9) en los siguientes términos:

Journals are the formal venues for most scholarly communication, and studying them as whole works is also one means of identifying productive elements of a research front. Of course, few journals are devoted to topical areas that are as narrowly defined as most domains under study. For example, even in the field of knowledge organization, the principle journal *Knowledge Organization* is devoted to the entire field. Thus, it would likely be the most cited journal in all domains within KO, but there are no journals devoted to specific narrow aspects of KO, such as “integrative levels,” “multilingual thesauri,” or “ethics in KO.

La segunda fuente principal de la investigación fue Google Citas, utilizada para identificar las publicaciones que citan los trabajos publicados en esta revista especializada. La combinación de ambas fuentes proporcionó la relación exhaustiva de todos los datos de autoría y la identificación del género asociado a los nombres de los autores (masculino o femenino) para analizar el liderazgo científico de las mujeres según el rol de autor de correspondencia (Moya-Anegón *et al.*, 2013).

Para la obtención de los registros se utilizó el software gratuito Publish or Perish (PoP) v.7, el cual permite extraer directamente las publicaciones de Google Académico, donde la revista se encuentra indexada. La estrategia de búsqueda utilizada consistió en colocar entre comillas el título de cada publicación en el campo title words de PoP. Posteriormente, con un gestor bibliográfico, se limpiaron y completaron algunas omisiones de los datos con ayuda de la consulta de la fuente primaria disponible en el sitio web oficial de la revista. La identificación del género de las autoras se realizó mediante una consulta a la plataforma Gender API (<https://gender-api.com>), cuya base de datos contiene 6.084.389 nombres validados de 191 países diferentes. Los datos de género fueron exportados en formato .csv para su posterior procesamiento e integración en la red de coautoría.

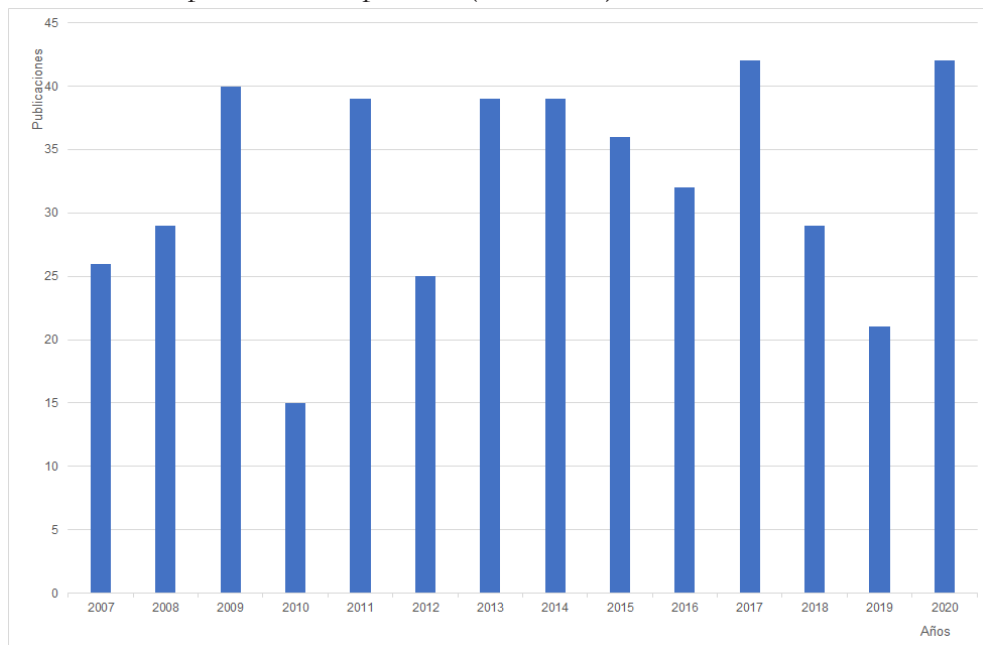
CARACTERÍSTICAS DEL CAMPO DE LA CURACIÓN DIGITAL

EVOLUCIÓN TEMPORAL

En el periodo comprendido entre los años 2007 a 2020 y en la revista IJDC se han publicado un total de 454 trabajos. Eso hace que en promedio se publican anualmente 32 trabajos. Como se aprecia en la figura 1, la tendencia es la publicación de un número similar de trabajos por año, si bien hay algunos años como el 2010 o 2019 donde el descenso se justifica ante la publicación de un único número en el año. Si bien, algunos años aun habiendo publicado un único número, en él se han concentrado un elevado número de publicaciones, como ocurre en el año 2020.

Figura 1.

