

OCTUBRE/DICIEMBRE,
2022, VOL. 36, N.º 93,
MÉXICO, ISSN 0187-358X

Investigación Bibliotecológica

ARCHIVONOMIA, BIBLIOTECOLOGIA E INFORMACION

ibi



Investigación Bibliotecológica

ARCHIVONOMÍA, BIBLIOTECOLOGÍA E INFORMACIÓN

Vol. 36, núm. 93, octubre/diciembre, 2022, México, ISSN: 2448-8321

DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2022.93>



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Instituto de Investigaciones Bibliotecológicas y de la Información

Contenido

INVESTIGACIÓN BIBLIOTECOLÓGICA, VOL. 36, NÚM. 93, OCTUBRE/DICIEMBRE, 2022, MÉXICO, ISSN: 2448-8321
DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2022.93>

ARTÍCULOS

- **Experiencia de clasificación automática de documentos sobre Ciencias de la Vida y Biomedicina obtenidos del Web of Science** [Experience of automatic classification of Life Sciences and Biomedicine documents obtained from Web of Science] 13-32
Luis Roberto Polo Bautista e Israel Polo Bautista
- **Modelo de fluxo de informação científica para as áreas de informação** [Scientific information flow model for information areas] 33-49
Henrique Denes Hilgenberg Fernandes e Jayme Leiro Vilan Filho
- **Acceso abierto al patrimonio informacional en humanidades y ciencias sociales de Instituciones de Educación Superior: propuesta conceptual y análisis de normatividades** [Open access to informational heritage in Humanities and Social Sciences of Higher Education Institutions: conceptual proposal and analysis of regulations] 51-71
Juan Miguel Palma Peña
- **Inteligencia artificial en los procesos documentales de los archivos digitales sonoros** [Artificial intelligence in the documentary processes of digital sound archives] 73-88
Georgina Sanabria Medina y Perla Olivia Rodríguez Reséndiz
- **Ontologías terminológicas en la organización temática de dominios de conocimiento** [Terminological ontologies in subject organization of knowledge domains] 89-113
Adriana Suárez Sánchez
- **Políticas de información en las bibliotecas del Congreso General de los Estados Unidos Mexicanos** [Information policy in the libraries of the General Congress of the United Mexican States] 115-134
Arturo Vargas Huerta
- **Políticas editoriales de revistas de Comunicación en acceso abierto indexadas en Scopus** [Editorial policies of Communications in Open Access journals indexed in Scopus] 135-151
Bexi Perdomo y Oscar Alberto Morales

- **Producción de libros académicos en Colombia entre 2013 y 2019: un avance para estudios comparados** [Production of academic books in Colombia between 2013 and 2019: a step forward in comparative studies]
Esteban Giraldo-González, Elea Giménez-Toledo y Juan Felipe Córdoba-Restrepo

153-176
- **Uso de las habilidades digitales en el proceso de enseñanza-aprendizaje en ciencias de la información en un entorno virtual durante la pandemia por Covid 19** [Use of digital skills in the teaching-learning process in a virtual environment during the Covid 19 pandemic]
Eduardo Oliva-Cruz y Adriana Mata-Puente

177-193
- **La producción científica de la Universidad Autónoma de Querétaro indizada en WoS, 2010-2020** [The scientific production of the Autonomous University of Querétaro indexed in WoS, 2010-2020]
Ramsés Jabín Oviedo Pérez y Ángel Bravo Vinaja

194-215

Investigación Bibliotecológica: archivonomía, bibliotecología e información [en línea] / ed. por el Centro Universitario de Investigaciones Bibliotecológicas. -Vol. 1, No. 1 (ago. 1986) - . Ciudad de México: Universidad Nacional Autónoma de México, Centro Universitario de Investigaciones Bibliotecológicas, 1986 - V. Semestral, 1986 – 2007; a partir de vol. 22, no. 44 (ene. / abr. 2008) - , la periodicidad cambió a cuatrimestral. A partir de enero de 2018, la periodicidad cambió a trimestral. Resúmenes en español e inglés, a partir del vol. 3, no. 1 (jul / dic. 1987) Disponible también en idioma inglés a partir del vol. 28, no. 62 (ene. / abr. 2014) Publicado por la misma dependencia bajo su nombre actual: Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliotecológicas y de la Información vol. 26, no. 56 (ene. / abr. 2012) - . Disponible para su consulta en línea a partir del vol. 1, no. 1 (ago. 1986) - . Publicado en formato electrónico a partir del vol. 30, no. 70 (sep. / dic. 2016) - . Todos los artículos cuentan con DOI en forma individual. Disponible en: <http://rev-ib.unam.mx/ib>
ISSN 0187-358X (impreso)
ISSN 2448-8321 (en línea)



Investigación Bibliotecológica: archivonomía, bibliotecología e información, Vol. 36, No. 93, octubre-diciembre 2022, es una publicación trimestral editada por la Universidad Nacional Autónoma de México, Ciudad Universitaria, Alcaldía Coyoacán, Ciudad de México, C.P. 04510, a través del Instituto de Investigaciones Bibliotecológicas y de la Información, Circuito Escolar s/n, Torre II de Humanidades, Piso 12, Ciudad Universitaria, Col. Copilco, Alcaldía Coyoacán, Ciudad de México, C.P. 04510, Tel. 56230349 y 56230337, <http://rev-ib.unam.mx/ib>, correos electrónicos: revista@iibi.unam.mx, drevista@iibi.unam.mx. Editor responsable: Dr. Egbert John Sánchez Vanderkast. Reserva de Derechos al uso Exclusivo No. 04-2016-041813344600-203, ISSN: 2448-8321, ambos otorgados por el Instituto Nacional del Derecho de Autor. Responsable de la última actualización de este número, Dra. María de Jesús Madera Jaramillo, Instituto de Investigaciones Bibliotecológicas y de la Información, Circuito Escolar s/n, Torre II de Humanidades, Piso 12, Ciudad Universitaria, Col. Copilco, Alcaldía Coyoacán, Ciudad de México, C.P. 04510, Tel. 56230337, fecha de la última modificación, noviembre 2022.

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación. Se autoriza la reproducción total o parcial de los textos aquí publicados siempre y cuando se cite la fuente completa y la dirección electrónica de la publicación.

DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2022.93>

REVISTA INDIZADA EN:

- Clarivate Analytics Web of Science
- Clarivate Analytics JCR
- Scopus
- SCImago Journal & Rank
- Índice de Revistas Mexicanas de Investigación Científica y Tecnológica del CONACyT
- SciELO Citation Index
- SciELO
- Google Academics
- Latindex
- DOAJ
- Elsevier Science Direct
- LISA
- LISTA Full Text
- LISS
- INFOBILA
- CSIC e-revist@s
- Dialnet
- CLASE
- HAPI

Esta revista está disponible en texto completo y en acceso abierto en:

- Revista IBI: <http://rev-ib.unam.mx/ib>
- SciELO: http://www.scielo.org.mx/scielo.php?script=sci_serial&pid=0187-358X&lng=es&nrm=iso
- Science Direct: <https://www.sciencedirect.com/journal/investigacion-bibliotecologica-archivonomia-bibliotecologia-e-informacion>

DIRECTOR DE LA REVISTA

DR. EGBERT JOHN SÁNCHEZ VANDERKAST

CONSEJO EDITORIAL

DRA. ESTELA MORALES CAMPOS
Universidad Nacional Autónoma de México
Ciudad de México, México

DR. CARLOS ADOLFO PIÑA GARCÍA
Universidad Veracruzana
Xalapa, México

PhD BIRGER HJØRLAND
University of Copenhagen
Copenhagen, Dinamarca

DR. JOSÉ LÓPEZ YEPES
Universidad Complutense de Madrid
Madrid, España

DRA. GLORIA PÉREZ SALMERÓN
Stichting IFLA Foundation
La Haya, Holanda

DRA. FERNANDA RIBEIRO
University of Porto
Porto, Portugal

DR. ELÍAS SANZ CASADO
Universidad Carlos III
Madrid, España

COLABORADORES EN ESTE NÚMERO

INVESTIGACIÓN BIBLIOTECOLÓGICA, vol. 36, núm. 93, octubre/diciembre, 2022, México,
ISSN: 2448-8321

Adriana Mata-Puente

Facultad de Ciencias de la Información
Universidad Autónoma de San Luis Potosí
Av. Industrias 101, Fracc. Talleres,
C.P. 78494, San Luis Potosí, S.L.P.
México
Tel. + 52 44 4832 1000 ext. 9019
amata@uaslp.mx

Adriana Suárez Sánchez

Instituto de Investigaciones Bibliotecológicas
y de la Información
Universidad Nacional Autónoma de México
Circuito Escolar, Torre II de Humanidades,
piso 12, Ciudad Universitaria, Alcaldía
Coyoacán, C.P. 04510, Ciudad de México
México
Tel. +52 55 5623 0379
asuarez@iibi.unam.mx

Ángel Bravo Vinaja

Programa Innovación en el Manejo de Recursos
Naturales, Colegio de Postgraduados
Campus San Luis Potosí
Iturbide 73, 78620 Salinas de Hidalgo, S. L. P.
México
Tel.: +52 44 4178 1630
abravo@colpos.mx

Arturo Vargas Huerta

Posgrado en Bibliotecología y Estudios de la
Información, Coordinación General de Posgrado
Universidad Nacional Autónoma de México
Unidad de Posgrado, Edificio H, 1er. nivel,
H-104, Circuito de Posgrados,
Ciudad Universitaria, Coyoacán, C.P. 04510,
Ciudad de México, México
Tel.: +52 55 5623 7032
arvargash@gmail.com

Bexi Perdomo

Centro de Investigación de la Creatividad
Universidad de Ciencias y Artes de América
Latina
Av. El Sol de la Molina. La Molina. Lima
Perú
Tel.: +51 1622 2222
bjperdomo@crear.ucal.edu.pe

Eduardo Oliva-Cruz

Facultad de Ciencias de la Información
Universidad Autónoma de San Luis Potosí
Av. Industrias 101, Fracc. Talleres,
C.P. 78494, San Luis Potosí, S.L.P.
México
Tel.: +52 44 4832 1000, ext. 9019
eduardo.oliva@uaslp.mx

Elea Giménez-Toledo

Centro de Ciencias Humanas y Sociales
Consejo Superior de Investigaciones Científicas
(CSIC) de España
C. de Albasanz, 26, 28037 Madrid
España
Tel.: +34 6 3926 6257
elea.gimenez@cchs.csic.es

Esteban Giraldo-González

Dpto. de Biblioteconomía y Documentación
Universidad Carlos III de Madrid
Calle Madrid, 126. 28903 Getafe
España
Tel.: +57 30 0658 1426
esteban.giraldo@alumnos.uc3m.es

Georgina Sanabria Medina

1ra. Cda. de Parotas, Mz. 124, Lt. 11, 2da.
Ampliación Santiago Acahualtepec, Alcaldía
Iztapalapa, C.P. 09609, Ciudad de México
México
Tel.: +52 55 3486 9896
gesame@hotmail.com

Henrique Denes Hilgenberg Fernandes
Instituto Brasileiro de Informação para
Ciência e Tecnologia
SAS Quadra 5, Lote 6, Bloco H,
Ibict – Sobreloja
CEP 70070-912, Brasília – DF
Brasil
Tel.: +55 61 3217 6114
denes@ibict.br

Israel Polo Bautista
Escuela Nacional de Biblioteconomía
y Archivonomía
Instituto Politécnico Nacional, México
Calz. Ticomán 645, La laguna Ticomán
Gustavo A. Madero, 07340,
Ciudad de México, México
Tel.: +52 56 1934 9344
israelpolo001@gmail.com

Juan Felipe Córdoba-Restrepo
Editorial
Universidad del Rosario
Calle 12C N° 6-25 - Bogotá D.C.
Colombia
Tel.: +57 31 1480 5345
juan.cordoba@urosario.edu.co

Juan Miguel Palma Peña
Coordinación de Humanidades
Universidad Nacional Autónoma de México
Circuito Mario de la Cueva S/N, Ciudad
Universitaria, Alcaldía Coyoacán, C.P. 04510,
Ciudad de México
México
Tel.: +52 56 2004 9381
jemajumi@hotmail.com

Jayme Leiro Vilan Filho
Faculdade de Ciência da Informação
Universidade de Brasília
Campus Universitário Darcy Ribeiro
Edifício da Biblioteca Central, Entrada Leste.
CEP: 70910-900, Brasília - DF
Brasil
Tel.: +55 61 3107 2640
leiro@unb.br

Luis Roberto Polo Bautista
Biblioteca Gregorio Torres Quintero
Universidad Pedagógica Nacional
Carr. Picacho-Ajusco 24, Equipamiento
Periférico Picacho Ajusco Canal 13, Tlalpan,
14200, Ciudad de México,
México
Tel.: +52 55 7851 0088
rpolo@upn.mx

Oscar Alberto Morales
Facultad de Odontología
Universidad de Los Andes
Calle 24, entre Av. 2 y 3. Mérida
Venezuela
Tel.: +58 27 4240 2379
geode.ula@gmail.com

Perla Olivia Rodríguez Reséndiz
Instituto de Investigaciones Bibliotecológicas
y de la Información
Universidad Nacional Autónoma de México
Circuito Escolar, Torre II de Humanidades,
piso 12, Ciudad Universitaria, Alcaldía
Coyoacán, C.P. 04510, Ciudad de México,
México
Tel.: +52 55 1846 5485
perlaolivia@gmail.com

Ramsés Jabín Oviedo Pérez
Facultad de Ciencias de Información
Universidad Autónoma de San Luis Potosí
Av. Industrias 101, Fracc. Talleres,
C.P. 78494, San Luis Potosí, S.L.P.
México
Tel.: +52 41 3123 6946
a332982@alumnos.uaslp.mx

Experiencia de clasificación automática de documentos sobre Ciencias de la Vida y Biomedicina obtenidos del Web of Science

Luis Roberto Polo Bautista*
Israel Polo Bautista**

Artículo recibido:
7 de marzo de 2022

Artículo aceptado:
1 de julio de 2022

Artículo de investigación

RESUMEN

Este artículo brinda una experiencia en el desarrollo y prueba de un algoritmo de clasificación que asigna automáticamente un tema a un documento. Se utilizó el algoritmo de clasificación Multinomial de Naïve Bayes para analizar la correlación entre los temas de investigación en Ciencias de la Vida y Biomedicina, y los resúmenes de un *corpus* de 10 167 artículos recuperados del Web of Science (WoS). Se realizó una prueba del rendimiento del algoritmo aplicada a 5 581 resúmenes para medir su precisión, exhaustividad y exactitud. Los resultados demuestran su utilidad en la organización de la información respecto a la asignación automática

* Biblioteca Gregorio Torres Quintero, Universidad Pedagógica Nacional, México
rpolo@upn.mx

** Escuela Nacional de Biblioteconomía y Archivonomía, Instituto Politécnico Nacional, México
israelpolo001@gmail.com

de temas a documentos obtenidos de un repositorio digital o base de datos bibliográfica. El algoritmo propuesto puede ser utilizado como alternativa a los métodos tradicionales de clasificación de documentos en un área específica del conocimiento; esto permitirá la creación de servicios especializados orientados al desarrollo de sistemas computacionales que apoyen la gestión de información digital y electrónica.

Palabras clave: Algoritmos; Clasificación automática de documentos; Naïve Bayes Multinomial; Ciencias de la Vida y Biomedicina

Automatic classification experience of documents about Life Sciences and Biomedicine obtained in the Web of Science

Luis Roberto Polo Bautista and Israel Polo Bautista

ABSTRACT

This article provides an experience in the development and proof of a classification algorithm that automatically assigns a theme to a document. The Naïve Bayes Multinomial classification was used to automatically analyze the correlation between the themes of research in Life Sciences and Biomedicine, and the result of a corpus of 10 167 articles recuperated from the Web of Science (WoS). A proof of the performance of the algorithm was applied to 5 581 reviews for measuring its precision, exhaustivity and accuracy. The results show its usefulness in the organization of information respect to the automatic assignation of themes to the documents obtained in a digital repository or a bibliographic data base. The algorithm proposed can be utilized as an alternative to the traditional methods of classification of documents in a specific area of knowledge; this will allow the creation of specialized services oriented to the development of computational services that support the digital and electronic information management.

Keywords: Algorithms; Automatic classification of documents; Naïve Bayes Multinomial; Sciences of Life and Biomedicine

INTRODUCCIÓN

Dentro de la bibliotecología y ciencias de la información, la clasificación de documentos se ha visto como una tarea intelectual y compleja. Tiene como finalidad facilitar la recuperación de documentos dentro de las unidades y centros de información, basada en una localización física del material, representada por áreas temáticas y sistemas alfanuméricos de clasificación.

Tradicionalmente, esta tarea es efectuada de forma manual por bibliotecarios profesionales, dando como resultado palabras clave en forma de descriptores o encabezamientos de materia que son utilizados como puntos de acceso temáticos en la búsqueda de información (Contreras 2018, 112).

Clasificar los documentos de forma manual es una tarea que conlleva tiempo y esfuerzo adicional por parte de los profesionales de la información, considerando que en algunos casos deben conocer o estar familiarizados con el área de conocimiento de un documento, con el objetivo de identificar y comprender con mayor precisión su contenido (Polo Bautista, y Martínez Acevedo 2021, 16).

El crecimiento masivo de documentos en los medios digitales también dificulta el proceso de clasificarlos e indizarlos manualmente. Este es el problema en la llamada sobrecarga de información: más que referirse al problema relacionado con la gran cantidad de información que existe actualmente, éste apunta hacia la dificultad para acceder a ella y gestionarla de manera adecuada (Levy 2005, 283).

Para solucionar estas problemáticas, se han desarrollado métodos para organizar y clasificar estos documentos de manera automática (Aljedani, Alotaibi, y Taileb 2020, 694). Estos métodos y modelos computacionales son relevantes en una gran variedad de tareas de organización y gestión de la información (Alfaro, y Allende 2020, 551), con énfasis en el análisis automático de grandes datos textuales y la identificación de patrones lingüísticos asociados con las principales áreas temáticas de un documento.

Los métodos de clasificación automática de documentos pueden diversificar y mejorar los procesos de organización de la información dentro del área de bibliotecología y ciencias de la información, permitiendo procesar una amplia gama de formatos de documentos y creando servicios especializados orientados en la gestión de información digital y electrónica.

El algoritmo implementado en este trabajo puede servir como base para el desarrollo de nuevos métodos de clasificación automática de documentos dentro del área de bibliotecología y ciencias de la información, a partir de la utilización de grandes modelos de lenguaje. De este modo, el rendimiento de estos algoritmos de clasificación puede ser indistinguible al compararse con la misma tarea realizada por un bibliotecario profesional.

“En la actualidad existen diversas técnicas de aprendizaje automático que son utilizadas para clasificar documentos de forma automática, y cada una tiene características propias que permiten resolver diferentes problemas” (Silva 2021, 27). La clasificación es un tema fundamental en el aprendizaje automático; su objetivo es construir algoritmos a partir del entrenamiento de un conjunto de datos que permitan predecir la etiqueta o área temática que le corresponde a un documento (Zhang 2004, 562).

El algoritmo Naïve Bayes es una técnica de aprendizaje automático supervisado, basada en el teorema de Bayes y la teoría de la probabilidad [...]. La clasificación bayesiana es una herramienta estadística para categorizar un conjunto de datos, a través de la predicción de una etiqueta o categoría definida para un documento (Marikani, y Shyamala 2020, 296).

Este algoritmo ha sido aplicado en varios dominios del conocimiento en las últimas décadas debido a su estructura sencilla y su rendimiento. Su simplicidad surge fundamentalmente con la suposición de que cada par de atributos es condicionalmente independiente con respecto a la información de la clase o área temática correspondiente (Harzevili, y Alizadeh, 2018: 516).

El algoritmo utilizado en este trabajo es el Multinomial de Naïve Bayes o NB multinomial, un modelo de aprendizaje automático probabilístico. La utilización del algoritmo permitió el desarrollo del sistema de clasificación automática de documentos en el área de Ciencias de la Vida y Biomedicina. Los detalles asociados a la fundamentación matemática del algoritmo se describen en la sección 2.2.1 sobre la metodología.

Seleccionamos el área de Ciencias de la Vida y Biomedicina por dos razones principales: i) es un área de investigación poco utilizada en este tipo de aplicaciones computacionales; y ii) es el campo de conocimiento que tiene más categorías (76) dentro de las áreas de investigación del Web of Science (Clarivate Analytics 2020). De las 76 categorías que implementa esta área de investigación, tomamos en cuenta 21 (*Tabla 1*).

Categorías del área de investigación en Ciencias de la Vida y Biomedicina				
Biología del Desarrollo	Biología marina y de agua	Biotecnología y Microbiología Aplicada	Medicina General e Interna	Pesca
Biofísica	Medicina Integrativa y Complementaria	Biología Evolutiva	Ciencias Ambientales y Ecología	Medicina de Urgencias
Medicina Legal	Biodiversidad y Conservación	Geriatría Gerontología	Hematología	Enfermedades infecciosas

Silvicultura	Ciencias y servicios sanitarios	Entomología	Dermatología	Inmunología
Ciencias de la Vida y Biomedicina - Otros Temas				

Tabla 1. Categorías del área de investigación en Ciencias de la Vida y Biomedicina

Fuente. Adaptado de *Áreas de investigación*

(*Categorías / Clasificación*) por Clarivate Analytics, 2020

El resto de este trabajo está organizado de la siguiente manera: en la sección dos se describe la metodología utilizada; en la sección tres se presentan los resultados obtenidos; en la sección cuatro se presenta una discusión relacionada con el estado del arte; y en la sección cinco se muestran las conclusiones generales y los trabajos futuros.

