

Ciencia de datos y la formación profesional en bibliotecología: un análisis textual y revisión de los currículos en Iberoamérica

Alonso Estrada Cuzcano*
Joel Alhuay-Quispe**

Artículo recibido:
10 de junio de 2024

Artículo aceptado:
22 de octubre de 2024

Artículo de investigación

RESUMEN

Los entornos digitales y la explosión informacional han transformado el trabajo profesional, lo cual ha requerido de cambios en la educación superior para desarrollar competencias en la era de los macrodatos y ciencia de datos. La literatura muestra cómo han evolucionado las actividades tradicionales de los profesionales de la bibliotecología hacia nuevos retos profesionales frente al surgimiento de técnicas y tendencias como la ciencia de datos. El artículo presenta una aproximación teórica-metodológica inicial sobre la inminente repercusión ontológica de la disciplina y su implicancia en la formación superior de la profesión. El objetivo es determinar, desde la enseñanza profesional

* Facultad de Letras y Ciencias Humanas, Universidad Nacional Mayor de San Marcos, Perú
mestradae@unmsm.edu.pe

** Facultad de Comunicaciones y Ciencias Administrativas, Universidad Privada San Juan Bautista, Perú
joel.alhuay@upsjb.edu.pe

de la bibliotecología, el estado actual de la incorporación de cursos en torno a la ciencia de datos en las universidades de la región, partiendo de una revisión de los planes de estudio de programas de pregrado en países de Iberoamérica. Además, busca identificar las tendencias de investigación sobre el manejo de grandes cantidades de datos así como los perfiles bibliotecarios relacionados estudiados en publicaciones científicas, a partir de la revisión de la literatura recuperada de la base de datos Web of Science entre 2015-2022 sobre “data science”, “big data”, “data scientist”, “databrarian”, “databrarianship” y “data librarian” en áreas de bibliotecología y ciencias de la información, seguido de un análisis de contenido textual usando técnicas de coocurrencia de términos y análisis de componentes principales.

Palabras clave: Bibliometría; Análisis textual; Educación en bibliotecología; Ciencia de datos

Data Science and Professional Training in Library Science: A Textual Analysis and Curriculum Review in Ibero-America

Alonso Estrada Cuzcano and Joel Albuay-Quispe

ABSTRACT

Current digital environments and the increasing information explosion have transformed the work of professionals, leading to the need for significant changes in higher education to develop competencies in the era of big data. The literature reveals how the traditional activities of library science professionals have evolved to face new challenges and emerging techniques and trends such as data science. This article presents a theoretical and methodological approach to the impending ontological impact on the discipline and its implications during the educational preparation in the field. The objective is to know the inclusion of data science courses in library science programs based on a review of undergraduate curricula from Ibero-American countries. Also, to identify research trends on big data management and related librarian profiles, as examined in scientific publications retrieved from the Web of Science database between 2015-2022, focusing on terms such as “data science,” “big data,” “data scientist,” “databrarian,” “databrarianship,” and “data librarian” within the fields

of library and information science, followed by a content analysis using co-occurrence techniques and principal component analysis.

Keywords: Bibliometrics; Textual Analysis; Library Science Education; Data Science

INTRODUCCIÓN

La ciencia de datos (del inglés, *data science*) es un campo interdisciplinario que requiere de habilidades en estadística, análisis de sistemas, programación y áreas temáticas específicas. Los conjuntos de habilidades esperados subrayan la importancia de tópicos como minería de datos y texto, análisis de macrodatos (*big data*), estadísticas, visualización de datos, codificación y aprendizaje automático. En este contexto es necesario realizar un análisis sobre la incidencia de la ciencia de datos en la formación del profesional de la información y sobre la especialización en el manejo de datos para transformarlos en formatos inteligibles utilizando técnicas analíticas (aprendizaje automático, aprendizaje profundo), analítica de texto, analítica web, lenguajes de programación, y aplicar análisis estadísticos complejos.

En el ámbito de la enseñanza superior han surgido nuevas áreas de investigación y aplicación tanto en la formación académica como en el quehacer profesional, originadas a partir de la manipulación y el análisis de los datos. Estos conceptos se definen, entre otros, por su origen en el inglés, como *big data* y *data science*. En el campo de las ciencias de la información (*id est*, la bibliotecología) y gestión de la información, aparecen conceptos como *open science*, *open data*, *open government*, los cuales están basados, por un lado, en la producción y la analítica de datos a partir de los productos y procesos de investigación científica; y por otro, en respuesta a políticas de transparencia de la gestión pública propiciadas por el crecimiento exponencial de la información electrónica.