Frecuencia de publicaciones por año (2007-2020)



Fuente: elaborada por los autores con Microsoft Excel.

TÉRMINOS EN LOS TÍTULOS

La figura 2 incluye con forma de nube de términos de los títulos tomando los términos utilizados en los títulos de todos los trabajos publicados en IJDC, sin considerar las palabras vacías y los términos que, aunque más frecuentes, son los que se esperaban encontrar considerando la temática de la revista. Los más frecuentes, conservando su idioma original, son data (273), digital (107), research (98), curation (84), management (73) y preservation (70).

Los autores analizados confirman que se puede definir la cobertura temática de la revista tanto por los términos frecuentes que no fueron incluidos en la nube de palabras, como por science, information, scientific, challenges, university, metadata o education. La mencionada figura 2 incluye un total de 1257 términos de los títulos. El tamaño de la palabra atiende a la frecuencia y se representa con un color diferente para mejorar la comprensión de la nube.

Figura 2.

Nube de términos incluidos en los títulos de los artículos publicados en IJDC.



Fuente: elaborada por los autores con Generador de Nubes de Palabras (<https://www.nubedepalabras.es/>)

PRODUCCIÓN DE LOS INVESTIGADORES

La élite de investigación se calculó con base a la ley de Price, considerando que “el número de productores prolíficos equivale a la raíz cuadrada del número total de autores” Price (1976, p. 3, traducción propia). La fórmula de la $\sqrt{1044}$ investigadores corresponden con aproximadamente los 32 autores más productivos. En la tabla 1 se listan los 21 autores que han publicado más de 5 publicaciones. Se utiliza el color rojo para las investigadoras; en azul para los investigadores.

Nombre del autor	Total de Publicaciones
Jones, Sarah	11
Ball, Alexander; Lyon, Liz; Matthews, Brian	9
Ludäscher, Bertram; Whyte, Angus; Donnelly, Martin;- Treloar, Andrew; Callaghan, Sarah	8
Brown, Geoffrey	7
Snow, Kellie; Pryor, Graham; Abrams, Stephen; Carl- son, Jake; Molloy, Laura	6
Day, Michael; Tedds, Jonathan; Mayernik, Matthew S.; Willoughby, Cerys; Knight, Gareth; Missier, Paolo	5

Tabla 1. Relación de autores que componen la Élite de investigación

Fuente: elaborada por los autores con Microsoft Excel.

Entre los autores más productivos en este campo emergente destaca la investigadora inglesa Sarah Jones del Digital Curation Centre, con un total de 11 publicaciones. En sus estudios se abordan las principales preocupaciones en cuanto a la gestión de datos, particularmente se especializa en investigaciones sobre la implementación de políticas y planes de gestión de datos en contextos institucionales.

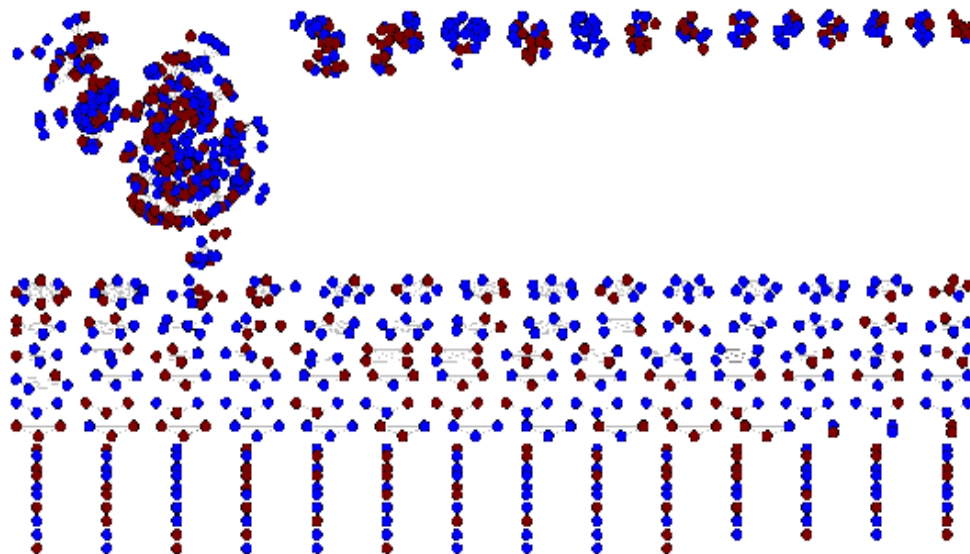
RED DE COAUTORÍA

Haciendo uso del Análisis de Redes Sociales, la red que considera las relaciones de coautoría está formada por 1044 nodos. La matriz de las relaciones es simétrica y ponderada. Simétrica, ya que un autor no puede ser coautor con otros sin que los otros sean coautores de ese autor. Y ponderada, atendiendo al número de ocasiones en el que ocurren las coautorías entre los autores. Las relaciones puntuales, o con un único documento, crean en esta red 2483 relaciones; en dos o más documentos, 182. Esos autores establecen un total de 2665 relaciones.