METODOLOGÍA

Utilizamos el algoritmo de clasificación Naïve Bayes Multinomial para analizar la correlación entre los temas en el área de Ciencias de la Vida y Biomedicina, y los resúmenes de un *corpus* de 10 167 artículos recuperados del Web of Science. Realizamos una prueba de rendimiento del algoritmo aplicada a 5 581 resúmenes, para medir su precisión, exhaustividad y exactitud.

Para el desarrollo y la implementación del algoritmo seguimos los siguientes procedimientos: i) Recopilación del *corpus* (Tabla 2); ii) Configuración del programa (Tabla 3); y iii) Codificación del algoritmo (Tabla 4).

Etapa	Procedimiento
1	Compilación del <i>corpus</i> de entrenamiento. Conformado por 10 167 artículos sobre Ciencias de la Vida y Biomedicina obtenidos a través del Web of Science.
2	Compilación del <i>corpus</i> de prueba. Conformado por 5 581 artículos sobre Ciencias de la Vida y Biomedicina obtenidos a través del Web of Science.

Tabla 2. Recopilación del *corpus*

Etapa	Procedimiento
1	Creación de una cuenta de correo electrónico de Google ¹ . Permitirá tener acceso gratuito a los servicios de Google Drive ² y Google Colaboratory ³ .
2	Subir los <i>corpus</i> de artículos de investigación (entrenamiento y prueba) a Google Drive.
3	Comenzar un nuevo proyecto (<i>notebook</i>) en Google Colaboratory.

Tabla 3. Configuración del programa

Etapa	Procedimiento
1	Importar los <i>corpus</i> de entrenamiento ⁴ y prueba ⁵ a través de Google Drive.
2	Implementar los módulos <i>Pandas</i> para el análisis de datos (McKinney 2010; The pandas development team 2020) y <i>Matplotlib</i> para generar visualizaciones de los datos (Hunter 2007).
3	Implementar el módulo <i>Scikit-learn</i> para utilizar el algoritmo Naïve Bayes Multinomial y generar métricas de rendimiento (Pedregosa et al. 2011).
4	Implementar el módulo <i>Seaborn</i> para apoyar en la generación de visualizaciones de datos (Waskom 2021).
5	Exportar el archivo que se generó como resultado tras la ejecución del algoritmo. ⁶

Tabla 4. Codificación del algoritmo

Detalles del corpus

La parte inicial de este trabajo consistió en la compilación de un *corpus* de registros bibliográficos de artículos en el área de Ciencias de la Vida y Biomedicina en el periodo 1997-2017, a través del Web of Science. Se recuperaron 10 167 artículos correspondientes a 21 (Tabla 1) de las 76 categorías del área de investigación antes mencionada.

Este *corpus* de artículos científicos contiene como idioma hegemónico el inglés, seguido de otros idiomas, como español, alemán, ruso y portugués. Fue utilizado en forma de datos de entrenamiento para analizar la correlación

1 Se puede crear una cuenta gratuita a través del siguiente enlace: <https://www.google.com/intl/es-419/gmail/about/>

2 Plataforma electrónica de almacenamiento de archivos basado en la nube: <https://www.google.com/intl/es/drive/>

3 Plataforma electrónica para escribir y ejecutar código Python en un navegador, sin configuración requerida e intuitivo de utilizar: <https://colab.research.google.com/?hl=es>

4 Se puede consultar el *corpus* de entrenamiento a través del siguiente enlace: <https://drive.google.com/file/d/15GNseM6xV8U5SKl7kAeUu6BaaB05uabK/view?usp=sharing>

5 Se puede consultar el *corpus* de prueba a través del siguiente enlace: https://drive.google.com/file/d/1C5nYJ7M4DMmeYtRNWrxTJ2L8le_6xmef/view?usp=sharing

6 Se puede consultar el archivo que se generó como resultado a través del siguiente enlace: <https://drive.google.com/file/d/1fSYNNkw1OO6jn0V2153HJVvmgOlQ9LqN/view?usp=sharing>

entre los contenidos de los resúmenes y los temas o categorías del área de investigación en Ciencias de la Vida y Biomedicina.

La prueba del rendimiento del algoritmo se aplicó a un *corpus* distinto del anterior, conformado por 5 581 registros bibliográficos de esta misma área de investigación, en el periodo 2002-2016. Contiene como idioma principal el inglés y en menor medida el español, el alemán y el ruso. Fue utilizado como datos de prueba para medir el grado de precisión, exhaustividad y exactitud del algoritmo.

Algoritmo Naïve Bayes Multinomial

Fundamentación matemática

Según Manning, Raghavan y Schütze (2009, 258) el algoritmo Naïve Bayes Multinomial describe la probabilidad de que un documento d esté en la clase c , tal como se muestra en la *Fórmula 1*:

$$P(c|d) \propto P(c) \prod_{1 \leq k \leq n_d} P(t_k|c) \quad [1]$$

Donde $P(t_k|c)$ es la probabilidad condicional del término t_k que se produce en un documento de la clase c . Interpretamos $P(t_k|c)$ como medida de la cantidad de pruebas t_k que contribuye a c como la clase correcta. $P(c)$ es la probabilidad a priori de que un documento pertenezca a la clase c . Si los términos de un documento no proporcionan una evidencia clara de una clase o categoría frente a otra, se elige la que tiene mayor probabilidad a priori (Manning, Raghavan y Schütze 2009, 258).

Supongamos que $(t_1, t_2, \dots, t_{n_d})$ son los *tokens* de d que forman parte del vocabulario que utilizamos para la clasificación y n_d es el número de *tokens* en d . Por ejemplo, $(t_1, t_2, \dots, t_{n_d})$ para la oración *Pekín y Taipéi se unen a la OMC* puede ser (Pekín, Taipéi, unen, OMC), con $n_d = 4$, si tratamos los términos y el artículo “el” como palabras vacías (Manning, Raghavan y Schütze 2009, 258).

De acuerdo con Manning, Raghavan y Schütze (2009, 258), “en la clasificación de textos, la finalidad es encontrar la mejor clase o categoría para un documento. La mejor clase en la clasificación NB es la clase más probable o máxima a posteriori (MAP) C_{map} , que está representada por la *Fórmula 2*”.

$$\arg \max \hat{P}(c|d) = \arg \max \hat{P}(c) \prod_{1 \leq k \leq n_d} \hat{P}(t_k|c) \quad [2]$$

“Se sustituye P por P^{\wedge} porque no se conocen los verdaderos valores de los parámetros $P(c)$ y $P(t_k | c)$, sino que se estiman a partir del conjunto de entrenamiento” (Manning, Raghavan y Schütze 2009, 258).

Manning, Raghavan y Schütze (2009, 258) mencionan que en la ecuación anterior se multiplican muchas probabilidades condicionales, una para cada posición $1 \leq k \leq n_d$. Esto puede dar lugar a un desbordamiento de punto flotante. Por lo tanto, es mejor realizar el cálculo sumando logaritmos en lugar de multiplicar las probabilidades. La clase con la mayor *log* de probabilidades sigue siendo la más probable; y la función del logaritmo es monótona. Por lo tanto, la maximización que realmente se hace en la mayoría de las implementaciones de NB se muestra en la *Fórmula 3*:

$$c_{map} = \arg \max [\log \hat{P}(c) + \sum_{1 \leq k \leq n_d} \log \hat{P}(t_k | c)] \quad [3]$$

Cada parámetro condicional $\log \hat{P}(t_k | c)$ es un peso que indica lo bueno que es un indicador t_k para c . Del mismo modo, $\hat{P}(c)$ es un peso que indica la frecuencia relativa de c . Las clases más frecuentes tienen más probabilidades de ser la clase correcta que las infrecuentes. La suma de las ponderaciones *log prior* y *termal* es entonces una medida de cuánta evidencia hay de que el documento está en la clase, y la ecuación anterior selecciona la clase para la que tenemos más evidencia (Manning, Raghavan y Schütze 2009, 259).

De acuerdo con Manning, Raghavan y Schütze (2009, 259) “los parámetros $\hat{P}(c)$ y $\hat{P}(t_k | c)$ se calculan probando la estimación de máxima verosimilitud, que es simplemente la frecuencia relativa y corresponde al valor más probable de cada parámetro dados los datos de entrenamiento. Para los priores esta estimación se representa a través de la *Fórmula 4*.”

$$\hat{P}(c) = \frac{N_c}{N} \quad [4]$$

“Donde, N_c es el número de documentos de la clase c y N es el número total de documentos. Estimamos la probabilidad condicional $\hat{P}(t_k | c)$ como la frecuencia relativa de término t en los documentos pertenecientes a la clase c ” tal como se muestra en la *Fórmula 5* (Manning, Raghavan y Schütze 2009, 259).

$$\hat{P}(t_k | c) = \frac{T_{ct}}{\sum t' \in VT_{ct} t'} \quad [5]$$

“Donde, T_{ct} es el número de ocurrencias de t en los documentos de entrenamiento de c , incluyendo múltiples apariciones de un término en un documento. En este caso se realizó la suposición de independencia posicional,

que describe lo siguiente: T_{ct} es un recuento de ocurrencias en todas las posiciones k en los documentos del conjunto de entrenamiento” (Manning, Raghavan y Schütze 2009, 260).

Manning, Raghavan y Schütze (2009, 260) señalan que el problema de estas estimaciones radica en que el resultado es cero para una combinación de término-clase que no aparece en los datos de entrenamiento [...]. Para la eliminación de los ceros se utilizó el suavizado de Laplace, que simplemente añade uno a cada recuento, como se puede observar en la *Fórmula 6*:

$$\hat{P}(t|c) = \frac{T_{tc} + 1}{\sum_{t' \in V} (T'_{ct} + 1)} = \frac{T_{tc} + 1}{(\sum_{t' \in V} T'_{ct}) + B'} \quad (6)$$

“Donde, $B = |V|$ es el número de términos del vocabulario. El alisado de adición puede interpretarse como una prioridad uniforme (cada término aparece una vez para cada clase) que se actualiza a medida que llegan las pruebas de los datos de entrenamiento” (Manning, Raghavan y Schütze 2009, 260).

Estructura

El algoritmo utilizado en este trabajo es el Naïve Bayes Multinomial descrito por Manning, Raghavan y Schütze (2009, 260). Consta de dos secciones, i) entrenamiento (*Tabla 5*), y ii) prueba (*Tabla 6*).

Naïve Bayes Multinomial (Entrenamiento) (C,D)	
1	$V \leftarrow$ Extraer el vocabulario (D)
2	$N \leftarrow$ Contar los documentos (D)
3	Para cada $c \in C$
4	Hacer $N_c \leftarrow$ Contar docs. en las clases (D, c)
5	$priori [c] \leftarrow \frac{N_c}{N}$
6	$texto_c \leftarrow$ Concatenar el texto de todos los docs. en la clase (D, c)
7	Para cada $t \in V$
8	Hacer $T_{ct} \leftarrow$ Contar los tokens de ($texto_c, t$)
9	Para cada $t \in V$
10	Hacer $prob [t][c] \leftarrow \frac{T_{tc} + 1}{\sum_{t' \in V} (T'_{ct} + 1)}$
11	Devolver $V, priori, prob$

Tabla 5. Naïve Bayes Multinomial (Entrenamiento)
Fuente. Adaptado de Manning et al. (2009, 260)

Naïve Bayes Multinomial (Prueba) (C, V, priori, prob, d)	
1	$W \leftarrow$ Extraer tokens de docs. (V, d)
2	Para cada $c \in C$
3	Hacer score $[c] \leftarrow \log \text{priori} [c]$
4	Para cada $t \in V$
5	Hacer score $[c] += \log \text{prob}[t][c]$
6	Devolver $\arg \max_{c \in C} \text{score} [c]$

Tabla 6. Naïve Bayes Multinomial (Prueba)
Fuente: Adaptado de Manning et al. (2009, 260)

PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

El desarrollo del algoritmo de clasificación requirió la utilización de dos conjuntos de datos (entrenamiento y prueba). Cada uno de los artículos de ambos conjuntos de datos ya tenían asignados temas de acuerdo con el área de investigación antes mencionada. En las Figuras 1 y 2 se presenta la distribución de los 21 temas que consideramos para el desarrollo del algoritmo.

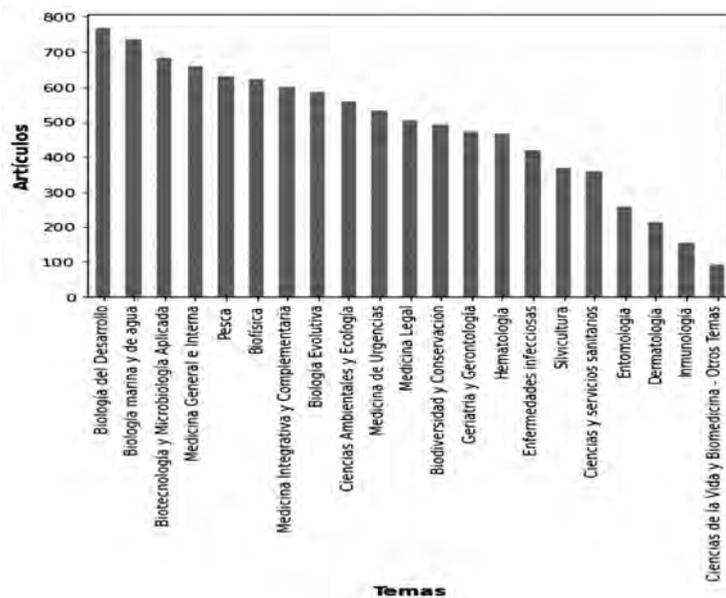


Figura 1. Distribución temática del corpus de entrenamiento

Los temas de la *Figura 1* representan datos de entrenamiento que permitieron que el algoritmo aprenda a predecir un tema determinado para cada artículo, con base en un cálculo probabilístico de la dispersión de palabras de los resúmenes. Los temas de la *Figura 2* son elementos de referencia que sirven como apoyo para medir el rendimiento del algoritmo, realizando comparaciones entre los temas asignados originalmente por el WoS y los temas calculados automáticamente.

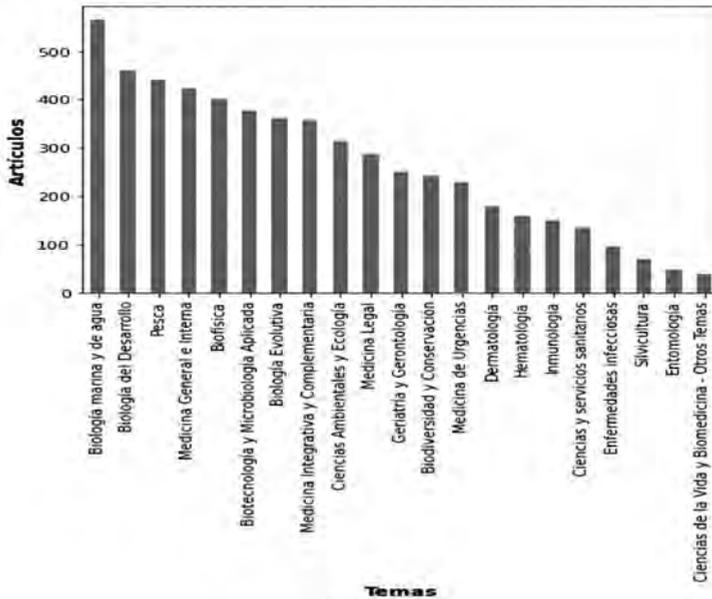


Figura 2. Distribución temática del *corpus* de prueba

Para evaluar el rendimiento general del algoritmo utilizamos las siguientes métricas:

Exactitud (Accuracy). De acuerdo con Arjaria, Rathore y Cherian (2021, 319), “El cálculo de la exactitud se utiliza para comparar la eficiencia del modelo. Tiene en cuenta el número total de predicciones correctas realizadas por el algoritmo. Se calcula como se muestra en la *Fórmula 7*.”

$$\text{Exactitud} = \frac{TP + TN}{TP + TN + FP + FN} \quad [7]$$

Exhaustividad (Recall). De acuerdo con Arjaria, Rathore y Cherian (2021, 319), “La exhaustividad se calcula tomando la proporción de entradas positivas

correctamente identificadas como positivas. Se calcula como se muestra en la *Fórmula 8*.”

$$\text{Exhaustividad} = \frac{TP}{TP + FN} \quad [8]$$

Precisión (Precision). De acuerdo con Arjaria, Rathore y Cherian (2021, 319), “La precisión es el número de casos positivos predichos correctamente por el algoritmo. Se calcula como se muestra en la *Fórmula 9*.”

$$\text{Precisión} = \frac{TP}{TP + FP} \quad [9]$$

Valor-F (F1 Score). De acuerdo con Arjaria, Rathore y Cherian (2021, 319), “Se define como una media ponderada de la precisión y la exhaustividad. Tiene su valor máximo en 1 y el peor en 0. Se calcula como se muestra en la *Fórmula 10*.”

$$\text{Valor F} = \frac{2 \times \text{Precisión} \times \text{exhaustividad}}{\text{Precisión} + \text{exhaustividad}} \quad [10]$$

- TP (Verdadero Positivo): Número de artículos que han sido correctamente clasificados.
- TN (Verdadero Negativo): Número de artículos correctamente clasificados que no pertenecen realmente al tema asignado.
- FP (Falso Positivo): Número de artículos clasificados erróneamente a un tema y que en realidad no es el tema que le corresponde.
- FN (Falso Negativo): Número de artículos clasificados erróneamente a un tema, pero que en realidad los artículos sí corresponden a ese tema.

Tomando lo anterior como base, en la *Tabla 7* se muestran los resultados del rendimiento del algoritmo, tras aplicarse en el *corpus* de prueba.

	Precisión	Exhaustividad	Valor-F
Biodiversidad y Conservación	0.98	0.38	0.54
Biofísica	0.95	0.95	0.95
Biología Evolutiva	0.83	0.93	0.88
Biología del Desarrollo	0.66	0.97	0.79
Biología marina y de agua	0.65	0.94	0.77
Biotecnología y Microbiología Aplicada	0.83	0.93	0.87
Ciencias Ambientales y Ecología	0.75	0.85	0.80

Ciencias de la Vida y Biomedicina - Otros Temas	0.00	0.00	0.00
Ciencias y servicios sanitarios	0.96	0.40	0.57
Dermatología	1.00	0.01	0.01
Enfermedades infecciosas	0.90	0.55	0.68
Entomología	0.00	0.00	0.00
Geriatría y Gerontología	0.93	0.80	0.86
Hematología	0.97	0.53	0.68
Inmunología	0.00	0.00	0.00
Medicina General e Interna	0.55	0.98	0.70
Medicina Integrativa y Complementaria	0.86	0.87	0.87
Medicina Legal	1.00	0.83	0.91
Medicina de Urgencias	0.91	0.74	0.82
Pesca	0.79	0.75	0.77
Silvicultura	1.00	0.27	0.43
Exactitud			0.77
Desempeño promedio	0.74	0.60	0.61
Media ponderada	0.78	0.77	0.73

Tabla 7. Resultados del rendimiento del algoritmo

En la *Tabla 4* se observan los cálculos realizados de cada uno de los artículos correspondientes a los 21 temas en el área de investigación en Ciencias de la Vida y Biomedicina, utilizando las métricas *Exactitud*, *Exhaustividad*, *Precisión* y *Valor-F*.

Los temas que tuvieron una precisión superior al 90% son Biodiversidad y Conservación (98%); Biofísica (95%); Ciencias y servicios sanitarios (96%); Dermatología (100%); Enfermedades infecciosas (90%); Medicina de Urgencias (91%), y Silvicultura (100%). Estos resultados se debieron a la cantidad de datos de entrenamiento que utilizamos para estos temas. Los temas que menor precisión tuvieron contaban con pocos datos de entrenamiento, como Ciencias de la Vida y Biomedicina - Otros Temas; Entomología, e Inmunología.

La métrica de exhaustividad indica la capacidad del algoritmo para identificar la probabilidad de que un tema determinado corresponda a un artículo, “de esta forma se evalúa la eficacia del algoritmo sólo en un tema” (Sokolova, Nathalie y Stan 2006, 3). Biofísica clasifica el 95% de artículos correctamente; Biología el 93%, Biología del Desarrollo 97%, entre otros.

La exactitud, “es una métrica que evalúa la eficacia global de un modelo de clasificación” (Sokolova, Nathalie y Stan 2006, 3). En este caso, el algoritmo propuesto tuvo un porcentaje de 77% de eficiencia. Considerando la diversidad de idiomas en los *corpus* de artículos, la extensión del contenido de los resúmenes y su falta de normalización, entre otras circunstancias, podemos estimar que el porcentaje es satisfactorio en cuanto a la clasificación de temas automática.

El algoritmo propuesto puede ser utilizado como alternativa a los métodos tradicionales de clasificación de documentos, en un dominio de conocimiento particular, permitiendo crear servicios especializados orientados al desarrollo de sistemas computacionales para utilizarse en la gestión de información digital y electrónica.

Para representar de forma visual el rendimiento del algoritmo se generó una matriz de confusión. Los elementos diagonales representan resultados correctamente clasificados. Los resultados mal clasificados están representados fuera de las diagonales de la matriz. El mejor algoritmo tendrá una matriz de confusión con sólo elementos diagonales y el resto de los elementos a cero (Arjaria, Rathore, y Cherian 2021, 318).

En la *Figura 3* se presenta la matriz de confusión del algoritmo propuesto.