Estudios sobre los programas de estudio en ciencia de datos

En su estudio, Wang y Lin (2019) reportaron 102 instituciones miembro en tres regiones –Asia Pacífico, Europa y África, y Norteamérica y Suramérica– que incluyen programas y cursos de ciencia de datos en 28 escuelas de información (*iSchools*), donde la mayoría de los cursos se concentraban en habilidades de análisis de datos; mientras que, la gestión de datos de investigación constituía un reducido porcentaje. Zhang *et al.* (2023) identificó en 12 programas de pregrado

en escuelas de información estadounidenses cursos de ciencia de datos, dirigidos a la minería y análisis de datos, estadística/matemáticas, programación, gestión de datos, datos, humanos, computadoras, sociedad, visualización de datos, introducción a la ciencia de datos, bases de datos, gestión de proyectos, sistemas de información y tecnología, aprendizaje automático y alfabetización de datos. Por otro lado, Urs y Minhaj (2023) revisaron y analizaron los sitios web de 122 escuelas de información donde identificaron 32 programas de posgrado que incluían explícitamente ciencia de datos y áreas relacionadas; sin embargo, los autores notaron que los planes de estudio que ofrecían cursos de ciencia de datos exhibían sesgos hacia la visualización de datos, el aprendizaje automático, la minería de datos, el procesamiento del lenguaje natural y la inteligencia artificial.

Estudios sobre ciencias de datos y bibliotecología

Si bien la contribución de las ciencias de la información (*library and information science*, en inglés) a la ciencia de datos en un estudio a partir de Web of Science (WoS) en el periodo 1980-2019 representaba apenas 3.4% del total (Virkus y Garoufallou, 2020), Chohdary MPhil, Asghar y Al Shaheer (2021) también utilizaron la misma base como fuente principal para extraer datos de los documentos publicados en el periodo 2010-2019 sobre “data science” y “library information science”, donde obtuvieron una muestra de 3 799 registros. Por su parte, Martínez Musiño (2021a) revisó la literatura científica y la intersección entre la ciencia de datos (“data science”) y la ciencia de la información (“information science”); su estudio recuperó 58 registros de Scopus y 49 de WoS en el periodo 1990-2020. En otra investigación del mismo autor donde solo empleó WoS, recuperó 49 referencias obtenidas de la colección principal en el periodo 1900-2020 (Martínez Musiño, 2021b).

De esta manera, el estudio busca responder medularmente a las siguientes preguntas de investigación:

- a) ¿Cuál es el estado de la enseñanza sobre ciencia de datos en las universidades con planes de estudio en bibliotecología, ciencias de la información, biblioteconomía y documentación?
- b) ¿Cuáles son las tendencias de investigación sobre ciencia de datos en publicaciones científicas de revistas de bibliotecología y ciencias de la información a partir de WoS?

MÉTODOS

Diseño y enfoque

La investigación tiene la naturaleza de estudio mixto: es cualitativo en cuanto emplea la revisión documental de material bibliográfico y la revisión de fuentes secundarias como técnicas para establecer una base de conocimiento sobre el tema; y descriptivo, ya que se apoya en análisis del corpus textual como técnica utilizada en estudios de enfoque bibliométrico.

Recuperación de datos

La recolección de los datos secundarios empleados en el estudio se constituye en dos etapas:

- a) Revisión de los sitios web de las universidades con planes de estudio en bibliotecología, ciencias de la información, biblioteconomía y documentación: esta etapa permitió conocer el estado de la cuestión y la aplicación práctica en el ámbito profesional de la cobertura geográfica de interés. Se realizó una búsqueda de los perfiles de programas profesionales y los cursos impartidos en universidades de América Latina y España que a nivel pregrado ofertaran bibliotecología, biblioteconomía, documentación e información, los cuales se obtuvieron a través de las páginas web institucionales.
- b) Selección de publicaciones en revistas de bibliotecología y ciencias de la información a partir de WoS: los artículos del estudio fueron recuperados de las colecciones ESCI y SSCI para el periodo 2015-2022, a partir de los términos de búsqueda “data science”, “big data”, “data scientist”, “databrarian”, “databrarianship” y “data librarian”. Estos resultados correspondientes al campo término (TS) de la base de datos se fusionaron con los registros procedentes de revistas de la categoría ‘Information Science & Library’ o que incluyeron términos como “library science”, “information and library”, “information science” y “librarianship” y que pertenecieran a la tipología de artículo (‘Article’) y revisión (‘Review’).

En la *Figura 1* se detalla el proceso de selección de los artículos de enseñanza sobre ciencia de datos, métodos, metodologías, técnicas y casos o experiencias de uso de tecnologías de medios virtuales, como en el aprendizaje virtual (*e-learning*) y el semipresencial (*b-learning*), o herramientas para el análisis de datos. Por otro lado, se excluyeron los trabajos de experiencias de negocios basados en macrodatos, estudios de corte bibliométrico y de analítica de datos (*data analytics*), con

excepción de trabajos generales de utilidad para la enseñanza académica. Además, se excluyeron artículos que mencionaran “big data era”, pero cuyos contenidos no correspondieran al tema propiamente o no incluyeran casos de uso para la enseñanza propiamente; así como estudios en los que se reseñaron iniciativas de utilización de los macrodatos para la gobernanza, pero con carácter administrativo mas no académico.