En la red de la figura 3 se visualizan 628 autores masculinos (color rojo) y 416 autoras (color azul). Eso permite contar con la perspectiva de género respecto de los autores que forman esa red.

Figura 3.

Red de coautoría de la revista IJDC para el periodo 2007-2020

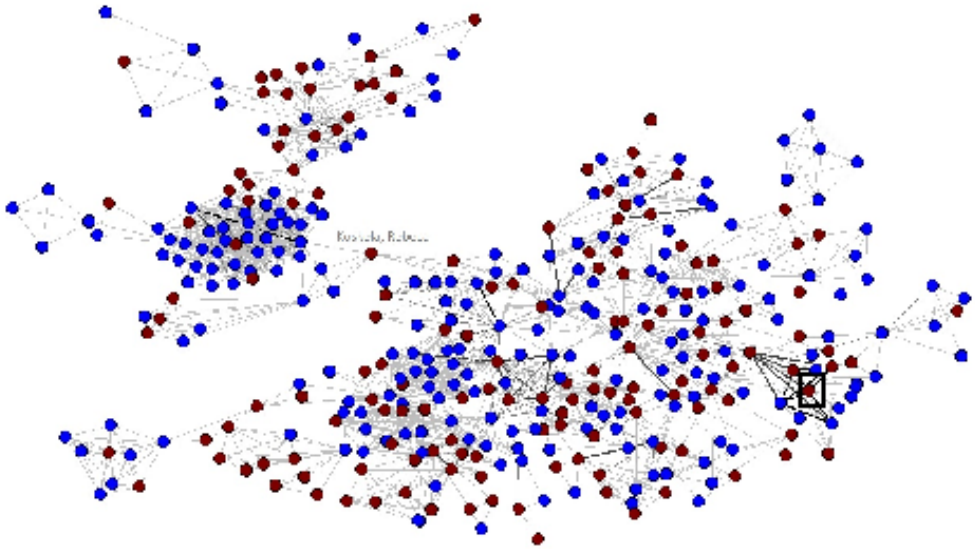


Fuente: elaborada por los autores con Pajek.

En esa misma figura y en la parte inferior, se visualizan aquellos nodos aislados, es decir, son autores pero no mantienen ninguna relación de coautoría con otros. Eso ocurrió con 52 autores. En la parte superior izquierda queda claramente identificado el componente principal. De entre los 192 componentes o subredes de toda la red de coautoría del IJDC, en ese componente quedan conectados 380 nodos.

Figura 4.

Componente principal de la red de coautoría de IJDC en 2007-2020



Fuente: elaborada por los autores con Pajek.

Como se ha indicado, la subred incluida en la figura 4 está formada por 380 nodos (226 vinculados al género masculino; 154 al género femenino) y en ella queda claramente identificado un vínculo débil en color rojo (que corresponde con la autora Koskela, Rebeca, ya que, si ella desapareciera, la red quedaría a su vez más fragmentada). Rebeca Koskela cumple la función de conectar la agrupación superior izquierda, de menor tamaño, con la agrupación derecha, de mayor tamaño. Esa autora conecta con el autor Bertram Ludascher que es el autor con mayor grado de la red (49). De igual manera que ocurría en toda la red, las relaciones de coautoría más frecuentes en la red ocurren en una ocasión en 1560 de ellas (gris claro); mientras que en más de una ocasión ocurre en 140 relaciones (gris oscuro). Así la red de coautoría que genera la autora Kirsty Merret, identificada con un rectángulo de color negro y en la parte inferior derecha de la figura, es la que más veces ocurre.