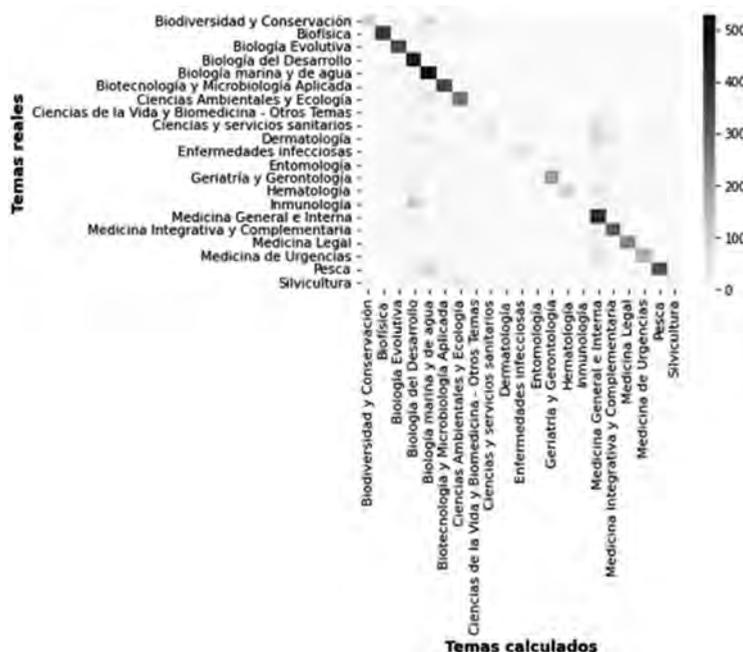


Figura 3. Matriz de confusión del algoritmo propuesto

En la *Figura 3* se observa una matriz que contiene en los ejes “X” e “Y”, los 21 temas en el área de Ciencias de la Vida y Biomedicina utilizados para desarrollar y evaluar el algoritmo. Los temas calculados representan aquellos que fueron generados automáticamente por el algoritmo, y los temas reales son los asignados originalmente por el Web of Science.

La matriz de confusión muestra que 16 de los 21 temas utilizados fueron clasificados correctamente, destacando los temas Medicina General e Interna; Biofísica; Biología Evolutiva; Biología del Desarrollo; Biología Marina y de Agua; Biotecnología y Microbiología Aplicada.

DISCUSIÓN

La clasificación de documentos automática generalmente se asocia con las ciencias de la computación, abarcando diversos enfoques como: Análisis de sentimientos, clasificación de imágenes, clasificación de textos estructurados y no estructurados, entre otros. Estos enfoques permiten la implementación de algoritmos en distintas áreas de conocimiento.

Los estudios más recientes sobre la clasificación automática de documentos utilizan modelos basados en Transformers⁷, como se muestra en Lehecka et al. (2020); Cai *et al.* (2020); Yu, Su y Luo (2019), y Liu, Wang y Ren (2021), mostrando resultados favorables en la clasificación de documentos con múltiples etiquetas.

Se han desarrollado diversas investigaciones sobre clasificación de textos utilizando modelos probabilísticos, como el Multinomial de Naïve Bayes. En el trabajo de Gao, Zeng y Yao (2019) se analizó la construcción y el mejoramiento del modelo Naïve Bayes, para la clasificación de documentos. Otro estudio similar es el propuesto por Marikani y Shyamala (2020), el cual se enfoca en la clasificación de documentos sobre predicción de enfermedades cardíacas.

En el trabajo de Chen *et al.* (2019), se utilizó una modificación del algoritmo Naïve Bayes que facilita la correlación general entre distintas clases o categorías. En el trabajo de Harzevili y Alizadeh (2018) se utilizó un clasificador Mixto Bayes Ingenuo Multinomial Latente para relajar el supuesto de independencia dentro de la clasificación de documentos.

Dentro de la bibliotecología y ciencias de la información, algunos de los trabajos más destacados son los siguientes: Contreras (2016) presentó un clasificador automático para documentos, basado en el área Z del Sistema de

7 Método que utiliza las redes neuronales para pre-entrenar un modelo de lenguaje que se utilizará en una tarea específica de procesamiento de lenguaje natural.

Clasificación de la Biblioteca del Congreso (L.C), teniendo resultados favorables en la clasificación del material bibliográfico.

Kragelj, Matjaž, y Mirjana (2021) desarrollaron un modelo de clasificación automática de documentos antiguos, basado en el Sistema de Clasificación Decimal Universal (UDC). Como parte final, en el trabajo de Cassidy (2020) se utilizó el modelo Naïve Bayes para la clasificación de patentes, considerando códigos especializados para estos recursos.

Como se puede observar, los estudios sobre clasificación de documentos en el área de las ciencias de la información utilizando modelos probabilísticos no son muy frecuentes. Es por ello por lo que este trabajo intenta utilizar este modelo, como alternativa a los métodos tradicionales de clasificación de documentos y a otras herramientas implementadas dentro de este campo, que permita facilitar la recuperación de información bibliográfica en línea, a través de biblioteca digitales, repositorios o bases de datos bibliográficas.

CONCLUSIONES Y TRABAJOS FUTUROS

A través de la utilización del clasificador Naïve Bayes Multinomial, presentamos una experiencia en el desarrollo y prueba de un algoritmo que asigna automáticamente un tema a un documento basándose en un cálculo probabilístico de la dispersión de las palabras de los resúmenes.

El funcionamiento general del algoritmo analiza los resúmenes de un conjunto de registros bibliográficos codificados en tablas en formato CSV (Valores separados por comas), y con base en cálculos probabilísticos asigna un tema del área de investigación en Ciencias de la Vida y Biomedicina a cada uno de los artículos, añadiéndolos a otra columna de la tabla.

La exactitud del algoritmo para la asignación de temas correspondientes a Ciencias de la Vida y Biomedicina es de 77%. Representa un rendimiento apropiado para considerar su utilización dentro de las unidades y centros de información. Esto reflejará un proceso de incorporación de las tecnologías de información y de aprendizaje automático supervisado.

La utilización del algoritmo implica un proceso de entrenamiento con datos de prueba sobre un área de conocimiento específico; de este modo, se puede diversificar su uso en otras disciplinas, permitiendo crear servicios especializados orientados al desarrollo de sistemas computacionales para utilizarse en la gestión de información digital y electrónica.

Tras la aplicación del algoritmo, la razón de que algunos temas se asignaron de forma correcta, en la mayoría de los casos se debió a la cantidad de datos de entrenamiento que incluimos en esas áreas, ya que el clasificador

logró comprender más sobre éstos, en comparación con los temas que tenían pocos datos de entrenamiento y, en su caso, la exactitud fue menor.

Como se mencionó anteriormente, el algoritmo asigna un tema a un documento con base en una distribución de las palabras de los resúmenes; de esta forma no afectaría de qué área del conocimiento se tratase, ni las brechas de idioma; sólo bastaría con establecer datos de entrenamiento suficientes para que el algoritmo tuviera un rendimiento similar o superior al presentado en este trabajo.

Utilizar este algoritmo no implica altos costos financieros, ya que se puede replicar en un entorno como Google Colaboratory, en el que se brindan recursos computacionales a través de sus servidores. De esta forma, el algoritmo propicia que el profesional de la información documental obtenga conocimientos y habilidades que le permitirán desarrollarse en la industria 4.0.

El algoritmo implementado en este trabajo puede servir como base para el desarrollo de nuevos métodos de clasificación automática de documentos dentro del área de bibliotecología y ciencias de la información, a partir de la utilización de grandes modelos de lenguaje basados en Transformes como BERT, GPT-3, etcétera, que aprovechen el aprendizaje profundo para minimizar la cantidad de datos de entrenamiento necesarios para el buen funcionamiento de un algoritmo de clasificación.

Estos métodos pre-entrenados pretenden ejercer un cambio de paradigma en los modelos de aprendizaje supervisado, modificándolos a aprendizaje auto-supervisado. De este modo, el rendimiento de estos algoritmos de clasificación puede ser indistinguible en comparación con la misma tarea realizada por un profesional de la información documental.

Como se mencionó anteriormente, uno de los objetivos principales de este trabajo es presentar una experiencia en el desarrollo de un algoritmo de clasificación, con la finalidad de promover este tipo de investigaciones dentro de nuestro gremio, para que se pueda no sólo desarrollar este tipo de algoritmos de clasificación, sino aprovecharse para el desarrollo de modelos inteligentes de creación de resúmenes automáticos, extracción de palabras clave, identificación y extracción automática de metadatos, implementación de chatbots, sistemas de recomendación, desarrollo de ontologías, implementación de análisis de satisfacción de los usuarios, o la creación de motores de búsqueda de última generación.

REFERENCIAS

- Alfaro, Rodrigo, y Héctor Allende. 2020. "Clasificación de Textos Multi-etiquetados con Modelo Bernoulli Multi-variado y Representación Dependiente de la Etiqueta". *Revista Signos* 53 (102): 549-567.
<https://doi.org/10.4067/S0718-09342020000300549>
- Aljedani, Nawal, Reem Alotaibi, y Mounira Taileb. 2020. "Multi-Label Arabic Text Classification: An Overview". *International Journal of Advanced Computer Science and Applications* 11 (10): 694-706.
<http://dx.doi.org/10.14569/IJACSA.2020.0111086>
- Arjaria, Siddhartha, Abhishek Singh Rathore, y Jincy Cherianc. 2021. "Kidney disease prediction using a machine learning approach: A comparative and comprehensive analysis". *Demystifying Big Data, Machine Learning, and Deep Learning for Healthcare Analytics*, editado por Pradeep, Sandeep Kautish, y Sheng-Lung Peng, 307-333. London: Elsevier.
<https://doi.org/10.1016/B978-0-12-821633-0.00006-4>
- Cai, Linkun, Yu Song, Tao Liu, y Kunli Zhang. 2020. "A Hybrid BERT Model That Incorporates Label Semantics via Adjustive Attention for Multi-Label Text Classification". *IEEE* 8.
<https://doi.org/10.1109/ACCESS.2020.3017382>
- Cassidy, Caitlin. 2020. "Parameter tuning Naïve Bayes for automatic patent classification". *World Patent Information* 61: 101968.
<https://doi.org/10.1016/j.wpi.2020.101968>
- Chen, Jiangning, Zhibo Dai, Juntao Duan, Heinrich Matzinger, y Ionel Popescu. 2019. "Improved Naive Bayes with optimal correlation factor for text classification". *SN Applied Sciences* 1 (9).
<https://doi.org/10.1007/s42452-019-1153-5>
- Clarivate Analytics. 2020. "Áreas de investigación (Categorías / Clasificación)". http://images.webofknowledge.com/WOKRS522_2R1/help/es_LA/WOS/hp_research_areas_easca.html
- Contreras Barrera, Marcial. 2018. "Aplicación del algoritmo RAKE en la indización de documentos digitales". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (75): 109-123.
- _____. 2016. "Minería de texto en la clasificación de material bibliográfico". *Biblios Journal of Librarianship and Information Science* 64.
<https://doi.org/10.5195/biblios.2016.309>
- Gao, Hongyi, Xi Zeng, y Chunhua Yao. 2019. "Application of improved distributed naive Bayesian algorithms in text classification". *Journal of Supercomputing* 7 (9): 5831-5847.
<https://doi.org/10.1007/s11227-019-02862-1>
- Harzevili, Nima y Sasan Alizadeh. 2018. "Mixture of latent multinomial naive Bayes classifier". *Applied Soft Computing* 69: 516-527.
<https://doi.org/10.1016/j.asoc.2018.04.020>
- Hunter, John. 2007. "Matplotlib: A 2D graphics environment". *Computing in Science & Engineering* 9 (3): 90-95.
<https://doi.org/10.1109/MCSE.2007.55>

- Kragelj, Matjaž, y Mirjana Kljaji Borštnar. 2021. "Automatic classification of older electronic texts into the Universal Decimal Classification–UDC". *Journal of Documentation* 77 (3):755–776.
<https://doi.org/10.1108/JD-06-2020-0092>
- Lehecka, Jan, Jan Švec, Pavel Ircing, y Luboš Šmídl. 2020. "Adjusting BERT's Pooling Layer for Large-Scale Multi-Label Text Classification". En *Text, Speech, and Dialogue (TSD)*, editado por Petr Sojka, Ivan Kopeček, Karel Pala, y Aleš Horák. Brno, Czech Republic: Springer, Cham.
https://doi.org/10.1007/978-3-030-58323-1_23
- Levy, David M. 2005. "To grow in wisdom: vannevar bush, information overload, and the life of leisure". Trabajo presentado en Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL, 05), Denver, CO, USA, 7-11 de junio. doi: 10.1145/1065385.1065450.
- Liu, Naiyin, Qianlong Wang, y Jiangtao Ren. 2021. "Label-Embedding Bi-directional Attentive Model for Multi-label Text Classification". *Neural Process Lett* 53: 375-389.
<https://doi.org/10.1007/s11063-020-10411-8>
- Manning, Christopher, Prabhakar Raghavan, y Hinrich Schütze. 2009. "Text classification and Naive Bayes". En *Introduction to Information Retrieval*, 253-287. Cambridge University Press.
- Marikani, y Shyamala. 2020. "Modified Multinomial Naive Bayes Algorithm for Heart Disease Prediction". En *Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, editado por S. Balaji, Yi-Nan Chung, y Álvaro Rocha, 294-300, Suiza: Springer, Cham.
https://doi.org/10.1007/978-3-030-28364-3_27
- McKinney, Wes. 2010. "Data structures for statistical computing in python". Trabajo presentado en Proceedings of the 9th Python in Science Conference, Austin, Texas, junio 28 a julio 3.
<https://doi.org/10.25080/Majora-92bf1922-00a>
- Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 12: 2825-2830.
- Polo Bautista, Luis Roberto, y Karen Vanessa Martínez Acevedo. 2021. "Algoritmo para el análisis temático de documentos digitales". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 35 (89): 13-31.
<http://dx.doi.org/10.22201/iibi.24488321xe.2021.89.58419>
- Silva Palacios, Daniel Andrés. 2021. "Clasificación Jerárquica Multiclase". Tesis doctoral, Universitat Politècnica de València, Departamento de Sistemas Informáticos y Computación.
- Sokolova, Marina, Nathalie Japkowicz, y Stan Szpakowicz. 2006. "Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation". En *AI 2006: Avances en Inteligencia Artificial*, editado por Abdul Sattar, y Byeong-ho Kang. Hobart, Australia: Springer.
https://doi.org/10.1007/11941439_114
- The pandas development team. 2022. "pandas-dev/pandas: Pandas". *Zenodo*.
<https://doi.org/10.5281/zenodo.6053272>

- Waskom, Michael. 2021. "Seaborn: statistical data visualization". *Journal of Open Source Software* 6 (60): 3021.
<https://doi.org/10.21105/joss.03021>
- Yu, Shanshan, Jindian Su, y Da Luo. 2019. "Improving BERT-Based Text Classification with Auxiliary Sentence and Domain Knowledge". *IEEE* 7: 176600 - 176612.
<https://doi.org/10.1109/ACCESS.2019.2953990>
- Zhang, Harry. 2004. "The Optimality of Naive Bayes". Trabajo presentado en Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, Miami Beach, Florida, 12 a 14 de mayo.

Para citar este texto:

- Polo Bautista, Luis Roberto, e Israel Polo Bautista. 2022. "Experiencia de clasificación automática de documentos sobre Ciencias de la Vida y Biomedicina obtenidos del Web of Science". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 13-32.
<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58607>

Anexo

Visualización y descarga del algoritmo

El algoritmo de clasificación temática presentado en este trabajo se puede visualizar y descargar completo por medio del siguiente enlace:

https://colab.research.google.com/drive/1E0utmZrL4H_g9QAlcRq25YG0qk8bN8q0?usp=sharing

Modelo de fluxo de informação científica para as áreas de informação

Henrique Denes Hilgenberg Fernandes*
Jayme Leiro Vilan Filho**

Artículo recibido:

14 de abril de 2022

Artículo aceptado:

29 de junio de 2022

Artículo de investigación

RESUMO

Esse artigo propõe um modelo para o fluxo da informação científica para as áreas de informação no Brasil (Arquivologia, Biblioteconomia, Ciência da Informação e Museologia). Foi concebido a partir de dados quantitativos extraídos de uma amostra aleatória de 266 projetos de pesquisa da Plataforma Lattes, desenvolvidos entre 2005 e 2019, por pesquisadores doutores e atuantes nas áreas de informação. A pesquisa considerou 1 298 publicações entre resumos, trabalhos completos, artigos de periódicos, capítulos de livros, livros, dissertações e teses relacionados aos projetos de pesquisa da amostra. Uma etapa qualitativa complementar

* Instituto Brasileiro de Informação para Ciência e Tecnologia, Brasília – DF, Brasil
denes@ibict.br

** Faculdade de Ciência da Informação, Universidade de Brasília, Brasília – DF, Brasil
leiro@unb.br

entrevistó 59 investigadores acerca de sus preferencias por tipo de documento para publicar. El resultado es un modelo apto a representar el uso de canales informales y formales, con los porcentajes de publicación de cada tipo de documento estudiado, una línea de tiempo indicando el período en que esos documentos son típicamente publicados y las razones para publicar en cada vehículo considerado. Un modelo gráfico resume todas esas informaciones. La importancia de los programas de pos-graduación en la investigación brasileña en las áreas de información también fue evaluada.

Palabras-clave: Comunicación científica; Modelos de flujo de información; Proyectos de investigación; Áreas de información

Modelo de flujo de información científica para las áreas de información

Henrique Denes Hilgenberg Fernandes e Jayme Leiro Vilan Filbo

RESUMEN

Este artículo tiene como objetivo proponer un modelo para el flujo de información científica en las áreas de información en Brasil (Archivología, Biblioteconomía, Ciencias de la Información y Museología). Se concibió a partir de datos cuantitativos extraídos de una muestra aleatoria de 266 proyectos de investigación en la base de datos de la Plataforma Lattes, que fueron desarrollados entre 2005 y 2019 por investigadores doctores que actúan en las áreas de información. Esta investigación consideró 1 298 publicaciones entre trabajos de resúmenes, artículos de conferencias, artículos de revistas, capítulos de libros, libros, disertaciones y tesis doctorales relacionadas con los proyectos de investigación de la muestra. Una investigación cualitativa complementaria entrevistó a 59 investigadores sobre sus preferencias al elegir un documento para publicar. El resultado es un modelo capaz de representar: el uso de canales informales y formales, con porcentajes de publicación para cada tipo de documento estudiado, una línea de tiempo con indicación de cuándo se publican típicamente estos documentos y las razones para publicarlo en cada vehículo considerado. Un modelo gráfico resume toda esta

información. También se evalúa la importancia de los programas de estudios de posgrado en las áreas de información en la investigación brasileña.

Palabras clave: Comunicación científica; Modelo de flujo de información; Proyecto de investigación; Áreas de información

Scientific information flux model for information areas

Henrique Denes Hilgenberg Fernandes e Jayme Leiro Vilan Filho

ABSTRACT

The objective of this article is to propose a model for the flux of scientific information in the areas of information in Brazil (Archival Science, Library Science, Information Sciences and Museology). It was conceived departing from quantitative data extracted from a random of 266 research projects in the data base of the Platform Lattes, which were developed between 2005 and 2019 for research doctors that act in the areas of information. This research considered 1 298 publications among resume works, conference articles, magazine articles, book chapters, books, dissertations, and doctoral thesis related with the research projects of the sample. Qualitative complementary research interviewed 59 researchers over their preferences for choosing a document for being published. The result is a model capable of representing: the use of informal and formal channels, with percentages of publication for each type of document studied, a line of time with an indication of when are typically published these documents and the reasons for publishing it in each considered vehicle. A graphic model resumes all this information. It is also evaluated the importance of the postgraduate program studies in the areas of information in the Brazilian research.

Keywords: Scientific communication; Information flow model; Research project; Information areas

INTRODUÇÃO

O estudo da comunicação científica frequentemente faz uso de modelos conceituais para representar o seu fluxo da informação, os quais são muito úteis na investigação dos fenômenos e processos envolvidos. Amplamente empregados na fundamentação teórica de diversos trabalhos, se aplicam ao monitoramento e previsibilidade da produção científica, servindo ainda como subsídio à definição de políticas de incentivo à pesquisa, assim como à tomada de decisões por parte de gestores e pesquisadores.

No caso do estudo da comunidade das áreas de informação no Brasil, que são aqui definidas como Arquivologia, Biblioteconomia, Ciência da Informação e Museologia, nenhum dos modelos disseminados na literatura científica, coberta pelos principais serviços de indexação, se adequa a representá-las, omitindo fenômenos peculiares às áreas e ao país. Um modelo aplicado às áreas, próprio e abrangente com relação às suas características, traz previsibilidade ao estudo da produção científica, podendo descrever um comportamento esperado, estabelecendo um padrão para as comunicações científicas e auxiliando o planejamento de cenários futuros. Visando a cobertura suficientemente ampla das investigações científicas brasileiras nas áreas de informação, propõe-se aqui um modelo para o fluxo da informação científica desenvolvido a partir da análise dos dados de projetos de pesquisa dessas áreas, executados entre 2005 e 2019, e das publicações a eles vinculadas.

Para Meadows (1999), o fluxo da informação ilustra quem consulta quem, destacando dessa forma os *gatekeepers* e as fontes de informação que esses atores têm acesso. Um modelo deve representar como se dá a transferência da informação entre os envolvidos no processo. Mueller (2000), no entanto, define o fluxo como a trajetória das sucessivas comunicações relacionadas a um projeto específico e representa o caminho percorrido pela pesquisa, desde que nasce uma ideia na mente de um pesquisador, passando pela publicação formal dos resultados (geralmente um artigo de periódico), e continua até que a informação sobre esse artigo possa ser recuperada na literatura secundária ou apareça como citação em outros trabalhos.

