

La investigación sobre datos en Bibliotecología, Ciencias de la Información y de la Documentación en México y España

Georgina Araceli Torres Vargas

(Universidad Nacional Autónoma de México, Instituto de Investigaciones
Bibliotecológicas y de la Información)

gatv@unam.mx

Introducción

El uso del término datos, puede rastrearse desde 1646. Sin embargo, es hasta finales del siglo XVIII, que los datos implican hechos en forma de evidencia científica recopilada por experimentos, observaciones y otras investigaciones. Es durante el siglo XX que ha presentado un crecimiento constante en la literatura. (Borgman, 2015, p. 17).

El advenimiento y masificación de las tecnologías digitales potencializan las posibilidades de captar digitalmente una gran diversidad de fenómenos, lo que favorece la proliferación exponencial de los datos a velocidades, variedades y volúmenes inusitados.

Los datos no son objetos puros o naturales con una esencia propia. Toman significado del contexto en el que existen y desde la perspectiva del espectador. (Borgman, 2015, p. 18). Tienen múltiples procedencias y no tendrá las mismas implicaciones de una disciplina a otra. En gran medida el tratamiento que requieran estará determinado por su contexto y el uso que se pretenda hacer de ellos. Esto deja de manifiesto la necesidad de determinar como campo disciplinario nuestro actuar frente a los datos.

En el entendido de que los datos son un tema presente en la bibliotecología, las ciencias de la información y en la documentación, en este estudio se pretende indagar cuándo surge el interés por abordar a los datos en estas áreas del conocimiento, así como desde qué enfoques se han abordado, sobre todo en países como México y España. Esto podrá contribuir a realizar el estado de la cuestión en cuanto a la investigación sobre datos y en particular a conocer si es un tema muy reciente o ha ido madurando con el tiempo.

Metodología

Por medio de la búsqueda en Web of Science y Scopus se identificó la producción de las investigaciones sobre datos, correspondientes a la categoría científica de *library and*

information science tanto de México como de España.

Resultados y discusión

A continuación, se señalan los resultados que de manera diferenciada se obtuvieron en Web of Science y en Scopus.

a) Web of Science

En la búsqueda sobre el término *dato* en general no arrojó resultados para la categoría *library and information science*. Después se refinó la búsqueda con el tema *library and information science* y en la categoría *library and informationscience*, el resultado fue 1,158 publicaciones.

Desde 1970 está presente el tema «dato» en la literatura de bibliotecología y ciencia de la información con una publicación en ese año y en 1976 dos. En años posteriores aparece el tema en una o dos publicaciones, pero a partir de 1991 crece a 3 y se observa un incremento más sostenido, tal y como se observa en la siguiente tabla:

1992 (11)	2002 (20)	2012 (53)
1993 (11)	2003 (18)	2013 (61)
1994 (14)	2004 (26)	2014 (56)
1995 (7)	2005 (21)	2015 (121)
1996 (18)	2006 (30)	2016 (119)
1997 (17)	2007 (31)	2017 (123)
1998 (19)	2008 (44)	2018 (131)
1999 (24)	2009 (40)	2019 (6)
2000 (22)	2010 (43)	
2001 (21)	2011 (43)	

Tabla 1. Publicaciones por año

Se observa cómo desde el año 2012 se tuvo una recurrencia del tema, que rebasa los cincuenta artículos y se va incrementando de forma sostenida. La cifra de 6 apariciones en 2019, obedece a que este texto se escribió a principios de ese año, lo que impide conocer de qué forma se manifestó el tema en las publicaciones hasta diciembre del 2019.

La siguiente gráfica muestra que hubo un crecimiento exponencial de publicaciones sobre datos y bibliotecología entre 2015 y 2018.