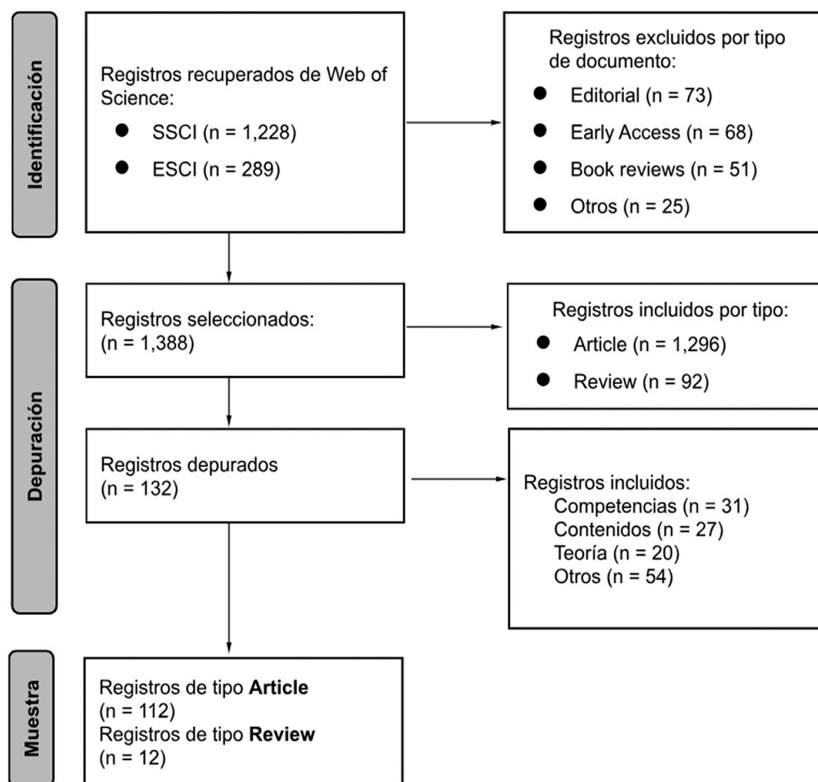


Figura 1. Procedimiento de selección de los registros
Fuente: elaboración de los autores (2024)

Elaboración de la taxonomía para el análisis de datos

Existen pocos estudios que refieren a la elaboración de una taxonomía para la ciencia de datos. Desai *et al.* (2022) hicieron un recuento de taxonomías descriptivas y normativas (Breiman, 2001; Chambers, 1993; Cleveland, 2001; Donoho, 2017), aunque se considera al enfoque de Cleveland (2001) como el más relevante para el estudio que desarrollamos, pues establece seis áreas para el analista de

datos; posteriormente esta taxonomía fue graficada por Porter (2015) y la denominó Cleveland's Data Science Taxonomy.

El análisis y sistematización de los registros obtenidos pudo realizarse a partir de una taxonomía obtenida a partir de la revisión de los propios registros y de las características de los contenidos. La revisión de literatura y antecedentes permitió establecer la siguiente taxonomía (*Figura 2*).



Figura 2. Conceptos de ciencia de datos en estudios de bibliotecología
Fuente: elaboración de los autores (2024)

Las categorías anteriores se agruparon de acuerdo con los criterios de los autores; se analizaron los contenidos y los respectivos enfoques de los estudios que comprendieron la muestra de la investigación. Los contenidos fueron organizados en cinco grupos: Teoría, métodos y fundamentos; Modelos y métodos; Investigación interdisciplinaria; Ética y aspectos jurídicos, y Pedagogía y enseñanza.

Procesamiento y análisis

El procesamiento de los datos implicó una etapa posterior a la recuperación de los datos: la depuración o preparación de los registros, como sugiere el estudio de Alhuay-Quispe, Estrada-Cuzcano y Bautista-Ynofuente (2022), el cual fue realizado

con el programa Publish or Perish (Harzing, 2007). Asimismo, como herramientas para los análisis de los datos bibliográficos y el procesamiento bibliométrico se usaron VOSviewer (van Eck y Waltman, 2010) y Bibliometrix (Aria y Cuccurullo, 2017), una librería de R empleada mediante el programa RStudio.

Los análisis de contenidos textuales se realizaron usando el método de coocurrencia de palabras, que identifica la frecuencia con la que dos términos aparecen juntos en el mismo conjunto de datos textuales o corpus. Este análisis se complementó con la técnica de centralidad de intermediación de aristas, definida por Girvan y Newman (2002), para identificar conglomerados dentro de la red de coocurrencia. También, se realizó un análisis de componentes principales, una técnica de análisis de conjuntos de datos que permite transformar un conjunto de variables posiblemente correlacionadas en un conjunto de valores de variables no correlacionadas llamadas componentes principales (Hendrix, 2010). La representación gráfica de este análisis se optimizó a partir del método de Louvain propuesto por Blondel *et al.* (2008), el cual permite la detección de comunidades en redes y optimiza una medida de modularidad de una red donde los nodos están más densamente conectados entre sí que con el resto de la red.

PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

Enseñanza de ciencia de datos en programas de bibliotecología de Iberoamérica

En Argentina, el Departamento de Bibliotecología y Ciencia de la Información de la Universidad de Buenos Aires (UBA) imparte la materia “Sistema de datos”, aunque esta proviene de la Licenciatura en Sistemas de Información de las Organizaciones de la Facultad de Ciencias Económicas de la misma institución y está orientado a la inteligencia empresarial (*business intelligence*).

En Brasil, el Programa de Biblioteconomía de la Universidad de Brasília (UnB) ofrece la asignatura “Tópicos especiais em biblioteconomia e ciência da informação: introdução a ciência de dados para a biblioteconomia”, en esta se considera a la ciencia de datos como uno de los tópicos especiales, aunque tiene carácter electivo u optativo. Por otro lado, el Programa de Biblioteconomía, perteneciente al núcleo del Programa en Ciencia de la Información, de la Universidad Federal del Rio Grande do Norte (UFRN) ofrece las asignaturas “Data science” y “Visualização da informação”, pero las sumillas no brindan detalles. El Programa de Biblioteconomía de la Universidad Federal del Cariri (UFCA) oferta la materia “Introdução à ciência de dados” y, por su parte, el Programa de Biblioteconomía y Ciencia de la Información de la Universidad de São Paulo (USP), la

asignatura “Inteligência artificial no tratamento da informação registrada”, en la cual convergen la inteligencia artificial y los servicios de información.

En Colombia, la Escuela Interamericana de Bibliotecología de la Universidad de Antioquia (UdeA) ofrece el curso “Big data”, descrito como conceptual o teórico y es optativo. Cabe mencionar que, aunque su plan de acción 2018-2021 menciona la irrupción de la ciencia de datos, aún no se han implementado cursos específicos en el plan de estudios (UdeA, 2018). La carrera en Ciencia de la Información, Bibliotecología y Archivística de la Pontificia Universidad Javeriana (PUJ) imparte el curso “Ciencia de datos: introducción a la ciencia de datos”, que brinda una perspectiva panorámica y tiene carácter de electivo.

En España, el Grado en Gestión Digital de Información y Documentación de la Universidad de La Coruña (UDC) oferta la materia “Ciencia de datos”, pero en realidad es un curso de estadística y se incluye a otro llamado “Taller de habilidades informacionales y digitales: representación gráfica y visualización digital de la información”, por los aspectos relacionados con la visualización de datos. En el Grado de Información y Documentación de la Universidad de Granada (UGR) se imparte el curso “Tratamiento masivo de datos”, aunque es optativo (*Tabla 1*).

Universidad	Curso	Créditos
Argentina		
Universidad de Buenos Aires (UBA). Bibliotecología y Ciencias de la Información	“Sistema de datos”	4
Brasil		
Universidade de Brasília (UnB). Programa de Biblioteconomia [<i>Curso de Biblioteconomia</i>]	“Tópicos especiais em biblioteconomia e ciência da informação: introdução a ciência de dados para a biblioteconomia”	4
Universidade Federal do Rio Grande do Norte (UFRN). Programa de Biblioteconomia del núcleo del Programa en Ciencia de la Información [<i>Curso de Graduação em Biblioteconomia pertencente ao Núcleo Específico de Curso de Graduação em Ciência da Informação</i>]	“Data science”	2
Universidade Federal do Cariri (UFCA). Programa de Biblioteconomia [<i>Curso de Biblioteconomia</i>]	“Visualização da informação” “Introdução à ciência de dados”	2
Universidade de São Paulo (USP). Programa de Biblioteconomia y Ciencia de la Información [<i>Curso de Biblioteconomia e Ciência da Informação</i>]	“Inteligência artificial no tratamento da informação registrada”	4

Colombia		
Universidad de Antioquia (UdeA). Escuela Interamericana de Bibliotecología	"Big data"	2
Pontificia Universidad Javeriana (PUJ). Carrera de Ciencia de la Información, Bibliotecología y Archivística	"Ciencia de datos: introducción a la ciencia de datos" "Visualización de datos: explorando y visualizando datos"	2
España		
Universidad de La Coruña (UDC). Grado en Gestión Digital de Información y Documentación	"Ciencia de datos" "Taller de habilidades informacionales y digitales: representación gráfica y visualización digital de la información"	6
Universidad de Granada (UGR). Grado de Información y Documentación	"Tratamiento masivo de datos"	6

Tabla 1. Ciencia de datos en programas de bibliotecología de América Latina y España
Fuente: elaboración de los autores (2024)

Análisis textual sobre ciencia de datos y bibliotecología

En un primer análisis, a partir de la agrupación conceptual de los contenidos de acuerdo con la taxonomía propuesta en la *Figura 2*, se encuentra que la mayor parte de las investigaciones que emplean los términos de datos y los perfiles bibliotecarios relacionados están orientados a los fundamentos y al estudio teórico de la ciencia de datos (40 %), seguido de las prácticas de enseñanza-aprendizaje y la evaluación de competencias (32 %). En cambio, los tópicos de investigación bibliométrica y los estudios de caso junto a la ética de datos apenas superan el 10 % (*Tabla 2*).