O trabalho seminal de Garvey e Griffith (1972), propôs um modelo para o fluxo da informação científica, a partir de dados de pesquisas dos associados da APA (*American Psychological Association*). Sob a forma de uma linha do tempo, esse modelo descreve os períodos em que as comunicações informais e formais tipicamente ocorrem, incluindo ainda a primeira citação por outro autor. Na mesma época, foi concebido também o modelo UNISIST (United Nations International Scientific Information System) (UNESCO, 1971), que representa o caminho da informação desde o seu produtor até o consumidor final,

destacando os papéis dos agentes envolvidos. Lancaster (1975) descreve o fluxo para um projeto de pesquisa hipotético, incluindo a sua formalização no sistema SSIE (*Smithsonian Science Information Exchange*).

Nas décadas que se seguem, Lievrouw (1992) propõe um Ciclo para a Comunicação Científica, composto pelas fases de Concepção, Documentação e Popularização dos resultados das pesquisas e Hurd (1996, 2000) apresenta proposições para lidar com o fluxo num panorama de muitas mudanças tecnológicas. Costa (2000) propõe um modelo híbrido, parte digital e parte impresso, uma vez que as previsões de Hurd (2000) pareciam estar distantes à época de sua publicação. Søndergaard, Andersen e Hjørland (2003) propõem uma atualização ao UNISIST com as mudanças tecnológicas e Bjork (2007) constrói um modelo bastante completo com uma metodologia para modelagem de processos.

Entre os mais recentes modelos propostos na literatura estão Khosrowjerdi (2011), Shehata, Ellis e Foster (2015) e Pinto e Costa (2018), enquanto Fernandes e Vilan Filho (2021) trazem uma ampla revisão acerca de modelos para o fluxo da informação científica. No Brasil, local da comunidade a ser estudada, identificaram-se os modelos propostos por Correia e Rocha (2008) para as grandes áreas do conhecimento e um modelo proposto por Melo (2014) para as áreas de informação no Brasil, porém este limitado às pesquisas de doutorado.

METODOLOGIA

A pesquisa usa métodos mistos, adotando a estratégia explanatória sequencial (Creswell, 2010) para a investigação, com uma fase quantitativa principal, seguida por uma fase qualitativa complementar. Na fase quantitativa, a técnica de pesquisa utilizada foi a bibliometria. Na fase qualitativa, por sua vez, foi utilizado o *survey*.

Como fonte de dados quantitativos, foram utilizados os projetos de pesquisa dos currículos da Plataforma Lattes (<http://lattes.cnpq.br>), que reúne informações sobre os pesquisadores brasileiros, onde cada projeto cadastrado inclui nome, descrição, anos de início e conclusão, a produção científica, tecnológica e orientações vinculadas a ele. Uma vantagem em se utilizar esses dados é que os projetos são cadastrados pelos próprios pesquisadores, que indicam quais itens de sua produção estão a eles vinculados. O experimento mostrou que pouco mais de um terço dos pesquisadores atribuiu a produção de sua autoria aos seus projetos de pesquisa.

Das fontes consideradas, o Lattes mostrou-se a mais adequada por compreender diversos tipos de publicações, em veículos nacionais e estrangeiros, sendo muito abrangente no Brasil por cobrir todas as áreas do conhecimento e seu uso ser praticamente obrigatório para se obter fomento, concorrer a posições de pesquisa e docência, bolsas, assim como pelo fato de que é o principal parâmetro para a avaliação da produtividade dos pesquisadores brasileiros. Em uso desde 1999, Vilan Filho, Maricato e Silva (2019) apuraram que o Lattes foi a fonte de informação brasileira mais usada para estudos métricos publicados em periódicos brasileiros das áreas de informação.

De início, a partir de uma relação dos pesquisadores doutores cadastrados na Plataforma Lattes, os dados foram filtrados de forma a conter apenas os currículos com data de atualização posterior a 2019, Grande Área de atuação informada como Ciências Sociais Aplicadas e Área de Atuação informada como Ciência da Informação¹ ou Museologia. Dos resultados, obteve-se 892 pesquisadores considerados atuantes nas áreas. No Lattes, cada pesquisador possui um identificador único a ele associado. Esses identificadores foram utilizados para recuperar os currículos com o uso do programa *scriptLattes* (Mena-Chalco e César Jr., 2009). O programa recuperou 2 426 projetos de pesquisa constantes dos currículos desses pesquisadores que foram iniciados e concluídos entre 2005 e 2019. Desses, 817 projetos possuíam a indicação da produção vinculada e 504 deles foram considerados elegíveis para a pesquisa, pois projetos com apenas uma publicação ou orientação não foram utilizados (se a única publicação não pudesse ser aproveitada, o projeto seria excluído da amostra, reduzindo a representatividade desta). Assim, com o uso do SPSS (*Statistical Package for Social Sciences*), foi selecionada uma amostra aleatória de 266 projetos, sendo analisadas 1 298 publicações.

Os dados coletados na amostra foram migrados para o Microsoft Excel, onde a etapa de análise foi, então, executada. Identificaram-se os percentuais de uso dos canais resumos, trabalhos completos, artigos de periódicos, capítulos de livros, livros, dissertações e teses. Livros e capítulos de livros, em sua maioria, apresentaram-se no formato impresso, ao passo que as demais publicações são majoritariamente digitais. Também foi descrita a cronologia da produção bibliográfica das áreas de informação, sob a forma de uma linha do tempo, demonstrando os períodos em que tipicamente são publicados cada um desses tipos documentais. Foram também estudadas as características das pesquisas realizadas nos programas de pós-graduação.

1 Na tabela de áreas da Capes (Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) as áreas de Arquivologia e Biblioteconomia estão contidas na Ciência da Informação.

De posse dos resultados quantitativos, iniciou-se a fase qualitativa da pesquisa, complementando os achados da fase anterior. Foram escolhidos dentro da amostra, e de forma intencional, 100 projetos cujos padrões de publicação sugerissem uma clara preferência por determinado tipo de documento, quando comparado ao restante da produção do próprio projeto ou ainda com os padrões dos demais projetos. Para cada um dos projetos selecionados, foi enviada uma entrevista por e-mail, questionando as razões da opção pelo tipo de documento preponderante de cada projeto. Foram respondidas 59 entrevistas, que foram analisadas através da análise de conteúdo e da categorização das respostas.

A presente pesquisa está limitada aos projetos de pesquisa registrados no Currículo Lattes e que possuíam a indicação da produção vinculada a eles. A amostra foi calculada considerando-se uma margem de erro de 5% e um nível de confiança de 95%.

RESULTADOS

Nessa seção, serão primeiramente apresentados os resultados da fase quantitativa da pesquisa, seguidos dos resultados da fase qualitativa e, por fim, a integração das duas fases, onde é apresentado o modelo de fluxo científico para as áreas de informação.

Resultados quantitativos

Na fase quantitativa da pesquisa, apurou-se que o tempo de vida médio dos projetos da amostra ficou em três anos e sete meses, com uma média de 4,88 publicações por projeto. Um percentual de 40,6% dos projetos possui ao menos uma tese ou uma dissertação entre as suas publicações vinculadas, fato que os caracteriza como projeto que foi desenvolvido com a participação da pós-graduação. Observou-se, ainda, que 82% de toda a produção está concentrada entre o primeiro e o quinto ano a partir do início do projeto, registrando-se o pico de produção no segundo ano.

A *Tabela 1* mostra a utilização dos canais pelos pesquisadores, atestando a preferência por trabalhos completos em eventos (26% da produção) e artigos de periódicos (24%), seguido por dissertações, capítulos de livros, resumos, teses e livros. O livro foi o tipo documental menos frequente na amostra, com apenas 40 ocorrências, sendo que 18 deles se tratam de livros organizados. A *Tabela 1* mostra os valores absolutos, percentuais e as médias por projeto para cada tipo de documento.

	Trab.	Art.	Dissert.	Cap.	Res.	Teses	Livros	Total
Total	338	308	211	171	140	90	40	1 298
Percentual (%)	26%	24%	16%	13%	11%	7%	3%	100%
Média por projeto	1,27	1,16	0,79	0,64	0,53	0,34	0,15	4,88

Tabela 1. Publicações por tipo documental

Fonte: Dados da pesquisa. Onde: Trab. - Trabalhos completos; Art. – Artigos de periódicos; Dissert. - Dissertações; Cap. - Capítulos de livros e Res. - Resumos

A Tabela 2, por sua vez, agrupa os volumes de publicação por tipo de documento (incluindo a totalização ao final) e, ao mesmo tempo, por ano de publicação. O início do projeto está identificado como ano 1. É possível observar que há publicações anteriores ao início do projeto, a partir do sexto ano que antecede o início. O volume de publicações anteriores ao início do projeto chega a 5,7% e abrange todos os tipos documentais. No entanto, a produção somente se torna significativa a partir do primeiro ano, atingindo o pico no segundo ano. No terceiro ano, inicia uma trajetória de queda, acentuando-se no quarto ano e tornando-se residual a partir do sexto ano. Ao décimo ano, a produção encerra-se completamente.

Ano	Trab.	Artigos	Dissert.	Cap.	Res.	Teses	Livros	Totais	%	% Ac
-6	1	0	1	0	0	0	0	2	0,15%	0,15%
-5	0	0	1	0	1	0	0	2	0,15%	0,31%
-4	2	1	1	0	0	0	0	4	0,31%	0,62%
-3	1	2	3	1	3	1	0	11	0,85%	1,46%
-2	5	3	4	2	5	1	0	20	1,54%	3,00%
-1	8	10	4	3	4	5	1	35	2,70%	5,70%
1	88	73	34	49	31	15	9	299	23,04%	28,74%
2	96	86	50	44	50	20	9	355	27,35%	56,09%
3	74	62	62	30	29	18	8	283	21,80%	77,89%
4	30	38	26	15	7	7	5	128	9,86%	87,75%
5	14	13	17	14	10	11	4	83	6,39%	94,14%
6	8	6	3	3	0	5	1	26	2,00%	96,15%
7	2	9	2	5	0	2	3	23	1,77%	97,92%
8	7	4	1	3	0	2	0	17	1,31%	99,23%
9	2	1	2	2	0	1	0	8	0,62%	99,85%
10	0	0	0	0	0	2	0	2	0,15%	100,00%
Total	338	308	211	171	140	90	40	1298	100,00%	-

Tabela 2. Produção vinculada aos projetos em relação ao ano de início

Fonte: Dados da pesquisa. Onde: Trab. - Trabalhos completos; Dissert. - Dissertações; Cap. - Capítulos de livros e Res. - Resumos

O gráfico da *Figura 1*, sob a forma de uma linha do tempo, ilustra os períodos em que a produção iniciou e cessou para cada tipo documental. A linha vertical delimita o início oficial do projeto e, para cada documento, o núcleo entre as linhas pontilhadas representa o período onde está concentrada 75% da produção. A área hachurada representa o ano em que foi atingida a mediana da produção de cada documento, representando o ano em que são tipicamente produzidos. Assim, conclui-se que os resumos, trabalhos completos, artigos de periódicos e capítulos de livros são tipicamente publicados no segundo ano após o início do projeto, ao passo em que aquelas publicações em formato monográfico, no caso livros, teses e dissertações são geralmente publicadas no terceiro ano. O valor ao lado da legenda representa o volume percentual de cada documento em relação à produção total, considerando todo o período analisado.

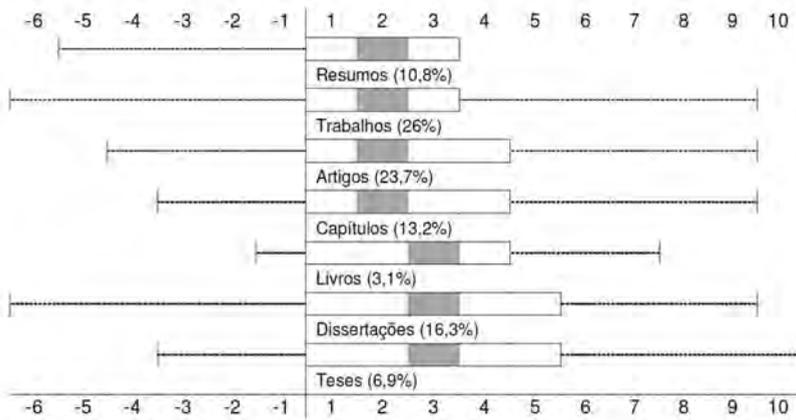


Figura 1. Cronologia da produção bibliográfica
 Fonte: Dados da pesquisa

A fim de estudar a participação da pós-graduação na pesquisa das áreas de informação no Brasil, todos os projetos com ao menos uma tese ou uma dissertação (40,6% dos projetos) foram copiados para um grupo distinto e comparados com o total da amostra. A *Tabela 3* traz esse comparativo, onde é possível ver números desses dois grupos. Tendo em vista que o volume de teses e dissertações é o mesmo nos dois grupos analisados, para se evitar a introdução de viés, o instrumento utilizado para comparação foi a média de publicações (para cada tipo de documento) por projeto. Os valores absolutos constam da tabela apenas a título ilustrativo, assim como os percentuais em relação aos valores do primeiro grupo (na quarta coluna).

A comparação mostra que os projetos com teses e dissertações possuem mais publicações por projeto do que aqueles do total da amostra, fato que ocorreu para todos os tipos de documentos, com exceção dos livros. Esse fenômeno sugere que os projetos com teses e dissertações são mais produtivos que os demais, pois 40,6% dos projetos reúnem 68,3% das publicações. Ou seja, das 1 298 publicações, 886 são encontradas nos projetos da pós-graduação. A média geral dos projetos com teses e dissertações (8,2 publicações/projeto) é significativamente superior à média geral do total da amostra (4,88 publicações/projeto)

	Total da amostra		Projetos com teses e dissertações	
	Total	Média por projeto	Total	Média por projeto
Dissertações	211	0,79	211 (100%)	1,95
Trabalhos	338	1,27	210 (62,1%)	1,94
Artigos	308	1,16	179 (58,1%)	1,66
Capítulos	171	0,64	111 (64,9%)	1,03
Teses	90	0,34	90 (100%)	0,83
Resumos	140	0,53	73 (52,1%)	0,68
Livros	40	0,15	12 (30,0%)	0,11
Total	1 298	4,88	886 (68,3%)	8,2

Tabela 3. Participação da pós-graduação na pesquisa (comparativo)

Fonte: Dados da pesquisa

Resultados qualitativos

Na pesquisa qualitativa, questionou-se aos pesquisadores as razões que os levaram a optar por determinado tipo de documento, com exceção de dissertações e teses que não foram avaliadas porque a razão para publicá-las é evidente: constituem requisitos legais para a obtenção dos graus de mestre e doutor. No caso dos resumos e trabalhos completos, ambos são publicados em anais de eventos, sendo o primeiro sob a forma resumida. Por essa razão, eles compartilham algumas características, de forma que algumas das razões apresentadas para se publicar em eventos (seja sob a forma resumida ou completa) são: 1) a apresentação de resultados parciais de um trabalho que ainda está em desenvolvimento e, portanto, sujeito à mudança; 2) a submissão aos pares, com o intuito de obter *feedback* e, assim, aprimorar a pesquisa; 3) a interação com outros pesquisadores e 4) a participação de estudantes de iniciação científica e de concluintes de cursos de graduação e mestrado.

Aqueles que publicaram especificamente sob a forma de resumo alegaram: 1) para que a pesquisa possa ser posteriormente publicada como artigo de periódico, uma vez que as políticas editoriais de muitas revistas brasileiras não permitem que sejam avaliados manuscritos que tenham sido previamente publicados como trabalhos completos em anais de eventos, ainda que publicados parcialmente; 2) o evento aceita somente resumos; 3) o evento não aceitava trabalhos completos de autor que não fosse doutor ou doutorando e 4) exigência do programa de bolsas acadêmicas. Já os que publicaram como trabalhos completos declararam que assim o fizeram a fim de fomentar discussões, divulgar a pesquisa ou por ser mais fácil de publicar. Os pesquisadores que optaram por publicar como artigos de periódicos alegaram que este tipo documental confere uma maior visibilidade ao trabalho, que goza de uma melhor reputação em comparação aos demais tipos, e que é o veículo que confere maior pontuação por ocasião da avaliação dos programas de pós-graduação.

Alguns pesquisadores que publicaram tanto sob a forma de capítulo em um livro organizado, quanto sob a forma de livro autoral alegaram que os livros atingem um público mais amplo que o acadêmico, que possuem maior liberdade para apresentar sua pesquisa e redigir o seu texto, e também alegaram terem sido convidados a publicar pelos editores ou organizadores. A motivação alegada para se publicar especificamente como capítulos de livros foi: 1) foi um trabalho em evento, cujos anais foram publicados sob o formato de livro; 2) o capítulo de livro é a forma preferencial de publicação na área de atuação do pesquisador (dentro das áreas de informação); 3) oportunidade de se publicar nesse formato e 4) fomento para a publicação de livros organizados. Por fim, aqueles que publicaram especificamente como livros alegaram: 1) a carência de publicações relativas ao assunto; 2) que os livros são a publicação mais adequada ao ensino de graduação; e 3) que os livros são o melhor veículo para o conteúdo que está sedimentado ao conhecimento científico.

Integração dos resultados

O modelo para o fluxo da informação científica dos projetos de pesquisa das áreas de informação no Brasil, mostrado na *Figura 2*, foi obtido a partir da integração dos resultados quantitativos e qualitativos. Nesse modelo, no primeiro ano o projeto é iniciado e são realizadas as atividades de revisão bibliográfica, desenvolvimento e testes da metodologia, coleta e análise de dados. No segundo ano, iniciam-se as publicações dos resultados, primeiramente sob os canais informais, com a publicação de trabalhos completos e resumos. Geralmente, as pesquisas que são publicadas sob o formato de resumos são

aprimoradas com as contribuições dos pares e dão origem a manuscritos que são submetidos a periódicos científicos. Os trabalhos completos também recebem *feedback* nos canais informais e posteriormente são publicados em canais formais. A publicação em canais formais também ocorre no segundo ano, com a publicação de artigos de periódicos e capítulos de livros.

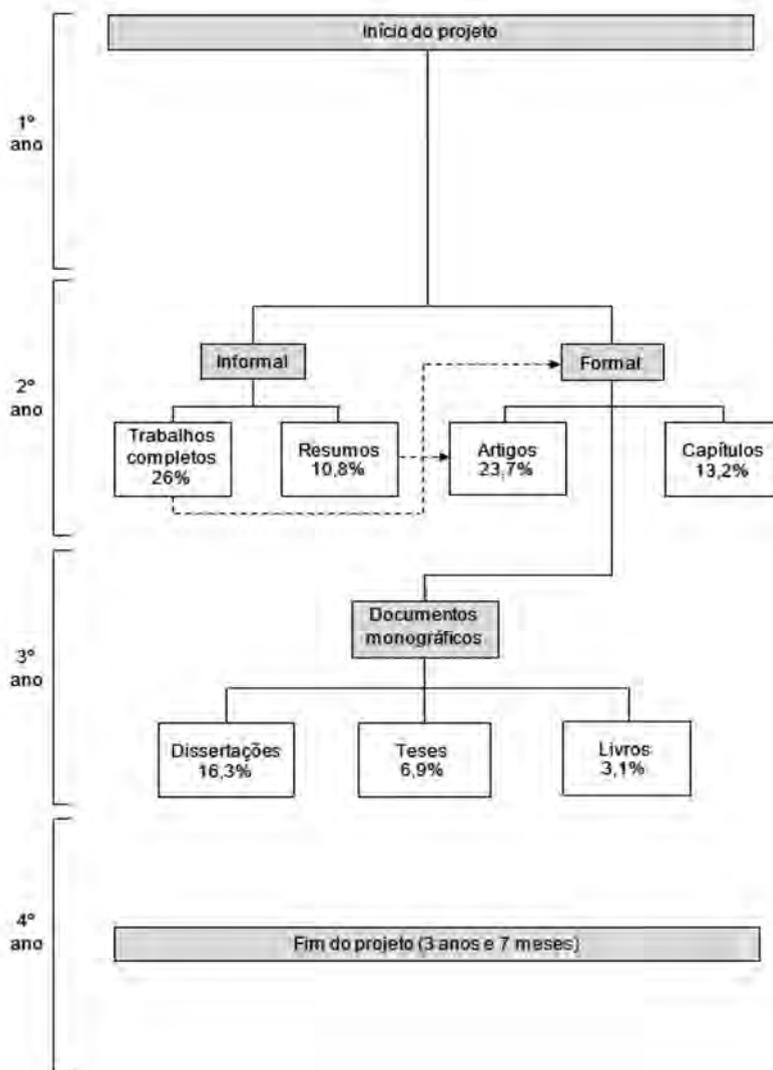


Figura 2. Modelo para o fluxo da informação científica
 Fonte: Dados da pesquisa

No terceiro ano, são publicadas as obras em formato monográfico, no caso os livros, dissertações e teses. O projeto é tipicamente encerrado no quarto ano, aos três anos e sete meses. Para cada um dos tipos documentais apresentados na *Figura 2*, são ilustrados também os percentuais que representam o seu volume de produção.

DISCUSSÃO

O modelo obtido a partir dos dados coletados por essa pesquisa, chamado aqui de Modelo FICAI-Br (Fluxo da Informação Científica das Áreas de Informação no Brasil), não inclui citações das publicações por outros autores na linha do tempo, mas é complementado com informações qualitativas, que descrevem a motivação para se publicar cada tipo de documento. Outra característica é a presença de documentos publicados antes do início formal do projeto.