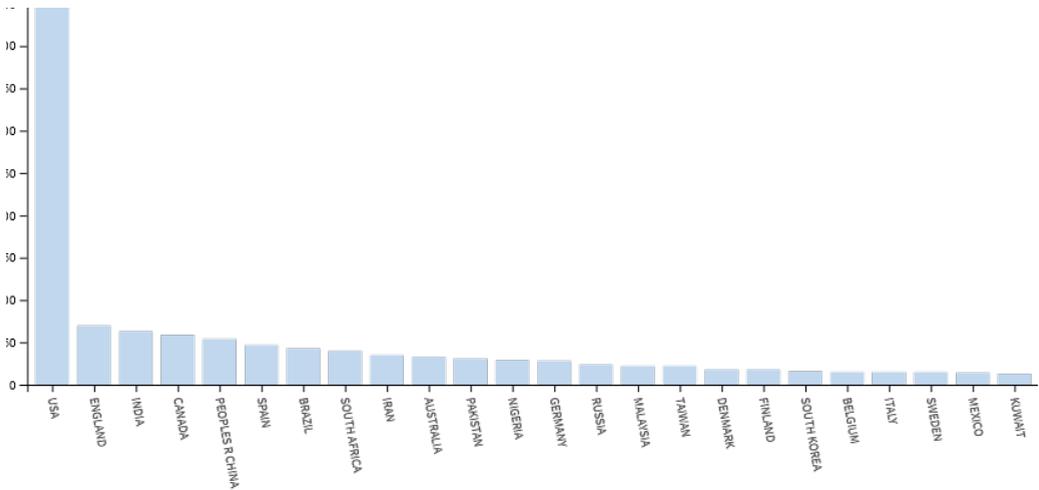


Fig. 4 Publicaciones por nacionalidad

Los siguientes son las revistas y libros en los que publicaron los autores mexicanos

E-CIENCIAS DE LA INFORMACION	4
INVESTIGACION BIBLIOTECOLOGICA	2
INFORMATION DEVELOPMENT	2
LIBRARY MANAGEMENT	1
INFORMATION AND LEARNING SCIENCES	1
INFORMACION: PERSPECTIVAS BIBLIOTECOLOGICAS Y DISTINCIONES INTERDISCIPLINARIAS (Capítulo)	1

Tabla 2. Revistas con publicaciones de autores mexicanos

Revistas en las que publicaron los autores españoles

PROFESIONAL DE LA INFORMACION	11
REVISTA ESPANOLA DE DOCUMENTACION CIENTIFICA	4
SCIENTOMETRICS	3
BID TEXTOS UNIVERSITARIAS DE BIBLIOTECONOMIA	2
I DOCUMENTACIO	2
ELECTRONIC LIBRARY	2
INFORMATION RESEARCH AN INTERNATIONAL	2
ELECTRONIC JOURNAL	2
INVESTIGACION BIBLIOTECOLOGICA	2
JOURNAL OF INFORMATION SCIENCE	2

Tabla 3. Revistas con dos o más publicaciones de autores españoles

Revistas con un solo artículo - autores españoles

14TH INTERNATIONAL SOCIETY OF SCIENTOMETRICS AND INFORMETRICS CONFERENCE ISSI
ASLIB JOURNAL OF INFORMATION MANAGEMENT
ASLIB PROCEEDINGS
HEALTH INFORMATION MANAGEMENT WHAT STRATEGIES
INFORMACAO SOCIEDADE ESTUDOS
INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT
INVESTIGACION BIBLIOTECOLOGIA
JOURNAL OF ACADEMIC LIBRARIANSHIP
JOURNAL OF INFORMETRICS
KNOWLEDGE ORGANIZATION
LEARNED PUBLISHING
LIBRARY HI TECH
LIBRI
PERFORMANCE MEASUREMENT AND METRICS
PORTAL LIBRARIES AND THE ACADEMY
PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON SCIENTOMETRICS AND INFORMETRICS
REVISTA GENERAL DE INFORMACION Y DOCUMENTACION