Concepto	Tema	Documentos	%
Ciencia de datos	Fundamentos	30	24.19 %
	Teoría	20	16.13 %
Pedagogía	Competencias	31	25.00 %
	Enseñanza	9	7.26 %

Investigación	Bibliometría	8	6.45 %
	Estudio de caso	8	6.45 %
Ética de datos	Ética	8	6.45 %
	Privacidad	6	4.84 %
Métodos	Herramientas	4	3.23 %

Tabla 2. Agrupación conceptual de las publicaciones analizadas
Fuente: elaboración de los autores (2024)

A través de un diagrama de Sankey, la *Figura 3* muestra la evolución temática de los conceptos recurrentes en los artículos analizados a partir de las palabras clave del autor en los periodos 2015-2019 y 2020-2022, representados mediante el método de coocurrencia de palabras y la técnica de centralidad de intermediación de aristas. Esto arrojó un decrecimiento en los estudios sobre “big data” hacia el segundo periodo, pero deja a consecuencia la diversificación hacia conceptos derivados como “data science” y “data mining”. Otros conceptos asociados, como “education” y “academic libraries” resultaron emergentes, esto podría relacionarse con tópicos de investigación o del quehacer profesional en bibliotecología. El concepto “artificial intelligence” se mantuvo en ambos lapsos, pero derivado en parte del tópico de “privacy”, esto estaría vinculado a asuntos de ética en el manejo de datos y el uso de inteligencia artificial generativa.

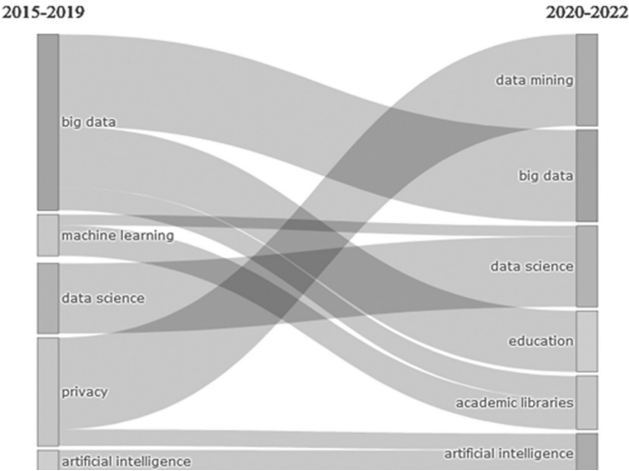


Figura 3. Diagrama de Sankey de las palabras clave
Fuente: elaboración de los autores (2024)

La *Figura 4* presenta un análisis de los componentes principales por medio del método de Louvain a partir de los términos clave de los documentos analizados;

muestra las categorías basadas en centralidad y densidad, dos dimensiones que crean un espacio bidimensional dividido en cuatro cuadrantes: temas nicho, temas motores, temas emergentes o en declive y temas básicos. Los temas con mayor presentación en el gráfico corresponden a tres conglomerados de los temas motores en el cuadrante superior derecho, de los cuales, un grupo está vinculado directamente con la bibliotecología, cuyos términos “data librarian”, “academic libraries” y “research data management” implican no solo un campo de estudio en la disciplina, sino una especialización desde la perspectiva laboral. En el cuadrante de temas emergentes del cuadrante inferior izquierdo destacan dos conceptos, “artificial intelligence” y “bibliometrics”, cada uno con un conglomerado que en estudios posteriores tendrá mayor diversidad o consolidación hacia otros cuadrantes. En el cuadrante superior izquierdo, de los temas nicho, si bien “information management” y “ethics” son del mismo conglomerado, estos temas resultan muy especializados o están bien desarrollados internamente, pero no se encuentran muy conectados con otros temas en los demás cuadrantes. Respecto a los temas básicos del cuadrante inferior derecho, los términos “education” y “digital preservation”, según el análisis, son temas fundamentales para el campo de estudios y están bien conectados con los temas de los otros cuadrantes; sin embargo, no están altamente desarrollados internamente, pero como constructos básicos son cruciales para la comprensión general del estudio de la ciencia de datos y su relación con la bibliotecología.

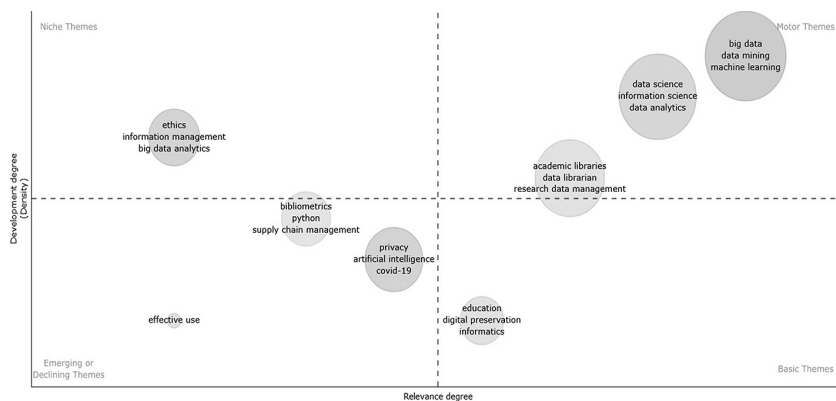


Figura 4. Mapa temático del análisis de componentes principales de las palabras clave
Fuente: elaboración de los autores (2024)