Na análise comparativa com os outros modelos descritos na literatura, percebe-se que, da mesma forma que o modelo de Garvey e Griffith (1972), o FICAI-Br foi elaborado a partir de dados quantitativos de pesquisas realizadas. No primeiro modelo, foram consideradas mais de 200 pesquisas desenvolvidas por pesquisadores associados a APA entre 1967 e 1968. No segundo modelo, foram utilizados 266 projetos de pesquisa executados por pesquisadores brasileiros das áreas de informação entre 2005 e 2019. Os dois modelos consideram a importância da utilização do *feedback* obtido dos pares em eventos científicos a fim de elaborar uma versão mais aprimorada do manuscrito para ser submetido a um periódico.

Da mesma forma que o UNISIST (UNESCO, 1971), o modelo proposto aqui inclui a literatura cinzenta, considerando teses e dissertações. Assim como Lancaster (1975), considera como data do início do projeto aquela em que este foi inserido no sistema de acompanhamento (SSIE no caso do Lancaster, e Lattes no FICAI-Br). Os fenômenos estudados são aqueles que poderiam ser descritos na fase de documentação do modelo de Lievrouw (1992), sendo mais abrangente com relação aos tipos documentais estudados, inclusive se considerar Hurd (1996, 2000), Buck, Flagan e Coles (1999) e Costa (2000), por considerar também livros, teses e dissertações. Shearer e Birdsall (2002) consideram apenas periódicos. Com relação aos modelos de Correia e Rocha (2008), contempla-se ainda no modelo proposto, teses e dissertações. Melo (2014) utiliza praticamente os mesmos tipos documentais, à exceção de resumos e dissertações, e é um modelo aplicado às áreas de informação no Brasil, mas está limitado às pesquisas realizadas por doutorandos, ignorando a data efetiva de início dos projetos.

Modelos	Resumos	Trabalhos	Artigos	Capítulos	Livros	Dissertações	Teses	Fim do projeto
Garvey e Griffith (1972)	-	2º ano	4º ano	-	-	-	-	2º ano
Lancaster (1975)	-			-	-	-	-	2º ano (24m)
Correia e Rocha (2008) - Humanas	-			1º ano	1º ano	-	-	-
Correia e Rocha (2008) – Sociais Apl.	-	1º ano	1º ano	3º ano	3º ano	-	-	-
Melo (2014)	-	4º ano	6º ano	5º ano	7º ano	-	4º ano (48m)	4º ano (48m)
FICAI-Br	2º ano	2º ano	2º ano	2º ano	3º ano	3º ano	3º ano	4º ano (43m)

Tabela 4. Cronologia dos eventos (comparativo)

Fonte: Dados da pesquisa

O *Tabela 4* compara a linha do tempo do modelo FICAI-Br com os demais modelos constantes da literatura que possuem indicadores da cronologia da produção de documentos. Foi observada a coincidência da publicação de trabalhos com os modelos de Garvey e Griffith (1972) e Lancaster (1975), no segundo ano do projeto; dos livros com o Modelo para Ciências Sociais Aplicadas de Correia e Rocha (2008), no terceiro ano e o final do projeto coincidindo com o modelo de Melo (2014). Resumos e dissertações são tipos documentais que estão presentes apenas no FICAI-Br. As diferenças nas linhas do tempo dos modelos comparados são explicadas pelo fato dos modelos representarem áreas diferentes do conhecimento e de alguns deles serem mais antigos, não representando o impacto das inovações tecnológicas ocorridas nas últimas décadas. Também há de se considerar que foram criados com dados e informações de países diferentes e que o modelo de Melo (2014) compreende apenas pesquisas desenvolvidas por doutorandos.

CONSIDERAÇÕES FINAIS

A análise dos projetos de pesquisa das áreas de informação constantes da Plataforma Lattes e das publicações vinculadas a eles, após a integração

dos resultados quantitativos e qualitativos, permitiu a elaboração do modelo FICAI-Br, capaz de representar o fluxo da informação científica dessas áreas no Brasil. Tal modelo descreve os padrões de comunicação e publicação desses projetos, subsidiando previsões e tendências e melhorando o planejamento e a organização das atividades de pesquisa.

A análise quantitativa permitiu determinar quais são os canais preferencialmente utilizados pelos pesquisadores, assim como definiu a cronologia em que cada documento foi tipicamente publicado, trazendo um diagnóstico das áreas de informação ao longo do período estudado. O tempo de vida médio dos projetos foi de três anos e sete meses com média de 4,88 publicações e pico de produção no segundo ano, com preferência por trabalhos completos em eventos (26%) e artigos de periódicos (24%), o que representa exatamente a metade da produção vinculada a projetos de pesquisa. Os documentos menos extensos (resumos, trabalhos completos, artigos de periódicos e capítulos de livros) foram tipicamente publicados no segundo ano após o início do projeto, enquanto as publicações em formato monográfico (livros, teses e dissertações) foram geralmente publicadas no terceiro ano. Quase dois terços da produção foi representada por documentos classificados como canais formais (63,2%).

O estudo mostrou ainda o impacto significativo dos programas de pós-graduação sobre a produção científica brasileira, estando presente em 40,6% dos projetos de pesquisa estudados e concentrando 68,26% de todas as publicações, caracterizando assim os projetos mais produtivos. Essa maior produtividade pode ser explicada pela atuação dos estudantes de mestrado e doutorado, que incrementam a força de trabalho dos pesquisadores.

Os estudos qualitativos, de caráter complementar, trouxeram os fatores que levaram os pesquisadores a optar por cada tipo de documento estudado. A opção por trabalhos em eventos (resumidos ou completos) tem por objeto apresentar resultados parciais, submetendo-os aos pares. Os artigos de periódicos conferem maior visibilidade ao trabalho, gozam de uma melhor reputação diante dos demais tipos de documentos e conferem maior pontuação por ocasião da avaliação dos programas de pós-graduação. Já os livros (tanto autorais como organizados), tiveram como atrativos atingir um público mais amplo que o acadêmico, possuir maior liberdade de apresentação e redação, e estar associados com demandas de editores ou organizadores.

A metodologia utilizada pode ser empregada de forma mais direta em outras áreas e disciplinas da pesquisa brasileira, que poderão obter os seus próprios diagnósticos com base nos dados da Plataforma Lattes, podendo ainda ser adaptada para estudar comunidades de outros países, especialmente na América Latina, permitindo assim comparações entre países e até mesmo

estudos mais amplos de âmbito regional. Outra possibilidade é que seja aperfeiçoada com a introdução de mais variáveis, como por exemplo, medição de percentuais de produção impressa e digital, acompanhando as tendências apontadas nos modelos de Hurd (2000), entre outras possibilidades.

Finalmente, o estudo evidencia as diferenças nas linhas do tempo dos modelos estudados, que refletem áreas diferentes do conhecimento, com suas peculiaridades temporais e de contexto, especialmente em relação às inovações tecnológicas ocorridas nas últimas décadas.

REFERÊNCIAS

- Bjork, Bo-Christer. 2007. "A model of scientific communication as a global distributed information system". *Information Research: an international electronic journal*. 12 (2): n. d.
- Buck, Anne, Richard Flagan, e Betsy COLES. 1999. "Scholar's Forum: A New Model for Scholarly Communication". Acessado 25 março, 2022.
<https://authors.library.caltech.edu/25880/2/scholarsforum.pdf>.
- Correia, Anna Edna da Silva, e Enivaldo da Rocha. 2008. "A disseminação da informação científica na UFPE". *Biblios* (Lima) 30 (n. d.): 1-13.
- Costa, Sely. 2000. "Mudanças no Processo de Comunicação Científica: O Impacto do uso de Novas Tecnologias". Em *Comunicação Científica*, editado por Suzana Mueller e Edilene Passos, 85-105. Brasília: Departamento de Ciência da Informação, Universidade de Brasília.
- Creswell, John. 2010. *Projeto de pesquisa: métodos qualitativo, quantitativo e misto*. 3ª. Edição. Porto Alegre: Artmed.
- Fernandes, Henrique Denes Hilgenberg, e Jayme Leiro Vilan Filho. 2021. "Fluxo da informação científica: uma revisão dos modelos propostos na literatura em Ciência da Informação". *Em Questão*. 27 (2): 138-163.
<https://doi.org/10.19132/1808-5245272.138-163>
- Garvey, William, e Belver Griffith. 1972. "Communication and information processing within scientific disciplines: Empirical findings for Psychology". *Information Storage and Retrieval*. 8 (3): 123-136.
[https://doi.org/10.1016/0020-0271\(72\)90041-1](https://doi.org/10.1016/0020-0271(72)90041-1)
- Hurd, Julie. 1996. "Models of Scientific Communications Systems". Em *From Print to Electronic: the transformation of scientific communication*, editado por Susan Crawford, Julie Hurd e Ann Weller, 9-33. Medford: ASIS.
- Hurd, Julie. 2000. "The Transformation of Scientific Communication: a model for 2020". *Journal of the American Society for Information Science*, 51 (14): 1279-1283.
[https://doi.org/10.1002/1097-4571\(2000\)9999:9999%3C::AID-AS11044%3E3.0.CO;2-1](https://doi.org/10.1002/1097-4571(2000)9999:9999%3C::AID-AS11044%3E3.0.CO;2-1)
- Khosrowjerdi, Mahmood. 2011. "Designing a viable scientific communication model: VSM approach". *Library Hi Tech*, 29 (2): 359-372.
<https://doi.org/10.1108/07378831111138224>
- Lancaster, Frederick. 1975. "Acessibilidade da informação na pesquisa científica em processo". *Ciência da Informação*, 4 (2): 109-117.

- Lievrouw, Leah. 1992. "Communication, representation, and scientific knowledge: a conceptual framework and case study". *Knowledge and Policy*, 5 (1): 6-28.
<https://doi.org/10.1007/BF02692789>
- Meadows, Arthur Jack. 1999. *A comunicação científica*. Brasília: Briquet de Lemos/Livros.
- Melo, Bárbara. 2014. "Fluxo da comunicação científica na área de Ciência da Informação no Brasil: análise da produção científica relacionada com teses defendidas de 2008 a 2010". Monografia, Universidade de Brasília, Faculdade de Ciência da Informação.
- Mena-Chalco, Jesus Pascual, e Roberto Marcondes Cesar Jr. 2009. "scriptLattes: An open-source knowledge extraction system from the Lattes platform". *Journal of the Brazilian Computer Society*. 15 (4): 31-39.
<https://doi.org/10.1007/BF03194511>
- Mueller, Suzana Pinheiro Machado. 2000. "A ciência, o sistema de comunicação científica e a literatura científica". Em *Fontes de Informação para Pesquisadores e Profissionais*, editado por Bernadete Santos Campello, Beatriz Valadares Cendón e Jeannette Marguerite Kremer, 21-34. Belo Horizonte: Ed. UFMG.
- Pinto, Carlos Sousa, e Joaquim Luís Costa. 2018. "Padrões de comunicação em diferentes comunidades científicas". Em *Comunicação da informação, gestão da informação e gestão do conhecimento*, editado por Sely Maria de Souza Costa, Fernando César Lima Leite e Rosemeire Barbosa Tavares, 145-159. Brasília: IBICT.
- Shearer, Kathleen, e Bill Birdsall. 2002., "The transition of scholarly communications in Canada". Acessado 25 março, 2022.
<https://www.moyak.com/papers/scholarly-communications-canada.pdf>
- Shehata, Ahmed, David Ellis, e Allen Edward Foster. 2015. "The impact of information and communication technologies on informal scientific communication". *Library Review*, 64 (6): 825-842.
<https://doi.org/10.1108/LR-09-2014-0102>
- Søndergaard, Trine Fjordback, Jack Andersen, e Birger Hjørland. 2003. "Documents and the communication of scientific and scholarly information: Revising and updating the UNISIST model". *Journal of Documentation*. 59 (3): 278-320.
<https://doi.org/10.1108/00220410310472509>
- UNESCO. 1971. *UNISIST: Study Report on the Feasibility of a World Science Information System*. Paris: UNESCO.
- Vilan Filho, Jayme Leiro, João de Melo Maricato, e Janine Barcelos de Moraes Silva. 2019. "Fontes de dados de estudos métricos publicados em periódicos brasileiros das áreas de informação e no ENANCIB (2015-2017)". Trabalho apresentado no 20º. Encontro Nacional de Pesquisa em Ciência da Informação, Florianópolis, 21-25 de outubro.

Para citar este texto:

Fernandes, Henrique Denes Hilgenberg, e Jayme Leiro Vilan Filho. 2022. "Modelo de fluxo de informação científica para as áreas de informação". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 33-49.
<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58637>

Acceso abierto al patrimonio informacional en Humanidades y Ciencias Sociales de Instituciones de Educación Superior: propuesta conceptual y análisis de normatividades

Juan Miguel Palma Peña*

Artículo recibido:
23 de marzo de 2022
Artículo aceptado:
8 de agosto de 2022

Artículo de investigación

RESUMEN:

Actualmente, las sociedades tienen a su disposición información como nunca antes la tuvieron, así como nunca como ahora tuvieron tanta necesidad de acceso abierto al patrimonio informacional de las instituciones para beneficiarse informativa, educativa y científicamente. El objetivo de esta investigación es estudiar por qué las publicaciones académicas en Humanidades y Ciencias Sociales (HyCS) desarrolladas en Instituciones de Educación Superior Públicas (IESP) son patrimonio informacional, así como justificar por qué dichas manifestaciones pueden estar en acceso abierto. La metodología para esta investigación es teórica,

* Coordinación de Humanidades, Universidad Nacional Autónoma de México, México
jemajumi@hotmail.com

y se implementan los niveles de investigación descriptivo, clasificatorio, explicativo y predictivo en función de describir y analizar conceptualmente el objeto de estudio y sus funciones para inducir la comprensión del tema. El presente documento se estructura en los siguientes apartados: primero, se presenta un entramado teórico sobre el patrimonio para fundamentar la propuesta conceptual del patrimonio informacional; segundo, se estudian los elementos teóricos del acceso abierto; y tercero, se presenta el análisis de contenido y los resultados del estudio de las normatividades del acceso abierto en Latinoamérica; por último, se presentan las conclusiones. Una conclusión general consiste en que el acceso abierto al patrimonio informacional en HyCS de IESP requiere pasar de ser realizado por buenas intenciones a implementarse con normatividades, dado que es una acción útil que beneficia a las sociedades al compartir los bienes de información derivados de la ciencia y la cultura para alcanzar objetivos a escala global.

Palabras clave: Acceso abierto; Patrimonio informacional; Instituciones de Educación Superior Públicas; Humanidades y ciencias sociales

Open access to informational heritage in Humanities and Social Sciences of Higher Education Institutions: conceptual proposal and analysis of regulations

Juan Miguel Palma Peña

ABSTRACT:

Currently, societies have information at his disposition like never before, as well never before there was such a need for open access to informational heritage of institutions to benefit informatively, educationally and scientifically. The aim of this research is to study why scholarly publications developed in the fields of Humanities and Social Sciences of Public Higher Education Institutions are informational heritage, as well as to justify why these manifestations can be in open access. Methodology for this research is theoretical and levels of research are implemented: descriptive, classificatory, explanatory, and predictive in order to describe

and conceptually analyze object of study and its functions to induce understanding of topic.

This document is structured in following sections: first, a theoretical framework on heritage is presented to support conceptual proposal of informational heritage; second, theoretical elements of open access are studied; and third, content analysis and findings of study of open access regulations in Latin America are presented; finally, conclusions are presented.

A general conclusion is that open access to informational heritage requires moving from being done with good intentions to being implemented with regulations, due to is a useful action that benefits societies by sharing information goods derived from science and culture to achieve global goals.

Keywords: Open Access; Informational heritage; Public Higher Education Institutions; Humanities and Social Sciences

INTRODUCCIÓN

La comunicación académica actualmente se encuentra en transición para fortalecer el acceso abierto a las publicaciones académicas desarrolladas en Humanidades y Ciencias Sociales (HyCS) de Instituciones de Educación Superior Públicas (IESP), acción relevante para fomentar su utilización, ya que son bienes de utilidad para el cambio, desarrollo y bienestar social (Moraes, 2011: 169).

Asimismo, en los últimos años las sociedades han tomado conciencia de la importancia del patrimonio de la humanidad, y en particular, de los beneficios que aporta el patrimonio informacional; y para tal propósito se requieren acciones inductoras para conocer y valorar sus manifestaciones, como lo es el acceso abierto, el cual es una acción que beneficia a IESP y a sociedades con la disponibilidad y visibilidad de los bienes de información.

Por lo anterior, la presente investigación contiene un análisis teórico y de contenido. El análisis teórico estudia los conceptos, tipos y atributos de las publicaciones en HyCS de IESP para plantear que aquéllas tienen los elementos para ser consideradas patrimonio informacional.

El segundo análisis plantea que el acceso abierto es una acción pertinente para socializar el patrimonio informacional, y se realiza un estudio de contenido sobre las normatividades del movimiento en Latinoamérica para sustentar por qué debe estar en acceso abierto dicho patrimonio.

Los planteamientos de la presente investigación analizan el acceso abierto al patrimonio informacional en HyCS de IESP; y el estudio de este movimiento en otros campos de estudio queda abierto a futuras investigaciones, dado que los recursos de información son sujetos a tratarse con enfoques de los contextos en que se producen y usan.

EL PATRIMONIO: ENTRAMADO TEÓRICO

Actualmente, las sociedades han recobrado conciencia de la relevancia del patrimonio de la humanidad, ya que se apoyan en sus manifestaciones para sustentar sus actividades sociales, culturales, informativas de cara al logro de objetivos mundiales.

En paralelo, la literatura sobre patrimonio enuncia los bienes, objetos o expresiones de éste; sin embargo, no profundiza en los conceptos, tipos, atributos y valores por los que aquél es significativo; además, estos estudios argumentan que los elementos del patrimonio son dinámicos para apoyar a la sociedad; por lo que, con base en la revaloración y situaciones sobre el objeto de estudio, es pertinente analizar el entramado teórico del patrimonio de acuerdo con las tendencias actuales.

Al respecto, un concepto general del patrimonio se refiere a éste como:

... ese acervo de elementos culturales –tangibles unos, intangibles los otros– que una sociedad determinada considera suyos y de los que echa mano para enfrentar sus problemas ...; para formular e intentar realizar sus aspiraciones y proyectos, para imaginar, gozar y expresarse. (Bonfil, 1997: 47)

El citado concepto muestra algunos factores sociales que conforman y sustentan los tipos de patrimonio, tales como: natural, histórico, cultural, urbano, documental, bibliográfico, entre otros; los cuales están constituidos por objetos, expresiones y elementos que son de estima para la comunidad que los desarrolla y usa.

En este sentido, denominar a los objetos y expresiones como patrimonio es resultado de análisis teóricos, pragmáticos y semánticos, así como por congregar adjetivos calificativos y demostrativos al sustantivo patrimonio para complementar la dialéctica del mismo.

Estos argumentos nos permiten plantear que las publicaciones académicas en HyCS de IESP son bienes que sustentan las funciones de las sociedades modernas y tienen los elementos para ser adjetivadas como patrimonio informacional; por tales motivos, resulta fundamental analizar y realizar una propuesta conceptual.

El patrimonio informacional en HyCS de IESP: propuesta conceptual

La producción académica en HyCS de las IESP se sitúa en una coyuntura debido a la necesidad de preservar no sólo manifestaciones antiguas, sino considerar recursos de información contemporáneos que por sus atributos y aportes son propios de reconocimiento como patrimonio informacional.

La propuesta conceptual del patrimonio informacional en HyCS se fundamenta con elementos teóricos de los bienes comunes, del patrimonio cultural, del patrimonio bibliográfico y documental, y en el hecho de que las publicaciones en dichos campos tienen el fin de apoyar la resolución, mejora e innovación de fenómenos sociales, los cuales son factores que fortalecen su adjetivación como memoria intelectual colectiva.

En este sentido, teóricamente se plantea que el patrimonio informacional se refiere a los objetos, bienes intelectuales, publicaciones, manifestaciones, expresiones, documentos y recursos de información desarrollados por determinados grupos humanos, objetivados en formatos impresos, electrónicos y digitales, y cuyo principal contenido son avances y/o resultados de investigación financiados con fondos públicos en HyCS de IESP.

La propuesta conceptual consiste en conjugar el adjetivo *informacional* al sustantivo *patrimonio*, la cual se presenta en la *Figura 1*:

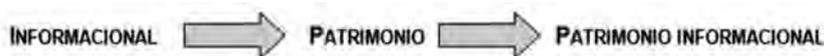


Figura 1. Patrimonio informacional en HyCS
Fuente: elaborado por el autor

El adjetivo *informacional* se refiere a que "... la información contenida en un objeto tiene un valor intrínseco..." (ALA, 2013: 134), y en esta investigación dicha determinación se emplea para sustentar la esencia de las publicaciones desarrolladas en HyCS.

El sustantivo *patrimonio* se refiere a los atributos históricos y contemporáneos de las manifestaciones, los cuales han sido determinados por los grupos humanos que elaboran aquéllas para usarlas.

Con base en los anteriores argumentos, resulta pertinente analizar los tipos de patrimonio informacional de IESP, en razón de complementar los estudios sobre el tema.

Tipos de patrimonio informacional en HyCS de IESP

Las manifestaciones que conforman el patrimonio informacional en HyCS de IESP, se basan en la tipología del Sistema HUMANINDEX (Universidad Nacional Autónoma de México, 2006).

Este apartado presenta los tipos de manifestaciones, conceptos, contenidos, objetivos y formatos que integran el patrimonio informacional en HyCS de IESP, los cuales se presentan de las *Tablas 1 a la 5*.