Tabla 3. Revistas con una publicación de autores españoles

Con respecto a los artículos recuperados de Web of Science, en el caso de México, se tiene que de un total de 11 únicamente 1 está relacionado:

- 2017 Repositorio de datos – Minería de datos

Mientras que, en el caso de España, de un total de 44 únicamente 7 están relacionados con el tema de datos y corresponden a los siguientes temas y años:

- 2018 Datos en el currículum de las iSchools
- 2018 Datos abiertos y big data en la Documentación periodística
- 2016 Big data, minería de texto y análisis de sentimientos
- 2016 Datos científicos abiertos
- 2014 Datos enlazados
- 2013 Datos abiertos enlazados
- 2011 Datos abiertos y datos abiertos enlazados

b) Scopus

En el caso de Scopus la búsqueda se complicó, porque en sus categorías no existe una enfocada a *library and informationscience*, así que se debió poner dicha oración, para excluir los otros resultados. Así se obtuvieron 6,451 documentos. Uno de los primeros criterios de exclusión fue enfocarse solo a revistas, teniendo como resultado 4,935 documentos. Además, al revisar el resultado, algunos de los documentos no se vinculan directamente con bibliotecología y estudios de la información.

El siguiente criterio de exclusión fue el de las revistas en las que apareció el artículo, siendo los excluidos, los que aparecieron en revistas de áreas del conocimiento que no fueran bibliotecología, ciencia de la información o documentación. Una vez hecha esa selección se tuvieron como resultados 3,658 documentos. En el caso de año de publicación, se tiene que el primer documento fue publicado en 1955 y el tema desde la década de 1990 ha tenido un crecimiento constante, hasta los 338 documentos en 2018.

La siguiente gráfica muestra el crecimiento que ha tenido la producción sobre datos en scopus:

Documents by year

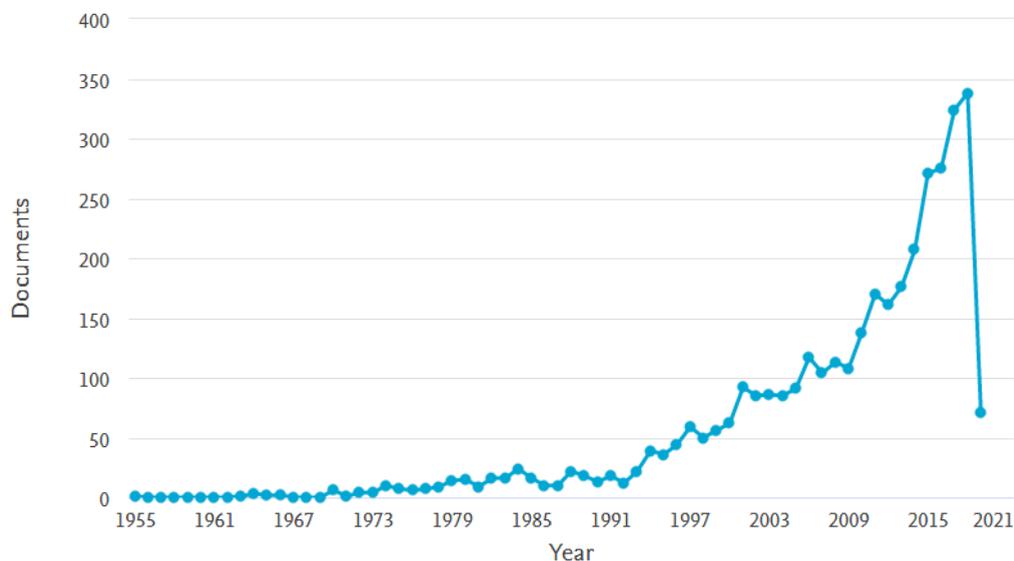


Fig. 5 Scopus. Crecimiento de la publicación en datos

Subject area ↓	Documents ↓
Social Sciences	1862
Medicine	1079
Computer Science	1036
Biochemistry, Genetics and Molecular Biology	319
Engineering	268
Agricultural and Biological Sciences	130
Arts and Humanities	123
Health Professions	110
Pharmacology, Toxicology and Pharmaceutics	106