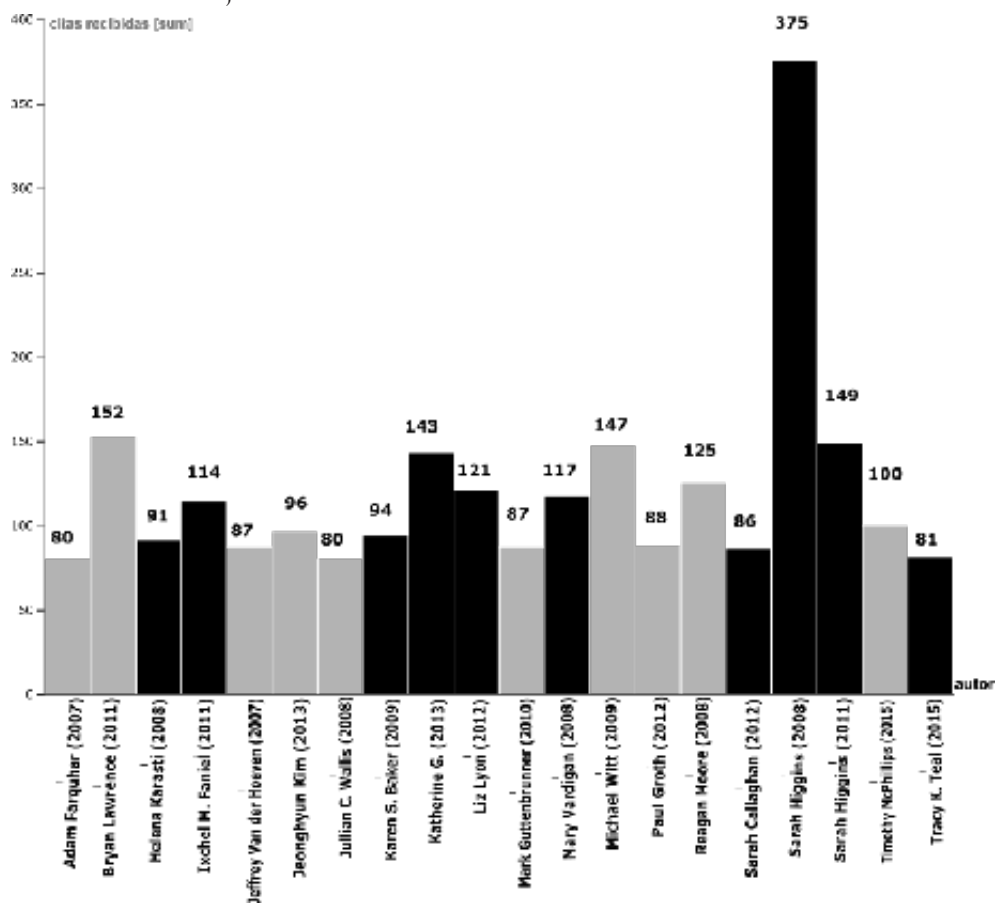
DINÁMICAS DE CITACIÓN

En la figura 5 se muestran los trabajos que han recibido más de 80 citas, identificados en el eje X por la referencia. Si analizamos el impacto de las

publicaciones, entre los 20 trabajos más citados, se encuentran 6 artículos que recibieron más de 100 citaciones y cuya composición de autoría está liderada por el género femenino. En este indicador la autora Sara Higgins obtiene el primer y tercer mayor número de citaciones recibidas a sus trabajos. Las discusiones sobre la emergencia de la Curación Digital como disciplina emergente y su propuesta de modelo de ciclo de vida constituyen pilares teóricos de este campo científico.

Figura 5.

Relación de trabajos con más de 80 citas recibidas



Fuente: elaborada por los autores con RawGraph y Inkscape.

CIENCIA Y LIDERAZGO

El análisis del liderazgo científico de las mujeres en el campo emergente de la curación digital como contribuyentes principales, permite afirmar que han sido autoras de correspondencia (Moya-Anegón, 2012) en una proporción superior a los autores de género masculino. Revisando los 92 trabajos publicados en el IJDC en el trienio 2018 a 2020, en 37 de ellos los autores de correspondencia eran hombres (tercera columna de la tabla 2) y 54 eran mujeres (segunda columna de la tabla 2), más un trabajo cuyo autor de correspondencia era una institución.

Año	Liderazgo científico mujeres	Liderazgo científico hombres
2020	25	16
2019	10	11
2018	19	10

Tabla 2. Liderazgo científico en los trabajos publicados en IJDC 2018-2020.

Fuente: elaboración propia por los autores.

CONCLUSIÓN

Las disparidades de género en la ciencia son una realidad (Larivière *et al.* 2013). Si bien el dominio de la curación digital presenta indicios de un pequeño cambio en esa tendencia. En el periodo temporal revisado, Sarah Jones es la autora y mujer más productiva, Rebeca Koskela cumple con una función como nodo para ampliar la cohesión de la red y Kirsty Merret trabaja frecuentemente en colaboración atendiendo al vínculo derivado de la coautoría, rompiendo la tendencia esperada (Kwiek & Roszka, 2020). Se resalta el año 2020 por el importante cambio de dinámica respecto de las mujeres autoras de IJDC, momento temporal en el que éstas desempeñan con más frecuencia funciones de liderazgo científico.