Publicación	Concepto y Contenido	Objetivo	Formatos
Libro	Publicación que tiene como mínimo 50 páginas e ISBN (International Standard Book Number). Su contenido es el análisis de literatura con métodos e instrumentos de investigación.	Investigación, Docencia, Difusión	Impreso, Electrónico, En línea
Libro compilado	Publicación que reúne los textos que constituirán el total de una obra.	Difusión	Impreso, Electrónico, En línea
Libro coordinado	Publicación colectiva en la que se llevan a cabo actividades en el contenido, en la edición, la compilación y la revisión de textos.	Difusión	Impreso, Electrónico, En línea
Libro editado	Publicación preparada con formas de estilo y fuentes tipográficas determinadas, con el objetivo de constituir una obra total.	Difusión	Impreso, Electrónico, En línea
Libro traducido	Publicación traducida a otro idioma diferente al que se publicó originalmente.	Investigación Docencia Difusión	Impreso, Electrónico, En línea
Antología	Contiene diversos textos de uno o varios autores.	Investigación Difusión	Impreso, Electrónico, En línea

Tabla 1. Libro y tipos de libro
Fuente: elaborado por el autor

Publicación	Concepto y Contenidos	Objetivo	Formatos
Artículo	Documentos con contenido académico publicados en revistas con ISSN (International Standard Serial Number).	Investigación Docencia Difusión	Impreso, Electrónico, En línea

Editorial	Documento con comentario breve del contenido y/o temática de una publicación periódica.	Difusión	Impreso, Electrónico, En línea
Presentación	Texto breve que da inicio a la temática tratada en alguna publicación periódica.	Difusión	Impreso, Electrónico, En línea
Nota de investigación	Texto breve que aborda alguna temática en el punto de vista de la investigación.	Investigación Difusión	Impreso, Electrónico, En línea
Opinión	Contenido con la opinión del autor acerca de un tema.	Difusión	Impreso, Electrónico, En línea

Tabla 2. Artículo y tipos de artículo

Fuente: elaborado por el autor

Publicación	Concepto y Contenido	Objetivo	Formatos
Capítulo de Libro	Publicaciones arbitradas que se ordenan en los libros para dividir su contenido por partes.	Investigación Docencia Difusión	Impreso, Electrónico, En línea

Tabla 3. Capítulos de libro

Fuente: elaborado por el autor

Publicación	Concepto y Contenido	Objetivo	Formatos
Ponencia en Memorias	Documentos presentados en eventos académicos: congresos, coloquios, encuentros, jornadas, entre otros, que son compilados en memorias para su publicación, tienen ISBN.	Investigación Docencia Difusión	Impreso, Electrónico, En línea

Tabla 4. Ponencias en memoria

Fuente: elaborado por el autor

Publicación	Concepto y Contenidos	Objetivo	Formatos
Digitales	Conjunto de objetos que incluye datos jurídicos, científicos, artísticos, entre otros, que son registrados en formatos electrónicos y digitales de origen.	Docencia Difusión	Electrónico, En línea
Multimedia	Son utilizados por las comunidades académicas para documentar videoconferencias programas, ponencias.	Docencia Difusión	Electrónico En Línea

Tabla 5. Objetos digitales y multimedia

Fuente: elaborado por el autor

Los tipos de patrimonio informacional expuestos comprueban que las publicaciones académicas en HyCS pueden ser adjetivadas como patrimonio informacional debido a sus atributos de investigación, docentes y de difusión. Asimismo, al exponer los formatos de las manifestaciones se comprueba que el patrimonio informacional requiere difundirse para ser conocido y valorado; por lo que, con base en este supuesto, es pertinente plantear que el acceso abierto es una acción con las condiciones para que el patrimonio esté disponible y visible para ser utilizado, por lo que es relevante analizar este movimiento.

ACCESO ABIERTO: ACTORES Y RUTAS

Diversos organismos internacionales (ONU, 2015; UNESCO, 2021) concuerdan en que el acceso abierto es un eje transversal para lograr objetivos mundiales y, con base en este argumento, se delimita que el acceso abierto y el patrimonio informacional tienen por objetivo común que las manifestaciones desarrolladas por IESP estén disponibles y visibles en el momento que se requieran, cuando se necesiten y sin barreras geográficas (UNESCO, 2015: 45).

Al respecto, un concepto general del acceso abierto consiste en que las sociedades tengan acceso sin limitaciones de ningún tipo a la información y conocimientos financiados con fondos públicos (Suber, 2008).

El acceso abierto puede organizarse en las siguientes etapas:

- 1ª: 1970 a 2001, tiene como base la crisis de las suscripciones (De Genaro, 1977: 69), la cual impulsó el desarrollo de canales para el acceso a la información producida con fondos públicos (Fausto, 2013).
- 2ª: 2002 a 2020, destaca por el desarrollo de normatividades (OAD, 2021); y en este período se publican las tres B del acceso abierto, conformadas por la Declaración de Budapest (Open Society Foundations, 2002), la Declaración de Bethesda (Howard, 2003) y la Declaración de Berlín (Max Planck Society, 2003).
- 3ª: 2021 a la fecha, se caracteriza por el tránsito a la ciencia abierta, (Bartling y Friesike, 2014: 9) con el objetivo de ampliar el acceso, reproducibilidad y gestión de datos y resultados de investigación (UNESCO, 2021: 7).

La periodización del acceso abierto permite identificar avances y delinear temas pendientes por estudiar por parte de los actores de este movimiento, los cuales son parte relevante de la cultura del libre acceso.

Actores del acceso abierto

El acceso abierto está conformado por diferentes actores que fomentan la investigación y el libre acceso a resultados y publicaciones académicas financiadas con fondos públicos.

Los actores del acceso abierto son los siguientes (ROARMAP):

- Agencias Gubernamentales. Incluye Gobiernos y Agencias de investigación.
- Financiadores de investigación. Incluye Consorcios y Organismos independientes.
- Organizaciones de investigación. Incluye Universidades e Institutos de investigación.
- Sub-organizaciones de investigación. Incluye Facultades, escuelas, departamentos.

Los actores listados nos permiten delimitar que gobiernos y organizaciones de investigación, incluidas IESP, son las que mayormente desarrollan e implementan acciones de acceso abierto, ya que son entidades en las que este movimiento tiene su punto de partida y en las que se realiza el mayor índice de productos académicos sujetos a estar disponibles en las rutas que mejor se adapten a las necesidades y objetivos de los contextos.

Rutas del acceso abierto

De acuerdo con la Declaración de Berlín, el acceso requiere de rutas para la circulación del patrimonio informacional en HyCS de IESP, las cuales son las siguientes:

- Ruta verde: su objetivo es la disponibilidad de las manifestaciones para su localización, acceso y uso sin restricciones. Ruta conformada por revistas y repositorios. Este canal trata recursos de información financiados con fondo públicos.
- Ruta dorada: su objetivo es almacenar, con accesos limitados, recursos publicados. Ruta compuesta por revistas comerciales y plataformas digitales. Su financiamiento es vía suscripción para el acceso y/o cargos por procesamiento de publicación.

La *Figura 2* presenta las rutas verde y dorada del acceso abierto:

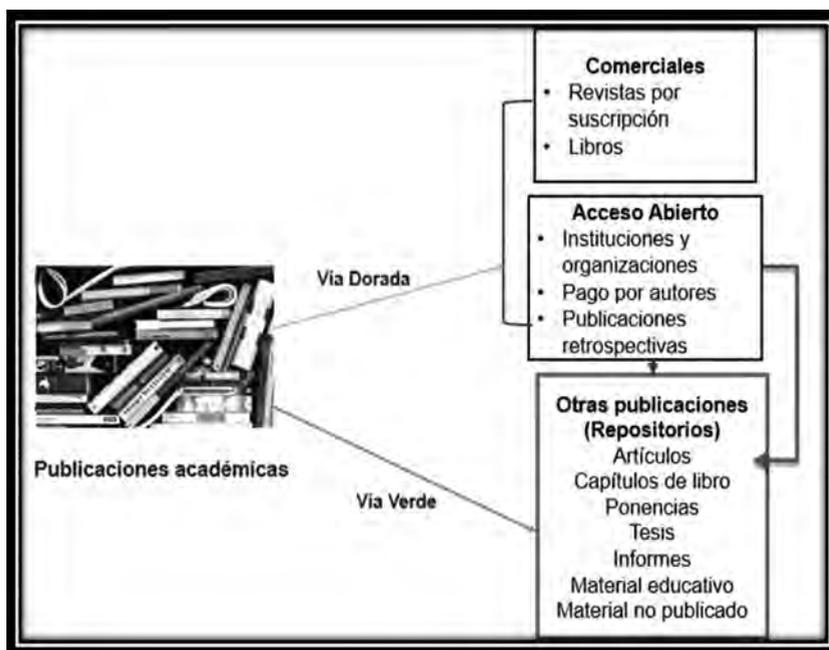


Figura 2. Rutas del acceso abierto
Fuente: (Martínez, 2017: 46)

Además, existen canales emergentes del acceso abierto, como la ruta diamante, la cual tiene las siguientes características (Arévalo, 2017):

La publicación y acceso a recursos publicados es libre de cargos para autores y lectores. La edición y revisión es mediante voluntarios de pares de alta calidad sin compensación financiera; las recompensas para los implicados son el reconocimiento académico y el bien social.

Lo expuesto anteriormente permite plantear que las rutas verde y dorada del acceso abierto tienen retos sobre redefinir y reordenar las prioridades de la disponibilidad, visibilidad y acceso a las manifestaciones informativas de IESP.

Asimismo, se considera que la ruta diamante es la que ofrecerá mayores beneficios a mediano plazo, ya que la tendencia a realizar actividades de comunicación académica en la Web fortalecerá la cultura del acceso abierto.

Paralelamente, resulta relevante referir los debates académicos (SPARC, 2021) centrados en debatir con enfoques cuantitativos las desventajas del

acceso abierto en ruta verde respecto a la ruta dorada (Langham, Bakker y Riegelman, 2021: 2); mismos estudios que dejan en lugar secundario los factores y beneficios sociales de las rutas verde y diamante del acceso abierto, las cuales presentan evidencia cualitativa comprobada de su efectividad en la sociedad actual.

En este sentido, podemos precisar que a 20 años de las declaraciones de las tres B y dado que el tránsito a la ciencia abierta demanda que el libre acceso sea realizado en el marco de la legalidad, existen temas de estudio pendientes, específicamente, sobre las normatividades del acceso abierto al patrimonio informacional.

ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

La presente investigación propone que un elemento del acceso abierto al patrimonio informacional es su formalización, la cual requiere asociarse con instrumentos legales, tales como los siguientes:

- *Declaraciones.* Recomendaciones sobre la disponibilidad de la información en acceso abierto; incluye la aceptación de los autores para el uso, reúso y distribución de sus manifestaciones.
- *Mandatos.* Lineamientos que exhortan que publicaciones sean de acceso abierto. Los mandatos tienen autoridad en facultades, escuelas, institutos o centros de investigación.
- *Políticas públicas.* Actividades decisionales para buscar el bienestar colectivo (Sánchez, E., 2010: 308); y existen tres tipos de políticas, y son las siguientes (Sánchez, N., 2010: 70):
 1. *Políticas educativas.* Políticas relacionadas con la calidad educativa (Vargas, 2014); tienen aplicación en Consejos de Ciencia y Tecnología e Instituciones de Educación.
 2. *Políticas científicas.* Medidas legislativas para actividades de descubrimiento, invención, innovación de un país (Sánchez, N., 2010: 71); están a cargo de Consejos de Ciencia y Tecnología.
 3. *Políticas de información.* “Plan de desarrollo de servicios y recursos de información y de su utilización más eficaz” (Sánchez, N., 2010: 70); tienen aplicación en Instituciones de Educación y sistemas de bibliotecas.

Los tipos de normatividades nos permiten tanto definir que para formalizar el acceso abierto es fundamental asociarlo con políticas educativas, científicas y de información, como comprobar la necesidad de reforzar los fundamentos legales del movimiento para su gestión, impacto y efectividad.

Metodología

Latinoamérica tiene amplia tradición en desarrollar acciones de acceso abierto al patrimonio informacional, lo cual se comprueba con los vastos sistemas de información y repositorios disponibles; asimismo, la región destaca por desarrollar normatividades de acceso abierto desde 2002 hasta el día de hoy (Babini y Rovelli, 2020: 47).

A la fecha, el acceso abierto se realiza de buena intención, y esto induce al cuestionamiento teórico, legal y social del movimiento; por lo que dicha situación requiere justificar que el patrimonio informacional es legalmente sujeto a estar disponibles en acceso abierto.

Al respecto, un elemento para fortalecer el anterior supuesto planteado es estudiar el contenido de los instrumentos relacionados con el contexto (Almada, 2012: 67); de ahí el planteamiento por analizar las normatividades del acceso abierto al patrimonio informacional en Latinoamérica, con el objetivo de obtener evidencias de las realidades y necesidades para su formalización.

La metodología de esta investigación consiste en buscar y recuperar las normatividades de Latinoamérica en las bases ROARMAP (Registry of Open Access Repository Mandates and Policies, por sus siglas en inglés) y MELIBEA (Directorio y estimador de políticas en favor del acceso abierto a la producción académica) (Universitat de Barcelona, 2010), las cuales son fuentes que proporcionan datos de las normatividades de acceso abierto a nivel global, información de tipos de actores que desarrollan las normas, particularidades de las normatividades, desarrollo cuantitativo, entre otros.

De acuerdo con ROARMAP, Latinoamérica tiene 68 políticas de acceso abierto registradas; y MELIBEA señala que Latinoamérica tiene 78 normatividades registradas; los datos de ambas bases corresponden a normatividades desarrolladas por diferentes tipos de actor y con tipos de aplicación obligatorio y recomendatorio.

Con la información recuperada se procede a realizar el análisis de contenido de las normatividades del acceso abierto en Latinoamérica, sobre la base de siete variables, que son las siguientes:

1. *Fechas*. Posibilita saber fecha en que se realizó la normatividad.
2. *Normatividades*. Permite referir principales mandatos y políticas.
3. *Tipo de actor*. Permite identificar actores que desarrollan las normatividades.
4. *Mandato y/o política*. Posibilita conocer si la normatividad es mandato y/o política.
5. *Tipo de aplicación*. Permite saber si la normatividad es obligatoria u opcional.

6. *Relación con biblioteca.* Posibilita saber si la normatividad colabora con bibliotecas.

7. *Políticas en ROARMAP.* Permite conocer las normatividades en la base de datos.

Los objetivos del análisis de contenido de las normatividades son los siguientes:

- Identificar actualidad de normatividades.
- Conocer principales normatividades.
- Identificar tipos de actores que producen las normatividades.
- Diferenciar si países tienen mandatos y/o políticas.
- Definir tipo de aplicación de las normatividades.
- Conocer si normatividades tienen colaboración con bibliotecas.
- Identificar si normatividades consideran proteger el patrimonio informacional.

Con base en delimitar los elementos de estudio, los resultados obtenidos se presentan a continuación.

Resultados

Los resultados del análisis de contenido de las normatividades de acceso abierto de Latinoamérica se presentan en orden alfabético por país. En cada resultado se describen las particularidades del actor en relación con el acceso abierto, y se presentan los datos obtenidos de acuerdo con las variables definidas. Por último, en cada resultado se expone una interpretación de la información.

Los resultados del análisis de contenido presentados en este apartado son síntesis de una investigación mayor, la cual contiene textos y especificidades de cada normatividad de acceso abierto (Palma, 2019: 120-147).

Argentina

País que con sus normatividades justifica la disponibilidad y visibilidad de la producción académica en acceso abierto; la *Tabla 6* muestra los resultados.

Fechas: 2011 (1), 2012 (1), 2013 (1), 2015 (1), 2017 (1)	Tipo de actor Agencias Gubernamentales (1) Organizaciones de investigación (7)
Normatividades: - Resolución 753-E/2016 / República Argentina. - Resolución CS No. 6323 / Universidad de Buenos Aires. - Política de Acceso Abierto a la Producción Académica y Científica / Universidad Nacional de Mar del Plata.	
Mandatos / políticas Mandato (1) Políticas (7)	Tipo de aplicación: Mandatorio (1) Opcional (7)
Relación con biblioteca Sin información	Políticas en ROARMAP: 8

Tabla 6. Normatividades de acceso abierto de Argentina
Fuente: elaborado por el autor

Las fechas de las normatividades son diferentes de acuerdo con su desarrollo, por ejemplo, la primera de 2011, y las otras fechas son de una normatividad. Los resultados muestran dos tipos de actores: primero, agencias gubernamentales y, segundo, organizaciones de investigación. En tipo de normatividades se indica un mandato y siete políticas. En el tipo de aplicación son: una obligatoria y siete opcionales. Sobre la relación con bibliotecas, las normatividades no refieren información. Este país tiene ocho normatividades registradas en ROARMAP.

Brasil

País de Latinoamérica pionero en regular legalmente el acceso abierto; la *Tabla 7* presenta los resultados.

Fechas: 2006 (1), 2007 (1), 2009 (1), 2010 (4), 2011 (1), 2012 (3), 2013 (3), 2014 (2), 2010 (1), 2015 (2)	Tipo de actor Agencias Gubernamentales (1) Organizaciones de investigación (26)
Normatividades: - Dispõe sobre o processo de disseminação da produção técnico-científica / Brasil. - Política de Acceso Abierto da Fiocruz / Fundação Oswaldo Cruz - Ordenanza 024/2010 Política de Repositorio Institucional / Universidad Federal de Bahía	
Mandatos / políticas Mandato (1) Políticas (26)	Tipo de aplicación: Mandatorio (1) Opcional (26)
Relación con biblioteca Sin información	Políticas en ROARMAP: 27

Tabla 7. Normatividades de acceso abierto de Brasil
Fuente: elaborado por el autor

Las normatividades de acceso abierto tienen fechas distintas acordes con su desarrollo, por ejemplo, la primera de 2006; 2010 tiene cuatro, 2012 tiene uno, 2013 tiene tres, y de 2014 a 2021 tienen uno. Estas normatividades tienen dos tipos de actor, primero, agencias gubernamentales, y segundo, organizaciones de investigación. En tipo de normatividad, tienen un mandato y veintiséis políticas. En tipo de aplicación, una es obligatoria y veintiséis opcionales. Acerca de la relación con bibliotecas, las normatividades no mencionan información. Brasil tiene 27 normatividades registradas en ROARMAP.

Colombia

Este país ha desarrollado normatividades para implementar el acceso abierto; y la *Tabla 8* expone los resultados.

Fechas: 2011 (1), 2015 (1), 2016 (1), 2018 (2)	Tipo de actor Organizaciones de investigación (6)
Normatividades: - Política Institucional de Acceso Abierto / Universidad de Antioquia - Política Institucional de Acceso Abierto a la producción científica, académica e institucional / Universidad del Rosario	
Mandatos / políticas Políticas (6)	Tipo de aplicación: Opcional (6)
Relación con biblioteca Sin información	Políticas en ROARMAP: 6

Tabla 8. Normatividades de acceso abierto de Colombia
Fuente: elaborado por el autor

Las normatividades de acceso abierto tienen fechas distintas relacionadas con su desarrollo, por ejemplo, la primera de 2011; una de 2015; una de 2016 y dos de 2018. El tipo de actor, son organizaciones de investigación. El tipo normatividad indica seis políticas. El tipo de aplicación son seis opcionales. Sobre la colaboración de bibliotecas, las normatividades no indican información. Colombia tiene seis políticas registradas en ROARMAP.

México

Este país tiene normatividades de acceso abierto; y sus resultados se muestran en la *Tabla 9*.

Fechas: 2012, (1), 2014 (2), 2015 (1), 2016 (1)	Tipo de actor Agencias Gubernamentales (1) Organizaciones de investigación (3)
Normatividades: - Ley de Ciencia y Tecnología / México - Lineamientos Generales para la Política de Acceso Abierto / Universidad Nacional Autónoma de México	
Mandatos / políticas Políticas (4)	Tipo de aplicación: Opcional (4)
Relación con biblioteca Sin información	Políticas en ROARMAP: 4

Tabla 9. Normatividades de acceso abierto de México

Fuente: elaborado por el autor

Las fechas de las normatividades son acordes con su desarrollo, por ejemplo, la primera de 2012; una de 2013; dos de 2014, y una de 2016. Los tipos de actor son dos, gubernamental e institucional. En tipo de normatividades, son cuatro políticas. El tipo de aplicación de las normatividades son opcionales. Sobre la colaboración de bibliotecas, las normatividades no refieren información. México tiene cuatro normatividades registradas en ROARMAP.

Perú

País pionero en regular el acceso abierto en Latinoamérica; y en la *Tabla 10* se presentan los resultados.

Fechas: 2011 (1), 2013 (2), 2014 (1), 2016 (1), 2020 (1)	Tipo de actor Agencias Gubernamentales (1) Organizaciones de investigación (8)
Normatividades: - Ley que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto / República de Perú - Políticas de organización y recuperación de trabajos de investigación / Universidad Católica Sedes Sapientiae (UCSS)	
Mandatos / políticas Políticas (9)	Tipo de aplicación: Mandatorio (1) Opcional (8)
Relación con biblioteca La gestión del Repositorio Institucional la realiza el Sistema de Bibliotecas de la UCSS.	Políticas en ROARMAP: 9

Tabla 10. Normatividades de acceso abierto de Perú

Fuente: elaborado por el autor

Las normatividades refieren fechas diferentes acordes con su desarrollo, por ejemplo, dos de 2010, 2011 una, 2013 dos, y de 2014 a 2020 tienen una. Las normatividades tienen dos tipos de actor, agencias gubernamentales, y organizaciones de investigación. En tipo de aplicación, una obligatoria y ocho opcionales. Este país tiene nueve políticas. Sobre la colaboración con bibliotecas, las normatividades señalan que apoyarán la gestión del repositorio institucional. Perú tiene registrados nueve normatividades en ROARMAP.