En el caso de área temática, es evidente que es un tema multidisciplinario, aunque por obvias razones lo encabeza ciencias sociales, en segundo lugar, está medicina, en tercero computación y en cuarto lugar bioquímica, genética y biología molecular. En el caso de computación la razón es que se trata del área que genera las herramientas que permiten el análisis de datos, mientras que llama la atención medicina, siendo la explicación pertinente la importancia que en biología se le ha dado a las bibliotecas y a la información, reflejada en Medline.

Fig. 6 Scopus. Áreas temáticas

Al igual que en Web of Science se limitó de 2015 a 2018, dando como resultado 1,208 documentos. En el caso del país de origen, es evidente el liderazgo de Estados Unidos, al respecto resalta que España esté en el octavo lugar. México está en el lugar 32, pero se incluyó, porque es parte central de este documento.

Documents by country or territory

Compare the document counts for up to 15 countries/territories

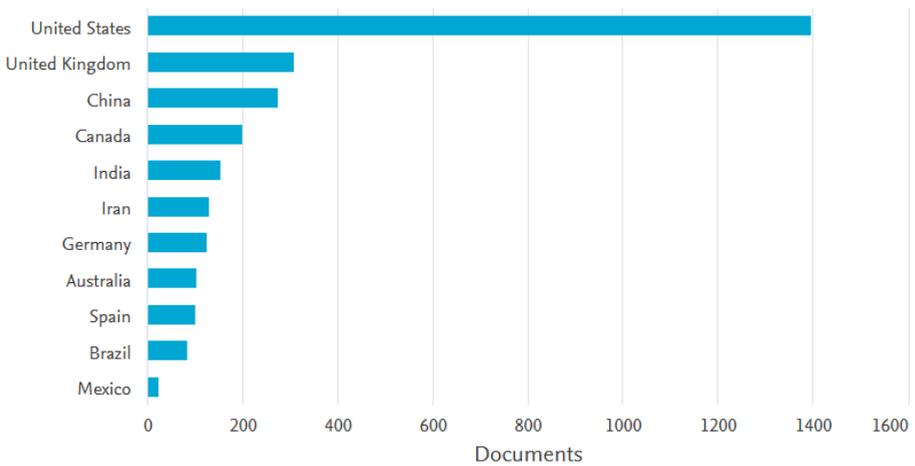


Fig. 7 Scopus. Publicaciones por nacionalidad

En la siguiente gráfica se observan los 10 autores con el mayor número de artículos.