En la *Figura 5* se muestran conglomerados o clústeres que contienen al menos 30 términos cada uno, los cuales son representados mediante el método no normalizado donde las coocurrencias se cuentan directamente sin ajustar por la frecuencia relativa de términos. Los datos del gráfico se representan con una

[illegible]

Los estudios seleccionados en la muestra del estudio reportan que la mayoría de las investigaciones publicadas en revistas científicas de la disciplina no se encuentran estrechamente emparentadas con la bibliotecología, la biblioteconomía o la documentación e información; más bien, pertenecen a diversos campos de estudio, como educación, gestión, administración pública, entre otras. Así también, se encuentra que la bibliotecología ha empezado a incorporar conceptos como “data librarian”, “research data management” y “data literacy”. Esto demuestra cómo la

disciplina está actualmente encaminada no solo hacia el manejo y tratamiento del documento y la información como objeto de estudio, sino también de los datos en sí mismos. El análisis de las tendencias temáticas en el campo de la ciencia de datos dentro de la bibliotecología resalta la evolución y adaptación del perfil profesional a las demandas tecnológicas emergentes y a las tendencias, lo cual refleja la necesidad de que tanto los estudiantes en formación como los profesionales bibliotecólogos adquieran nuevas competencias y habilidades que permitan gestionar grandes volúmenes de información y desarrollar habilidades críticas en áreas de disciplinas conexas.

DISCUSIÓN

Los estudios previos documentan retos actuales y desafíos importantes para la profesión de la bibliotecología en materia de recolección, análisis, tratamiento y divulgación de grandes cantidades de datos, no solo desde la perspectiva de la investigación disciplinar, como la bibliometría, sino también para el rol del bibliotecario más allá de las bibliotecas tradicionales; es decir, en la gestión de datos abiertos de investigación, en la privacidad y ética de los datos, en el papel del bibliotecario de datos y en la alfabetización en datos.

La educación superior universitaria en ciencias de la información y bibliotecología implica el manejo de datos masivos y de diversos conjuntos de datos, capacidades que añaden valor a la toma de decisiones y establecimiento de políticas públicas, y para lo cual se precisa una formación adecuada y experiencia específica. En ese sentido, ya se ha propuesto implementar la alfabetización en datos (Koltay, 2015) y la gestión de datos de investigación (Andrikopoulou, Rowley y Walton, 2021) en las bibliotecas universitarias. En cambio, Ndukwe y Daniel (2020) formularon un curso de analítica de enseñanza (*teaching analytics*) para la formación en ciencia de datos; este estaría compuesto por tres elementos: *data*, *visualization* y *analysis*, al tiempo de que un bibliotecario de datos debe difundir los hallazgos importantes relacionados con la investigación recopilando datos relevantes de varias fuentes (Semeler, Pinto y Rozados, 2019).

Por lo que, para futuros estudios no solo debería realizarse un análisis a partir de las sumillas, sino elaborar una revisión de los sílabos de cada programa de estudios, como propone Khan (2020), que permita orientar sustancialmente las decisiones de los cambios en los planes de estudio basadas en los datos. Al respecto, Wang (2018) señala que la misión y la naturaleza de la ciencia de datos son consistentes con las de la ciencia de la información, que están estrechamente interrelacionadas y que juntas forman los componentes de la investigación de la “cadena de información”, pues desde la perspectiva

del proceso tienen preocupaciones similares (1246). El autor las denomina como disciplinas gemelas, pero es necesario prever la demanda de formación en campos tan especializados y, a la vez, tan multidisciplinarios.

Es importante el desarrollo de la ética de datos en la inclusión de la responsabilidad, innovación, programación, piratería y códigos profesionales (Floridi y Taddeo, 2016); en la bibliotecología y ciencias de la información debe ser un aspecto importante que privilegie la confidencialidad (datos personales) y la privacidad (uso de la información), lo cual involucra una permanente actualización de los códigos de ética de los profesionales (Francis, 2021; Estrada-Cuzcano y Alfaro-Mendives, 2021). En ese contexto, también resulta innegable la reflexión sobre el impacto de la inteligencia artificial en el quehacer bibliotecario como una tarea urgente, puesto que los avances de la apropiación de la inteligencia artificial están recientemente desarrollados en algunos sectores de la profesión, pero aún no se han materializado a gran escala (Cox y Mazumdar, 2024).

Finalmente, en este nuevo escenario de proliferación de técnicas avanzadas para el tratamiento de grandes volúmenes de datos y el aprovechamiento de la inteligencia artificial generativa en el quehacer profesional, que origina el surgimiento de disciplinas convergentes a las ciencias de la información, como las humanidades digitales, también repercute tanto en las nuevas competencias de la formación profesional y el rol profesional en bibliotecología, los cuales requieren revisarse y replantearse desde los planes de estudios hasta el mercado laboral.