Los resultados generales del análisis de contenido de las normatividades de acceso abierto en Latinoamérica se presentan comparativamente en la *Figura 3*.

Pais	 Argentina	 Brasil	 Colombia	 México	 Perú
Países con mandatos y políticas obligatorias	✓	✓			✓
Países con políticas opcionales	✓	✓	✓	✓	✓
Países con mandatos y políticas de Agencias Gubernamentales	✓	✓		✓	✓
Países con mandatos y políticas de Organizaciones de Investigación	✓	✓	✓	✓	✓
Países con mandatos y políticas que colaboran con bibliotecas					✓

Figura 3. Resultados del análisis de normatividades de acceso abierto de Latinoamérica

Fuente: elaborado por el autor

Los datos de la *Figura 3* indican que tres países tienen mandatos y políticas obligatorias; cinco países tienen políticas opcionales; cuatro países tienen mandatos y políticas de agencias gubernamentales; cinco países tienen mandatos y políticas de organizaciones de investigación; y un país refiere que sus mandatos y políticas colaboran con bibliotecas.

Análisis de resultados

Los resultados permiten definir que las normatividades de acceso abierto de Latinoamérica se inclinan por fomentar el flujo de información en ruta verde. Asimismo, los resultados permiten plantear la pertinencia de formalizar el acceso abierto para justificar los cuestionamientos sobre las rutas abiertas.

Además, los resultados, de acuerdo con las variables definidas, nos permiten realizar las siguientes interpretaciones:

Los resultados de la variable país comprueban que el acceso abierto en Latinoamérica necesita pasar de ser realizado de buena intención a ser implementado con normatividades para apoyar el tránsito hacia la ciencia abierta.

Los resultados de la variable sobre las fechas de las normatividades nos permiten constatar mayor interés de los actores por formalizar el libre acceso; misma tendencia que actualmente está representada por los acuerdos transformadores.

Los resultados sobre los tipos de actor nos permiten identificar que Gobiernos e IESP son actores que mayormente desarrollan instrumentos legales para formalizar el acceso abierto, debido a los aportes del movimiento para sus objetivos centrales.

Los resultados de la variable sobre los tipos de normatividades nos permiten conocer el índice de instrumentos producidos en la región como avance para formalizar el acceso abierto al patrimonio informacional.

Los resultados de la variable sobre el tipo de aplicación de las normatividades demuestran que la mayoría son implementadas opcionalmente y nos permiten delimitar la necesidad de endurecer las normas para el libre acceso.

Los resultados de la variable sobre la colaboración de bibliotecas en las normatividades de acceso abierto nos permiten comprobar que escasamente se contempla esta vinculación; mismo resultado que puede considerarse es derivado del desconocimiento de las contribuciones que las bibliotecas aportan al movimiento.

CONCLUSIONES

El estudio teórico del patrimonio informacional permite como conclusión la necesidad de difundir estas manifestaciones para conocer y valorar su utilidad; y para tal propósito, el acceso abierto es una acción sustantiva.

Analizar el acceso abierto permite concluir también que es un instrumento mundialmente útil, ya que la distribución y circulación sin restricciones del patrimonio informacional beneficiará a las sociedades para que conozcan, usen y apliquen aquél para mejorar y solucionar sus necesidades básicas y específicas.

Los resultados del análisis de contenido de las normatividades en Latinoamérica comprueban que es pertinente formalizar el acceso abierto al patrimonio y, para ello, se concluye que es indispensable armonizar los derechos y políticas para justificar legalmente el usar y el compartir sin restricciones el patrimonio informacional derivado de la ciencia y la cultura.

La presente investigación comprueba que las IESP tienen los elementos para gestionar y tomar decisiones sobre el usufructo del patrimonio informacional en HyCS que desarrolla su comunidad, y se concluye que es factible considerar la colaboración formal de bibliotecas universitarias y especializadas, las cuales de buena intención aportan funciones significativas para promover el libre acceso a las manifestaciones.

REFERENCIAS

- ALA. 2013. *Glossary of Library and Information Science*. Edited by Michael Levine-Clark, Toni M. Carter. United States of America: ALA editions.
- Almada Navarro, Margarita. 2012. “Bases teóricas para comprender las Políticas de Información”. En *La naturaleza objetiva y subjetiva de las Políticas de Información*, coordinado por Egbert J. Sánchez Vanderkast, 3-26. Ciudad de México: UNAM. http://ru.iibi.unam.mx/jspui/bitstream/IIBI_UNAM/L29/1/naturaleza_politicas_informacion_c.pdf
- Arévalo, Julio. 2017. “¿Qué es la ruta diamante al acceso abierto?”. *Universo Abierto: Blog de la biblioteca de Traducción y Documentación de la Universidad de Salamanca*. <https://universoabierto.org/2017/06/09/que-es-la-ruta-diamante-de-acceso-abierto/comment-page-1/>
- Babini, Dominique, y Rovelli, Laura. 2020. *Tendencias recientes en las políticas científicas de ciencia abierta y acceso abierto en Iberoamérica*. Ciudad Autónoma de Buenos Aires: CLACSO; Fundación.
- <https://www.clacso.org/wp-content/uploads/2020/12/Ciencia-Abierta-1.pdf>
- Bartling, Sönke, y Friesike, Sascha. 2014. “Towards Another Scientific Revolution”. En *Opening Science: The evolving guide on how the Internet is changing research, collaboration and scholarly publishing*, 3-15. Springer Open.
- <https://link.springer.com/book/10.1007%2F978-3-319-00026-8>
- Bonfil Batalla, Guillermo. 1997. “Nuestro patrimonio cultural: un laberinto de significados”. En *El patrimonio nacional de México*, coordinado por Enrique Florescano, 28-56. Ciudad de México: Consejo Nacional para la Cultura y las Artes.
- De Gennaro, Richard. 1977. “Escalating Journal Prices: Time to Fight Back”. *American Libraries*. 8 (2): 69-74.
- <http://www.jstor.org/stable/25620959>
- Fausto, Sibebe. 2013. “Evolución del Acceso Abierto: breve histórico”. *Blog Scielo*. <https://blog.scielo.org/es/2013/10/21/evolucion-del-acceso-abierto-breve-historico/#.XKKAAdL0mCh>
- Howard Hughes Medical Institute. 2003. *Bethesda Statement on Open Access Publishing*. <http://legacy.earlham.edu/~peters/fos/bethesda.htm>
- Langham-Putrow, A., C. Bakker, y A. Riegelman. 2021. “Is the open access citation advantage real? A systematic review of the citation of open access and subscription-based articles”. *PLoS ONE* 16 (6): e0253129.
- <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0253129>
- Martínez Arellano, Felipe. 2017. “Metadatos y repositorios institucionales”. *Bibliotecas y Archivos* 2 (4): 44-53.
- http://www.publicaciones.enba.ipn.mx/revista/files/bibliotecas_archivos_2017.pdf

- Max Planck Society. 2003. *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*.
<http://openaccess.mpg.de/Berlin-Declaration>
- Morales Campos, Estela. 2011. "La universidad pública y su compromiso social en la producción de conocimiento". En *Derecho a la información, bien público y bien privado: acceso comunitario y acceso individual*, coordinado por Estela Morales Campos. 163-174. Ciudad de México: UNAM.
http://ru.iibi.unam.mx/jspui/bitstream/IIBI_UNAM/L9/1/derecho_a_la_informacion.pdf
- OAD. 2021. *Declarations in support of OA*.
http://oad.simmons.edu/oadwiki/Declarations_in_support_of_OA
- ONU. 2015. *Transforming our world: the agenda for sustainable development*.
<https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>
- Open Society Foundations. 2002. *Budapest Open Initiative*.
<https://www.budapestopenaccessinitiative.org/read/>
- Palma Peña, Juan. 2019. "Acceso abierto al patrimonio informacional de instituciones de educación superior: recomendaciones y estrategias". Tesis de doctorado. Universidad Nacional Autónoma de México.
<http://132.248.9.195/ptd2019/octubre/0796776/Index.html>
- Sánchez, Egbert. 2010. *Políticas de información en Universidades Públicas Estatales*. Ciudad de México: UNAM.
https://ru.iibi.unam.mx/jspui/bitstream/IIBI_UNAM/L47/1/politicas_informacion.pdf
- Sánchez, N. 2010. "Política de acceso abierto a la producción científica del Sistema Nacional de Salud de Cuba". Tesis de doctorado. Universidad de Granada; Universidad de la Habana.
<http://tesis.sld.cu/FileStorage/000418-6CDC-NancySanchezTarragoPhD2010.pdf>
- SPARC. 2021. *Trampling on Green in a Rush for Gold*.
<https://sparcopen.org/news/2021/trampling-on-green-in-a-rush-for-gold/>
- Suber, Peter. 2008. "Gratis and libre open Access". *SPARC Open Access Newsletter*.
<https://sparcopen.org/our-work/gratis-and-libre-open-access/>
- ROARMAP. *Registry of Open Access Repository Mandates and Policies*.
<http://roarmap.eprints.org/>
- UNESCO. 2015. *Scholarly Communication 1. Open Access for Researchers*. Paris: UNESCO.
<https://unesdoc.unesco.org/ark:/48223/pf0000231938>
- UNESCO. 2021. *Recommendation on Open Science*. UNESCO.
<https://unesdoc.unesco.org/ark:/48223/pf0000379949?4=null&queryId=5ea-03b1a-8267-4a66-a381-53142c2e4737>
- Universidad Nacional Autónoma de México. 2006. *HUMANINDEX: base de datos bibliográfica de humanidades y ciencias sociales*. México.
http://www.humanindex.unam.mx/humanindex/pagina/paginas_graficas_sub-sistema.php?par=1&par1=1&idi=1
- Universitat de Barcelona. 2010. *Políticas MELIBEA: directorio y estimador de políticas en favor del acceso abierto a la producción académica*.
<https://politicas.opensciencespain.org/>

Vargas, Emilio. 2014. “El papel de la política educativa”. *Acento*.
<https://acento.com.do/opinion/el-papel-de-la-politica-educativa-8157032.html>

Para citar este texto:

Palma Peña, Juan Miguel. 2022. “Acceso abierto al patrimonio informacio-
nal en humanidades y ciencias sociales de Instituciones de Educación
Superior: propuesta conceptual y análisis de normatividades”. *Investigación
Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 51-71.
<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58615>

DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58615>

Inteligencia artificial en los procesos documentales de los archivos digitales sonoros

Georgina Sanabria Medina*
Perla Olivia Rodríguez Reséndiz*

Artículo recibido:
24 de marzo de 2022
Artículo aceptado:
11 de agosto de 2022

Artículo de investigación

RESUMEN

La recuperación de información digital a través de la inteligencia artificial es en la actualidad una de las aplicaciones más prometedoras para los archivos que acopian grandes volúmenes de contenidos digitales. Las experiencias pioneras en este campo se han comenzado a desarrollar en archivos digitales sonoros. La presente investigación tuvo un enfoque cualitativo en tanto que se llevó a cabo un análisis de literatura científica reciente. Además, se identificaron cinco archivos sonoros en América, Asia y Europa, pioneros en el uso de sistemas inteligentes para la automatización de algunos procesos documentales que intervienen en la preservación sonora y se realizó una entrevista.

* Instituto de Investigaciones Bibliotecológicas y de la Información, Universidad Nacional Autónoma de México, México
gesame@hotmail.com
perlaolivia@gmail.com

Con ello, fue posible observar algunas de las primeras aplicaciones en esta materia. Este trabajo tiene como objetivo exponer experiencias de uso de la inteligencia artificial en archivos sonoros digitales, a fin de evidenciar cómo su uso comienza a extenderse a través de los procesos documentales que intervienen en el ciclo de vida por el cual transitan los objetos digitales.

Palabras clave: Objetos sonoros digitales; Archivos digitales sonoros; Procesos documentales; Inteligencia artificial

Artificial intelligence in the documentary processes of digital sound archives

Georgina Sanabria Medina and Perla Olivia Rodríguez Reséndiz

ABSTRACT

Digital information retrieval through artificial intelligence is currently one of the most promising applications for archives that collect large volumes of digital content. Pioneering experiences in this field have begun to be developed in digital sound archives. This research had a qualitative approach in that an analysis of recent scientific literature was carried out. In addition, five sound archives in the Americas, Asia and Europe were identified as pioneers in the use of intelligent systems for the automation of some of the documentary processes involved in sound preservation, and an interview was conducted. With this, it was possible to observe some of the first applications in this area. The objective of this work is to expose experiences of the use of artificial intelligence in digital sound archives, in order to show how its use begins to spread through the documentary processes that intervene in the life cycle through which digital objects transit.

Keywords: Digital sound objects; Digital sound archives; Documentary processes; Artificial intelligence

INTRODUCCIÓN

La inteligencia artificial (IA) es una disciplina que desde su creación se ha utilizado para resolver algunas de las problemáticas que aquejan a las sociedades. Su desarrollo, aplicación y evolución en áreas como la medicina, la economía, la política, la ingeniería o el medio ambiente, constituyen algunos ejemplos de su impacto en la vida cotidiana. Aunque no de forma tan avanzada, como en los ámbitos anteriormente descritos, su implementación en algunos de los procesos que se llevan a cabo en los archivos sonoros y audiovisuales comienza a ser visible.

Los archivos han tenido cambios disruptivos que han afectado sus procesos documentales; de ellos, destacan la digitalización de los documentos sonoros y la creación de documentos de origen digital, lo cual derivó en el creciente incremento de objetos digitales.

La transformación de los archivos sonoros ha ido a la par tanto del avance tecnológico, como del incremento de grandes volúmenes de datos que hacen más complejos y lentos los procesos de preservación. Estas condiciones conllevan a considerar que la utilización de algoritmos inteligentes en estos repositorios no se puede postergar más. “Desde principios del siglo XIX, nuestra capacidad de generación de datos nunca había sido tan intensa y constante. Cada año se produce más información digital que el anterior. El uso y manejo de grandes volúmenes de información digital es un signo contemporáneo” (Rodríguez, 2020: IX).

En este contexto, el uso de la IA dentro de estos archivos cobra especial relevancia debido a que los sistemas inteligentes trabajan con datos, los necesitan para realizar las tareas establecidas con éxito. Tomando en cuenta esto, los archivos sonoros digitales constituyen un nicho favorable para los algoritmos inteligentes debido a la gran cantidad de datos que almacenan.

Actualmente es posible diseñar sistemas inteligentes que permitan hacer más eficientes las tareas que se llevan a cabo y los servicios que se ofrecen en los archivos sonoros digitales. El objetivo de este artículo es determinar la forma en cómo se ha comenzado a utilizar la IA para mejorar los procesos documentales en los archivos sonoros de medios públicos, con lo cual se evidenciará cómo su uso comienza a extenderse a través de los procesos documentales que intervienen en el ciclo de vida por medio de los cuales transitan los objetos digitales.

METODOLOGÍA

Para el desarrollo de esta investigación se realizó una revisión detallada de literatura científica con la finalidad de localizar archivos digitales sonoros de medios públicos que utilizan algoritmos de IA dentro de sus procesos documentales. Esta búsqueda ayudó a identificar 5 archivos pioneros en el uso de sistemas inteligentes para la automatización de algunas tareas de preservación sonora. Se trata del sistema de radiodifusión Nippon Hoso Kyokai, el archivo de la Radio y Televisión Suiza, el archivo del Instituto Holandés de Imagen y Sonido, el archivo de Radio Televisión Española y el archivo de Radio Televisión Nacional de Colombia. Además, se realizó una entrevista con la responsable de Proyectos del Fondo Documental de Radio Televisión Española. Con ello, fue posible observar experiencias de uso de la IA en ciertos procesos documentales que intervienen en el ciclo documental.

PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

Los objetos digitales

Los objetos digitales son “documentos cuyo origen es sólo digital, es decir, que no tienen un equivalente en formato analógico o bien que no han sido digitalizados de un soporte analógico” (Rodríguez, 2016: 198), su contenido se almacena en plataformas digitales para su conservación y acceso.

El uso popularizado de computadoras y dispositivos móviles, así como las variadas aplicaciones y *softwares* para la grabación y edición de audio propician un incremento constante de este tipo de documentos.

La llegada de la era digital avivó la creación de espacios en donde es posible garantizar la preservación tanto de los documentos de origen digital como de los digitalizados. Con la aparición del disco compacto en 1980, iniciaron las primeras reflexiones en relación con la preservación digital de documentos sonoros. Hasta hace algunos años los archivos no estaban preparados para el almacenamiento y gestión de estos documentos, la tecnología digital trajo consigo nuevos desafíos (Rodríguez, 2016).

La digitalización, como una tarea cada vez más constante dentro de los archivos, y el aumento de la producción de los objetos digitales, han contribuido con el incremento diario del acervo. Esta es sin duda una de las razones por las cuales las instituciones responsables del resguardo de la memoria sonora se han visto en la necesidad de contar con almacenamientos cada vez más robustos que permitan el resguardo y la preservación de este tipo de documentos.

De este modo es como surgen los archivos sonoros digitales, los cuales cuentan con tecnología para el almacenamiento masivo digital.

Archivos digitales sonoros

Debido al incremento de los objetos digitales, actualmente se considera que “todo archivo que haya emprendido procesos de digitalización [o que resguarde objetos digitales] tarde o temprano estará ante la definición y puesta en marcha de un archivo o repositorio digital” (Rodríguez, 2016: 215).

Se recomienda que los archivos sonoros cuenten con Sistemas de Gestión y Almacenamiento Masivo Digital. La Asociación Internacional de Archivos Sonoros y Audiovisuales (IASA) los define como aquellos sistemas que “están basados en las tecnologías de la información y la comunicación (TIC), los cuales están completamente automatizados y diseñados para almacenar, administrar, mantener, distribuir y preservar un complejo conjunto de objetos digitales” (IASA, 2006: 54).

En gran medida, con estos sistemas se han establecido las bases con las que operan los archivos que resguardan colecciones digitales. El advenimiento de este tipo de objetos transformó las tareas y los procesos principales que se llevan a cabo para garantizar la preservación de los documentos: identificación, ingreso, inventario, catalogación, almacenamiento y difusión.

Los archivos digitales sonoros tienen el reto de garantizar la preservación de uno de los documentos más frágiles y vulnerables con los que cuenta la humanidad. El riesgo de pérdida del contenido digital es muy alto, de ahí que se requiera la transferencia constante de estos objetos a la par del avance tecnológico.

Es en las instituciones en donde recae la responsabilidad de que el contenido de los objetos digitales no se pierda y que, en consecuencia, se garantice su acceso para la posteridad. Por eso es prioritario realizar acciones para asegurar el ciclo de vida de los objetos sonoros digitales.

El ciclo de vida de los objetos sonoros digitales

De acuerdo con Wilson *et al.* (2006), De Jong (2016) y Rodríguez (2017), los documentos sonoros digitales tienen un ciclo de vida, el cual se desarrolla en cuatro etapas, desde que se producen hasta que se utilizan y son aprovechados (*Figura 1*).



Figura 1. Etapas del ciclo de vida del documento sonoro digital
Fuente: Elaboración propia con información de Rodríguez (2017)

La *Figura 1* muestra el ciclo de vida del documento sonoro digital. La *creación de documentos* se refiere a la producción del material sonoro. La *ingesta y evaluación* tienen que ver con verificar la integridad de los materiales digitales. La *curaduría, conservación y almacenamiento* se relaciona con garantizar la permanencia del documento en el almacenamiento digital. El *acceso, uso y reuso* es la etapa a partir de la cual un paquete de información se transforma en un nuevo *set* de datos (Rodríguez, 2017).

El ciclo de vida está determinado por el flujo de trabajo que se desarrolla dentro de los archivos sonoros digitales, es decir, por los procesos documentales, que consisten en un conjunto de operaciones y técnicas para administrar el flujo de los objetos digitales.

Conocer estos procesos es de utilidad para tener una idea clara sobre cómo es que se realiza la gestión de los documentos sonoros digitales en los archivos con el fin de garantizar su preservación. Con base en lo anterior, se comienza a vislumbrar de qué manera la IA puede contribuir para mejorar estos procesos, ya que, por ejemplo, “la adopción de soluciones de inteligencia artificial en distintas áreas de una cadena de radio y televisión es una magnífica oportunidad para beneficiar las tareas propias de los archivos” (Bazán, 2020: 128).

En la *Figura 2* se hace alusión a los procesos documentales que se llevan a cabo en los archivos sonoros de acuerdo con las etapas que conforman el ciclo de vida de los documentos sonoros digitales.