Documents by author

Compare the document counts for up to 15 authors

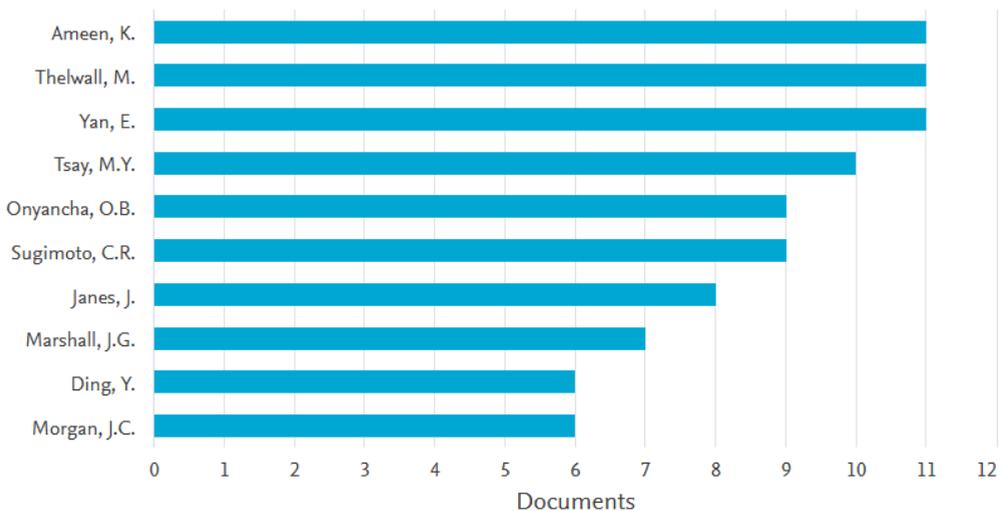


Fig. 8 Scopus. Autores con mayor producción

Como se observa, el análisis que arroja la recuperación de la producción científica en Scopus tampoco refleja el posicionamiento del término dato entre los más utilizados. Es claro que en la bibliotecología, documentación y ciencias de la información si bien se nota un incremento del estudio del tema relativo a los datos, son otras áreas en las que se manifiesta mayor número de artículos.

Esto no quiere decir que los datos no sean objeto de estudio de las áreas que se dedican al estudio de la información, sino que se requiere que estos estudios sean más numerosos al interior de dichas áreas.

Conclusiones

El estudio se realizó con base al análisis de las publicaciones que se registran en SCOPUS desde 1955 y en WOS desde 1970, en relación con las áreas de conocimiento: bibliotecología, ciencia de la información o documentación. El análisis permite arribar a diversas conclusiones, entre las que se tiene que en el ámbito de la bibliotecología y estudios de la información el análisis de los datos, como parte sustancial de la información ha sido de creciente interés para diversos investigadores de este área del conocimiento.

Como podemos ver por los resultados arrojados en WOS y SCOPUS, el tema ha tenido un crecimiento constante desde la década de 1990. Sin embargo, su crecimiento exponencial se ha dado a partir del año 2012. Si bien es a partir de esta época que se incrementa la producción, es factible derivar que el estudio de los datos es un tema que desde hace tiempo ha sido abordado. Incluso es posible considerar que varios de los métodos, técnicas y herramientas que se usan actualmente en el análisis de los datos tuvieron su origen en la bibliotecología y los estudios de la información, siendo el mejor ejemplo la bibliometría, que en la actualidad se conoce como ciencimetría y recientemente se han incorporado las métricas de la web y redes sociales.

En lo concerniente a las aportaciones de trabajos sobre el análisis de los datos se tiene que Estados Unidos es el líder en la producción sobre el tema, cuestión que es debatible, por diversas cuestiones, como el idioma de dichas bases de datos y que no se presta atención al origen de los autores, sino a la institución a la que pertenecen. El análisis de datos requiere de grandes bases de datos y una infraestructura de alto nivel con la que cuentan varias de las instituciones de ese país. A pesar de las cuestiones señaladas, destaca que España ocupa el octavo lugar y México el 32.

Por último, el debate no debe ser sobre si el análisis de datos es una nueva disciplina o área del conocimiento, sino si en la bibliotecología y estudios de la información se cuentan con los elementos necesarios para contribuir a dicho estudio. Asimismo, hay que

reflexionar en qué medida se requiere replantear los contenidos de los programas de estudio para que los egresados de estas carreras tengan claridad sobre su participación en el análisis de datos, ya sea como investigadores, profesores o profesionales.

Bibliografía

Borgman, C. L. (2015). *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, Massachusetts: MIT Press.

García, A. G., Ibáñez, A. P., Sapena, A. F., Mancebo, M. F. P., & Moreno, L. M. G. (2015). Herramientas de análisis de datos bibliográficos y construcción de mapas de conocimiento: Bibexcel y Pajek. *BiD: Textos universitaris de biblioteconomia i documentació*, (34), 11.

Van Eck, N., & Waltman, L. (2009). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.