REFERENCIAS

Alves, B.H., Dalessandro, R.C., & Santos, F.B.d. (2019). Colaboração Científica no Periódico Knowledge Organization: Elementos para Caracterização de um Domínio. In T.H.B. Barros & N.B. Tognoli (Eds.), *Organização do Conhecimento*

Responsável: Promovendo Sociedades Democráticas e Inclusivas (pp. 137-144). Ed. da UFPA.

- Bhargava, R. (2009). *Manifesto For The Content Curator: The Next Big Social Media Job Of The Future?*. <https://www.rohitbhargava.com/2009/09/manifesto-for-the-content-curator-the-next-big-social-media-job-of-the-future.html>
- Guallar, J., Codina, L., & Abadal, E. (2020). La investigación sobre curación de contenidos: análisis de la producción académica. *Ibersid*, 14(1), 13-22.
- Guimarães, J.A.C., Martínez-Ávila, D., & Alves, B.H. (2015). “Epistemic Communities in Knowledge Organization: An Analysis of Research Trends in the Knowledge Organization Journal.” In *ISKO UK biennial conference 13th–14th July 2015*, London.
- Hjørland, B. (2002). Domain analysis in information science: Eleven approaches – traditional as well as innovative. *Journal of Documentation*, 58(4), 422-462 <https://doi.org/10.1108/00220410210431136>
- Hjørland, B. (2017). Domain Analysis. *Knowledge Organization*, 44(6), 436-464.
- Hjørland, B., Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, 46(6), 400-425.
- Kwiek, M., & Roszka, W. (2020). Gender Disparities in International Research Collaboration: A Study of 25,000 University Professors. *Journal of Economic Surveys*. doi: 10.1111/joes.12395
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature News*, 504(7479), 211.
- Martínez-Ávila, D., Ibekwe, F., & Bochi, F. (2020). The Epistemic Communities and Evolution of Knowledge Domains: A Domain Analysis of the Journal Education for Information. In: M. Lykke, T. Svarre, M. Skov, & D. Martínez-Ávila (Eds.), *Knowledge Organization at the Interface* (pp. 264-273). Ergon Verlag. <https://doi.org/10.5771/9783956507762-264>
- Meiksins, P., Layne, P., Beddoes, K., Lewis, M., Masters, A.S., & Deters, J. (2019). Women in Engineering: A Review of the 2018 Literature. *SWE Magazine*. <https://alltogether.swe.org/2019/04/women-in-engineering-a-review-of-the-2018-literature/>
- Miguel, S., Hidalgo, M., Stubbs, E., Posadas, P., & Jaureguizar, E.O. (2013). Estudio bibliométrico de género en la paleontología de vertebrados. El caso de la revista argentina Ameghiniana (1957-2011). *Investigación Bibliotecológica*, 27(61), 133-155.
- Moya-Anegón, F. (2012). Liderazgo y excelencia de la ciencia española. El *Profesional de la Información*, 21(2), 125-128. <http://dx.doi.org/10.3145/epi.2012.mar.01>

- Moya-Anegón, F., Guerrero-Bote, V.P., Bornmann, L., & Moed, H.F. (2013). The research guarantors of scientific papers and the output counting: a promising new approach. *Scientometrics*, 97, 421-434.
- Oliveira, E.F.T.d, Alves, B.H., Prado, M.R.d., & Pavanelli, M.A. (2017). Produção Científica e Inserção Internacional da Revista Scire no Período de 2006 a 2014. *Scire*, 23(1), 47-56.
- Palomba, R. (2006). Does gender matter in scientific leadership? *Scientific Careers*, 133-137.
- Price DS. (1976). A General Theory of Bibliometric and Other Cumulative Advantage Processes. *Journal of the American Society for Information Science*, 292-306. doi: 10.1002/asi.4630270505
- Rusbridge, C., Burnhill, P., Ross, S., Buneman, P., Giaretta, D., Lyon, L., & Atkinson, M. (2005). The Digital Curation Centre: A Vision for Digital Curation. In: 2005 IEEE *International Symposium on Mass Storage Systems and Technology* (pp. 31-41). IEEE. doi: 10.1109/LGDI.2005.1612461
- Smiraglia, R.P. (2012). Shifting Intension in Knowledge Organization: An Editorial. *Knowledge Organization*, 39(6), 405-408.
- Smiraglia, R.P. (2015). *Domain Analysis for Knowledge Organization: Tools for Ontology Extraction*. Oxford: Chandos Publishing.