CONCLUSIONES

La incorporación de la enseñanza en materias de ciencia de datos en escuelas de bibliotecología o biblioteconomía de universidades de Iberoamérica aún se encuentra en un estado inicial, con apenas cuatro casos en Brasil, dos instituciones en Colombia y en España, y solo una universidad en Argentina. Sin embargo, este estudio permite entender que la educación universitaria en bibliotecología implica el manejo de datos masivos y diversos conjuntos de datos, capacidades que añaden valor a la toma de decisiones y al establecimiento de políticas públicas, que requieren de una formación adecuada y experiencia específica.

Los estudios de tendencias a partir de los análisis textuales han permitido conocer las repercusiones a nivel teórico de la ciencia de datos en el campo de la bibliotecología. La mayoría de las investigaciones analizadas evidencian que los estudios sobre el uso de los macrodatos y perfiles bibliotecarios en torno a los datos aún están escasamente relacionados con la bibliotecología, la biblioteconomía, la documentación o los estudios de la información y, más bien, proceden de áreas como educación y administración; en cambio, pocos trabajos se centran en

la utilización de datos para la investigación o para mejorar la calidad de la educación, sin detenerse en la repercusión de la enseñanza o el aprendizaje a nivel del pregrado. Por lo que, esta revisión, tanto a nivel de literatura previa y de planes de estudio, contribuirá a orientar la actualización de los currículos y el perfil del egresado universitario en bibliotecología y ciencias de la información.

REFERENCIAS

- Alhuay-Quispe, Joel, Alonso Estrada-Cuzcano y Lourdes Bautista-Ynofuente. 2022. "Análisis y visualización de datos en estudios bibliométricos". *JLIS.it* 13 (2): 58-73. <https://doi.org/10.36253/jlis.it-461>
- Andrikopoulou, Angeliki, Jennifer Rowley y Geoff Walton. 2021. "Research Data Management (RDM) and the Evolving Identity of Academic Libraries and Librarians: A Literature Review". *New Review of Academic Librarianship* 28 (4): 349-65. <https://doi.org/10.1080/13614533.2021.1964549>
- Aria, Massimo, y Corrado Cuccurullo. 2017. "Bibliometrix: An R-Tool for Comprehensive Science Mapping Analysis". *Journal of Informetrics* 11 (4): 959-75. <https://doi.org/10.1016/j.joi.2017.08.007>
- Blondel, Vincent, Jean-Loup Guillaume, Renaud Lambiotte y Etienne Lefebvre. 2008. "Fast Unfolding of Communities in Large Networks". *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10), p10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Breiman, Leo. 2001. "Statistical Modeling: The Two Cultures (With Comments and a Rejoinder by the Author)". *Statistical Science* 16 (3): 199-231. <https://doi.org/10.1214/ss/1009213726>
- Chambers, John. 1993. "Greater or Lesser Statistics: A Choice for Future Research". *Statistics and Computing* 3 (4): 182-84. <https://doi.org/10.1007/BF00141776>
- Chohdary MPhil, Nazia Ismail, Mahe Bushra Asghar y Muhammad Ahmad Al Shaheer. 2021. "Predicting LIS Scholarly Research Directions in the Era of Data Science". *Library Philosophy and Practice*, e6328. <https://digitalcommons.unl.edu/libphilprac/6328/>
- Cleveland, William. 2001. "Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics". *International Statistical Review* 69 (1): 21-26. <https://doi.org/10.1111/j.1751-5823.2001.tb00477.x>
- Cox, Andrew, y Suvoodeep Mazumdar. 2024. "Defining Artificial Intelligence for Librarians". *Journal of Librarianship and Information Science* 56 (2): 330-40. <https://doi.org/10.1177/09610006221142029>
- Desai, Jules, David Watson, Vincent Wang, Mariarosaria Taddeo y Luciano Floridi. 2022. "The Epistemological Foundations of Data Science: A Critical Analysis". *SSRN*, e4008316. <https://doi.org/10.2139/ssrn.4008316>
- Donoho, David. 2017. "50 Years of Data Science". *Journal of Computational and Graphical Statistics* 26 (4): 745-66. <https://doi.org/10.1080/10618600.2017.1384734>

- Estrada-Cuzcano, Alonso, y Karen Lizeth Alfaro-Mendives. 2021. "Analysis of Professional Secrecy in Ibero-America: Ethical and Legal Perspectives". *IFLA Journal* 48 (3): 432-38.
<https://doi.org/10.1177/03400352211057149>
- Floridi, Luciano, y Mariarosaria Taddeo. 2016. "What is Data Ethics?". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2083), e20160360.
<https://doi.org/10.1098/rsta.2016.0360>
- Francis, Mary. 2021. "The Treatment of Privacy in Professional Codes of Ethics: An International Survey". *The Library Quarterly* 91 (3): 304-21.
<https://doi.org/10.1086/714320>
- Girvan, Michelle, y Mark Newman. 2002. "Community Structure in Social and Biological Networks". *Proceedings of the National Academy of Sciences* 99 (12): 7821-26.
<https://doi.org/10.1073/pnas.122653799>
- Harzing, Anne-Wil. 2007. Publish or Perish. V8. Windows.
<https://harzing.com/resources/publish-or-perish>
- Hendrix, Dean. 2010. "Relationships Between Association of Research Libraries (ARL) Statistics and Bibliometric Indicators: A Principal Components Analysis". *College and Research Libraries* 71 (1): 32-41.
<https://doi.org/10.5860/0710032>
- Hernández-Pérez, Tony, y María-Antonia García-Moreno. 2013. "Datos abiertos y repositorios de datos: nuevo reto para los bibliotecarios". *El Profesional de la Información* 22 (3): 259-63.
<https://doi.org/10.3145/epi.2013.may.10>
- Khan, Hammad. 2020. "Exploring Data Science Learning Objectives in LIS Education". *Proceedings for the Association for Library and Information Science Education Annual Conference: ALISE 2020*, 317-25.
<https://hdl.handle.net/2142/108828>
- Koltay, Tibor. 2015. "Data Literacy for Researchers and Data Librarians". *Journal of Librarianship and Information Science* 49 (1): 3-14.
<https://doi.org/10.1177/0961000615616450>
- Martínez Musiño, Celso. 2021a. "Ciencia de datos versus ciencia de la información: informetría y análisis textual". *Scire: Representación y Organización del Conocimiento* 27 (2): 31-42.
<https://doi.org/10.54886/scire.v27i2.4731>
- Martínez Musiño, Celso. 2021b. "La informetría y el análisis del discurso aplicados a la producción científica en la ciencia de datos y ciencia de la información". *E-Ciencias de la Información* 11 (2), e45234.
<https://doi.org/10.15517/eci.v11i2.45234>
- Ndukwe, Ifeanyi Glory, y Ben Kei Daniel. 2020. "Teaching Analytics, Value and Tools for Teacher Data Literacy: A Systematic and Tripartite Approach". *International Journal of Educational Technology in Higher Education* 17 (1), 22.
<https://doi.org/10.1186/s41239-020-00201-6>
- Porter, Tanya Thames. 2015. "Identifying the Data Scientist Amongst STEM Educators: An Introspective Survey of Work Skills". Tesis de doctorado, Universidad de Capella.
<https://doi.org/10.13140/RG.2.1.3531.0168>

- Semeler, Alexandre Ribas, Adilson Luiz Pinto y Helen Beatriz Frota Rozados. 2019. "Data Science in Data Librarianship: Core Competencies of a Data Librarian". *Journal of Librarianship and Information Science* 51 (3): 771-80.
<https://doi.org/10.1177/0961000617742465>
- UdeA (Universidad de Antioquia). 2018. *Plan de Acción Escuela Interamericana de Bibliotecología 2018-2021*. Escuela Interamericana de Bibliotecología, Universidad de Antioquia.
<https://bit.ly/2VYGE1d>
- Urs, Shalini, y Mohamed Minhaj. 2023. "Evolution of Data Science and Its Education in iSchools: An Impressionistic Study Using Curriculum Analysis". *Journal of the Association for Information Science and Technology* 74 (6): 606-22.
<https://doi.org/10.1002/asi.24649>
- Van Eck, Nees Jan, y Ludo Waltman. 2010. "Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping". *Scientometrics* 84 (2): 523-38.
<https://doi.org/10.1007/s11192-009-0146-3>
- Virkus, Sirje, y Emmanouel Garoufallou. 2020. "Data Science and Its Relationship to Library and Information Science: A Content Analysis". *Data Technologies and Applications* 54 (5): 643-63.
<https://doi.org/10.1108/DTA-07-2020-0167>
- Wang, Lin. 2018. "Twinning Data Science With Information Science in Schools of Library and Information Science". *Journal of Documentation* 74 (6): 1243-57.
<https://doi.org/10.1108/JD-02-2018-0036>
- Wang, Yi-Yu, y Chi-Shiou Lin. 2019. "A Survey of Data Science Programs and Courses in the iSchools". *Proceedings of the Association for Information Science and Technology* 56 (1): 801-02.
<https://doi.org/10.1002/pa2.184>
- Zhang, Yin, Dan Wu, Loni Hagen, Il-Yeol Song, Javed Mostafa, Sam Oh, Theresa Anderson, et al. 2023. "Data Science Curriculum in the iField". *Journal of the Association for Information Science and Technology* 74 (6): 641-62.
<https://doi.org/10.1002/asi.24701>

Para citar este texto:

- Estrada Cuzcano, Alonso, y Joel Alhuay-Quispe. 2025. "Ciencia de datos y la formación profesional en bibliotecología: un análisis textual y revisión de los currículos en Iberoamérica". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 39 (102): 203-220.
<http://dx.doi.org/10.22201/iibi.24488321xe.2025.102.58943>