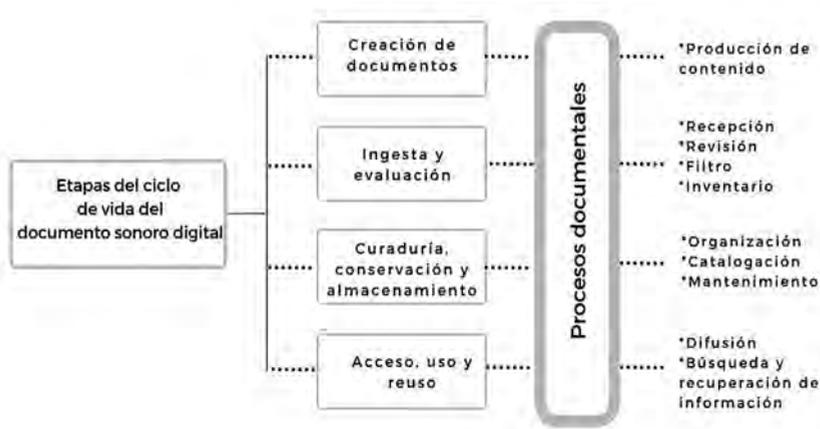


Figura 2. Procesos documentales

Fuente: Elaboración propia con información de Rodríguez (2016, 2017)

Una vez enunciados los procesos documentales y, considerando que la optimización de la gestión documental sirve para mejorar la eficacia de dichos procesos, se advierte cómo se están utilizando algoritmos inteligentes dentro algunos archivos sonoros para asegurar la preservación de sus objetos digitales.

La IA se define como una herramienta que “se basa en algoritmos que resuelven problemas. Estos algoritmos son de muy distinto tipo: algunos realizan búsquedas, otros siguen algún tipo de razonamiento, otros aprenden, algunos siguen reglas lógicas, mientras que otros se basan en probabilidades” (Oribe, 2017: 91).

A continuación, presentamos cómo la IA se está utilizando en algunos de los procesos documentales de los archivos sonoros digitales, de acuerdo con las etapas del ciclo de vida del documento digital.

El uso de la IA en los procesos documentales

Creación de documento

La IA ha tenido un avance en los últimos años en cuanto a la creación de documentos sonoros digitales. Algunas estaciones de radio comienzan a utilizar sistemas inteligentes para la creación de contenido automatizado. Las radiodifusoras han visto “en la aplicación e inclusión de algoritmos de IA oportunidades creativas durante la producción de los programas, lo cual ayuda a transmitir información a los espectadores de forma rápida, precisa y automática” (ITU-R, 2019: 1).

Un ejemplo de esto es la creación de programas de noticias del sistema de radiodifusión pública de Japón, Nippon Hoso Kyokai (NHK). Ellos han experimentado con la realización de programas noticiosos, de radio y televisión, a través de un locutor de IA.

La voz con la que el locutor lee las noticias tiene un tempo más suave y más fácil de entender que una voz humana normal. Para lograr tal voz, la IA se entrenó en funciones de voz contextuales utilizando un gran volumen de guiones de noticias y discursos de locutores. El uso del discurso de los locutores, que encarna la experiencia en anunciar, como datos de entrenamiento, da como resultado una voz sintetizada que tiene un tono noticioso natural (NHK, 2018).

Los archivos sonoros digitales que crean contenido podrán explorar el uso de las herramientas con IA que ya existen para generar producciones automáticas con el acervo que resguardan y así hacer difusión de éste. Incluso, aquellos archivos que sólo resguardan los documentos sonoros han de estar atentos porque es probable que en un futuro este tipo de objetos digitales forme parte de sus colecciones.

Ingesta y evaluación

Parte de los procesos documentales que se desarrollan en esta etapa son la recepción, revisión, filtro e inventario de los objetos digitales que se van a almacenar en el archivo.

En lo referente a la revisión de los documentos digitales, es posible desarrollar algoritmos de revisión de calidad para identificar defectos específicos en los archivos digitales (Rezzonico, s/f).

Para el documentalista hacer el registro y revisión de todos los documentos sonoros digitales que ingresan día a día implica mucho tiempo, incluso en algunos casos podrían ser años de trabajo. En este sentido, la ventaja que ofrecen los algoritmos es que pueden detectar fallas de una manera más rápida y con mayor precisión.

En el archivo de la Radio y Televisión Suiza ya se han desarrollado algoritmos de control de calidad para identificar defectos específicos de los archivos digitales. “El control de calidad manual de 4 000 archivos digitales supondría años de trabajo. En lugar de eso, una interfaz permite a los humanos examinar los resultados de los algoritmos de modo más rápido” (Rezzonico, s/f).

Otro ejemplo se observa “en el Instituto Holandés de Imagen y Sonido, donde, ya no se analiza manualmente ningún contenido salvo que realmente sea estrictamente necesario hacerlo” (V. Bazán, comunicación personal, 15 de junio de 2021).

Curaduría, conservación y almacenamiento

Dentro de estas tres etapas podemos encontrar los procesos de organización y catalogación de los documentos sonoros digitales, pero también lo referente al mantenimiento de los sistemas de gestión y almacenamiento masivo digital.

Los archivos digitales sonoros se comienzan a beneficiar con el uso de algoritmos inteligentes para la catalogación. Este proceso documental es una de las tareas más arduas que se realiza en un archivo, pues la extracción de palabras clave de un documento y su descripción son fundamentales para garantizar su posterior recuperación. Con el incremento de objetos digitales, esta labor se vuelve más lenta y compleja.

En este sentido, Cariani y Oives (2020) reconocen que para encontrar información en las colecciones es necesario describirlas y catalogarlas, pero la catalogación humana es demasiado lenta dado el volumen del material que se agrega constantemente, por ello, es importante advertir que hay una gran oportunidad para utilizar aprendizaje automático y sistemas inteligentes para ayudar a crear metadatos que mejoren la visibilidad de los documentos sonoros digitales.

Actualmente hay algoritmos que transcriben de voz a texto, extraen palabras clave y categorizan el contenido, reduciendo los errores y el tiempo que tiene que disponer un catalogador para cada uno de los documentos con los que trabaja.

Los futuros procesos de generación automática de metadatos en los archivos se fundamentarán en tres tecnologías complementarias: visión artificial, tecnologías del habla y procesamiento del lenguaje natural. La aplicación de estas tecnologías permitirá alcanzar un nivel de detalle en el análisis hasta ahora impensable (Bazán, et al., 2020: 323).

En el archivo de Radio y Televisión Suiza se han comenzado a desarrollar sistemas inteligentes para la catalogación automática de sus documentos. “Ante la falta de metadatos sobre el contenido, hemos creado un equipo de desarrollo durante los últimos dos años y medio que utiliza tecnologías de inteligencia artificial para realizar la extracción automática de metadatos” (Rezzonico, s/f).

Por otro lado, el archivo de Radio Televisión Española también está haciendo uso de sistemas para la catalogación automática:

En RTVE ya tenemos un desarrollo inteligente que permite descargar todos los días los programas que se emiten en directo y generar automáticamente los metadatos. La forma en la que participa el documentalista es en supervisar el contenido que la máquina ha generado (V. Bazán, comunicación personal, 15 de junio de 2021).

En América Latina, el archivo que también ha creado desarrollos inteligentes para este proceso documental es el que pertenece a Radio Televisión Nacional de Colombia.

La experiencia que han tenido hasta el momento les permite afirmar que “con la catalogación automática, se obtiene un mayor detalle en precisión sobre los segmentos de voz para los diferentes locutores, así como la precisión de los segmentos de música” (Ramírez y Murillo, 2020: 146).

Acceso, uso y reúso

De las propuestas que se están generando en el trabajo colaborativo entre los archivistas y los desarrolladores de sistemas inteligentes, la que ha causado más impacto es la que tiene que ver con el acceso.

“El intento ha sido desarrollar sistemas de interacción humano-computadoras inteligentes y eficientes, que permitan al usuario acceder a grandes cantidades de información heterogénea” (Avrithis, et al., 2002: 1). Sobre esto, Carrive (2019) enfatiza que con la aparición de tecnologías inteligentes se facilita la búsqueda de información y, por lo tanto, se mejora el servicio a los usuarios.

En la actualidad, sin tecnología inteligente se vuelve complicado optimizar el acceso dada la cantidad de documentos que ingresan a este tipo de archivos, por eso “es necesario y fundamental desarrollar herramientas multidisciplinarias con el objeto de presentar las últimas tecnologías que permitan superar este rezago para darle una respuesta oportuna a los requerimientos de la ciudadanía” (Ramírez y Murillo, 2020: 141). Además, “actualmente los grandes centros de información en el mundo reportan un rezago entre la cantidad de información generada y la cantidad de información procesada y puesta al servicio público” (Ramírez y Murillo, 2020: 141).

Para tratar de resolver esta problemática, la radiodifusora NHK ha desarrollado un sistema de búsqueda de información que es “revolucionario porque puede manejar el gran volumen de recursos que posee la emisora, con lo cual se hace eficaz buscar el documento deseado” (Mochizuki, 2018).

Ejemplos como el anterior dan cuenta del interés que se tiene en los archivos sonoros digitales por crear sistemas que permitan hacer una búsqueda rápida en todo el acervo y que logren recuperar la información lo más cercana posible a lo que el usuario necesita. La IA sirve para “encontrar gemas ocultas, programas y contenido que de otra manera no se verían de nuevo o se pasarían por alto. [Además] podemos utilizar la velocidad de las máquinas para ayudarnos, ya que la mirada de una computadora es mucho más rápida que la nuestra” (Altón, citado en Cowlshaw, 2019).

Con relación al reúso de los documentos sonoros digitales, hoy en día se cuenta con tecnologías inteligentes que permiten reutilizar el contenido del acervo para, a partir de esa información, generar un objeto digital nuevo.

La BBC de Londres está utilizando algoritmos inteligentes para reutilizar algunos de los documentos que resguarda y crear, a partir de ellos, nuevos objetos digitales. Lo que hicieron fue “producir una hora de programación, seleccionada del archivo de la BBC de forma automática, utilizando herramientas de *machine learning*. Crear un programa con un mínimo de intervención humana fue un proyecto ambicioso y sin precedentes” (Cowlshaw, 2019).

Una vez presentados algunos ejemplos sobre cómo la IA está contribuyendo para mejorar los procesos documentales que se gestan dentro de las cuatro etapas del ciclo de vida de los documentos sonoros digitales, se propone un modelo (Figura 3) que transparente cómo es que los sistemas inteligentes podrían intervenir en las cuatro etapas.



Figura 3. Etapas del ciclo de vida de los documentos sonoros digitales con la incorporación de IA
Fuente: Elaboración propia

Como lo podemos visualizar, cuando se incluyen algoritmos de inteligencia artificial, el ciclo de vida de los documentos digitales sonoros se vuelve más rápido y dinámico porque acelera el tiempo de ejecución de los procesos documentales. Además, se obtienen resultados en tiempo real. Al ser un ciclo no acaba porque, a partir de las recomendaciones automáticas, se generan nuevos contenidos y vuelve a comenzar el proceso.

DISCUSIÓN

Tal como se ha mostrado, algunos de los involucrados en la gestión de los archivos sonoros y audiovisuales ven en los algoritmos inteligentes una gran oportunidad para perfeccionar los procesos documentales del ciclo de vida de los documentos sonoros digitales.

El futuro de la IA en los archivos sonoros depende de aprender a integrarla en los trabajos sustanciales de preservación, ya que su uso favorecerá a quienes laboran en estos archivos y su impacto será en beneficio de la sociedad.

Sin duda, esto es un trabajo que requiere de mucho tiempo, sobre todo para poder entrenar a los algoritmos, lo sabemos; no obstante, aspiramos a que los agentes que intervienen en las tareas de los archivos digitales sonoros, así como los desarrolladores de las tecnologías inteligentes puedan tener un panorama completo sobre cómo la IA es totalmente aplicable a todo el conjunto de procesos que se gestan para la preservación de los documentos sonoros digitales.

Visualizar y promover acciones para que el archivo digital sonoro se optimice con IA se va a volver cada vez más necesario. En esta investigación se develó el uso de la IA para optimizar un proceso documental en específico, lo cual sin duda es un gran avance, pero es insuficiente si se considera que la visión a futuro y la tendencia natural de los archivos digitales sonoros, dado el volumen de datos que almacenan, es integrar sistemas inteligentes en todos los procesos documentales.

Debido a que diariamente se está generando por diversas vías información sonora, hay una “importancia de que los archivos y las bibliotecas colaboren con expertos de la comunidad de IA, dado que es un desafío mejorar el acceso a las colecciones digitales que están en constante crecimiento” (Cariani y Oives, 2020: 101).

Sin duda se necesita que haya un trabajo colegiado y multidisciplinar en donde se involucren tanto a archivistas y bibliotecarios, como a ingenieros, informáticos y matemáticos.

Desde las décadas de 1970-1980 los ingenieros, principalmente, han trabajado intensamente en el desarrollo de técnicas y algoritmos para la categorización e identificación de datos. Inicialmente fue con el tratamiento de textos con el fin de clasificar automáticamente documentos, para pasar posteriormente al tratamiento de sonido e imágenes en sus diferentes formatos (Gil, Díaz y Rodríguez, 2019: 12).

Este trabajo en conjunto brindará beneficios tanto para los desarrolladores de la IA como para los responsables de los archivos, ya que:

La comunidad de IA necesita buenos conjuntos de datos. La comunidad de archivos tiene conjuntos de datos que deben de mejorarse. Al trabajar juntos, los conjuntos de datos se pueden mejorar mejorando la calidad de los metadatos y haciendo que el conjunto de datos sea más viable para el análisis (Cariani y Oives, 2020: 102).

Por lo anterior, se deben construir relaciones sólidas y recíprocas entre los responsables de los archivos digitales sonoros y los encargados de los desarrollos de sistemas inteligentes debido, entre otras cosas, a que “los archivistas son expertos en la recopilación y el análisis de datos de una manera que los científicos informáticos no lo son, y esto coloca a los archivos en una posición única para la colaboración” (Cecchine, 2021).

Los archivos sonoros con visión de futuro tendrían que estar formando colaboraciones multidisciplinarias para ampliar los límites de su uso, ya que “deben de estar pensados para que sean resilientes, inclusivos y abiertos, para ello se requiere experimentación y colaboración entre instituciones de la memoria y múltiples disciplinas científicas” (Bocyt y Oomen, 2020).

De tal forma que este cambio profundo por el que atraviesan los archivos digitales sonoros ha de ir a la par del avance tecnológico, pues “deben de estar a la vanguardia de su propio futuro para que puedan dirigir, guiar y no perder. Las vastas masas de información en los archivos proporcionan una excelente plataforma para la explotación de la inteligencia artificial” (Hegehdus, 2020: 57).

CONSIDERACIONES FINALES

El resultado de este artículo evidencia que los archivos sonoros presentados están usando herramientas de inteligencia artificial para la automatización en la creación de documentos, en el control de calidad, en la catalogación y en la búsqueda de información. Sin embargo, es evidente que hay un campo muy rico aún por explorar sobre las aportaciones que pueden tener los sistemas inteligentes dentro de los archivos digitales sonoros, lo cierto es que ya hay un avance considerable con relación a su uso, tanto en archivos de Europa, Asia y América, y eso puede dar pie a que otros se motiven a estudiar la viabilidad de implementar algoritmos de IA para garantizar la preservación de los objetos digitales que resguardan día a día.

Algunas consideraciones finales para el uso de la IA en un archivo sonoro digital son las siguientes:

1. Es necesario reconocer, dentro del archivo digital, que hay un apuro por optimizar los procesos documentales.
2. Dejar a un lado la resistencia al cambio sobre la forma en la que se ha estado trabajando con el archivo.
3. Adentrarse en el conocimiento sobre los desarrollos inteligentes que se están llevando a cabo en los archivos sonoros digitales del mundo.
4. Comenzar a crear redes de conocimiento que vinculen a los archivistas con los expertos en IA a fin de explorar las posibilidades de uso en el archivo digital sonoro.
5. Incentivar el uso de recursos públicos para la incorporación de la IA en la preservación del patrimonio sonoro.

Agradecimientos

El artículo recupera resultados de la investigación *El uso de la inteligencia artificial en los archivos sonoros digitales para su aprovechamiento educativo*, desarrollada por Georgina Sanabria Medina en el posgrado de Bibliotecología y Estudios de la Información de la UNAM. Se expresa el más sincero agradecimiento al Conacyt por todo el apoyo para la realización de la investigación.

REFERENCIAS

- Avrithis, Yanis, Giorgos Stamou, Anastasios Delopoulos, y Stefanos Kollias. 2002. "Methods and Applications of Artificial Intelligence". Trabajo presentado en Second Hellenic Conference on AI, Thessaloniki, Greece, 11-12 de abril.
- Bazán Gil, Virginia. 2020. "Reflexiones en torno a la inteligencia artificial: el caso de los archivos de RTV". En *Inteligencia artificial y datos masivos en archivos digitales sonoros y audiovisuales*, coordinado por Perla Olivia Rodríguez Reséndiz, 117-132. Ciudad de México: IIBI-UNAM.
- Bazán Gil, Virginia, Eduardo Lleida, Carmen Pérez, Manuel Gómez, y Alberto de Prada. 2020. "Tecnologías del habla: nuevas oportunidades para los archivos de televisión". En *Actas del IV Congreso ISKO España y Portugal 2019*, editado por Jesús Tramullas, Piedad Garrido-Picazo y Gonzalo Marco-Cuenca, 323-335. Zaragoza: ISKO
<https://doi.org/10.5281/zenodo.3738614>
- Bocytte, Rasa, y Johan Oomen. 2020. "Content Adaptation, Personalisation and Fine-grained Retrieval: Applying AI to Support Engagement with and Reuse of Archival Content at Scale". *Proceedings of the 12th International Conference on Agents and Artificial Intelligence* (1): 506-51. doi: 10.5220/0009188505060511
- Cariani, Karen, y David Oives. 2020. "Using Computational Tools and Experts to Improve Access to Digital Media Archives". En *Inteligencia artificial y datos masivos en archivos digitales sonoros y audiovisuales*, coordinado por Perla Olivia Rodríguez Reséndiz, 101-115. Ciudad de México: IIBI-UNAM.

- Carrive, Jean. 2019. "Using Artificial Intelligence to Preserve Audiovisual Archives: New Horizons, More Questions". Trabajo presentado en MM'19: The 27th ACM International Conference on Multimedia, New York, United States, 21-25 de octubre.
- Cecchine, Randi. "AI and Audiovisual Archive: introduction to a blog series." *Sound and visión*, abril 21, 2021.
<https://beeldengeluid.nl/en/knowledge/blog/ai-and-audiovisual-archives-introduction-blog-series>
- Cowlshaw, Tim. "Using Artificial Intelligence to Search the Archive." *BBC*, octubre 22, 2017.
<https://www.bbc.co.uk/rd/blog/2018-10-artificial-intelligence-archive-television-bbc4>
- De Jong, Annemieke. 2016. *Digital Preservation Sound and Vision: Policy, Standards and Procedures*. Netherlands: Institute for Sound and Vision.
- Gil Leiva, Isidoro, Pedro Díaz Ortuño, y José Vicente Rodríguez Muñoz. 2019. "Técnicas y usos en la clasificación automática de imágenes". Trabajo presentado en 14º Congreso ISKO España, Barcelona, 10-11 de julio.
- Hegedus, Istvan. 2020. "How Artificial Intelligence and Machine Learning can help rethink Archives?". Trabajo presentado en 30th Conference 'International Archival Day', en línea, 7-8 de diciembre.
- IASA. 2006. *Lineamientos para la producción y preservación de objetos de audio digitales TC-04*, editado por Kevin Bradley, Ciudad de México: Radio Educación.
- ITU-R. 2019. *Artificial intelligence systems for programme production and Exchange*. Ginebra: International Telecommunication Union.
- Mochizuki, Takahiro. "Archive Search System Using Image Analysis Technology." *NHK*, diciembre, 2018.
<https://www.nhk.or.jp/str/english/publica/bt/71/2.html>
- NHK. "AI announcer has made a debut on 'News Check 11' program." *NHK*, noviembre, 2018.
<https://www.nhk.or.jp/str/english/publica/bt/74/8.html>
- Oribe, Antonio. 2017. *Una mirada al futuro. Inteligencia, abundancia, empleo y sociedad*. México: Alfaomega.
- Ramírez Acosta, Alejandro, y Juan Carlos Murillo Pabón. 2020. "Catalogación automática: más cerca que lejos". En *Inteligencia artificial y datos masivos en archivos digitales sonoros y audiovisuales*, coordinado por Perla Olivia Rodríguez Reséndiz, 133-148. Ciudad de México: IIBI-UNAM.
- Rezzonico, Pietro. "Artificial Intelligence at the service of the RTS Audiovisual Archives." *FIAT-IFTA*, s/f.
<https://fiatifta.org/index.php/artificial-intelligence-at-the-service-of-the-rts-audiovisual-archives/>
- Rodríguez Reséndiz, Perla Olivia. 2016. "El OASIS en la preservación digital de archivos sonoros". *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* (70): 197-220.
<http://dx.doi.org/10.1016/j.ibbai.2016.10.009>
- Rodríguez Reséndiz, Perla Olivia. 2017. "Gestión documental sonora: el ciclo de vida digital en los archivos sonoros". *Palabra Clave (La Plata)* 7 (1): 1-14.
<https://doi.org/10.24215/18539912e030>

- Rodríguez Reséndiz, Perla Olivia. 2020. "Aproximaciones a la inteligencia artificial y a los datos masivos en archivos digitales sonoros y audiovisuales". En *Inteligencia artificial y datos masivos en archivos digitales sonoros y audiovisuales*, coordinado por Perla Olivia Rodríguez Reséndiz, IX-XVIII. Ciudad de México: IIBI-UNAM.
- Wilson, Andrew, Richard Wright, Malcolm Polfreman, Sheila Anderson, Simon Tanner, y Emma Beer. 2006. *Digital Moving Images and Sound Archiving Study*. England: AHDS.

Para citar este texto:

- Sanabria Medina, Georgina, y Perla Olivia Rodríguez Reséndiz. 2022. "Inteligencia artificial en los procesos documentales de los archivos digitales sonoros". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 73-88.
<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58618>

Ontologías terminológicas en la organización temática de dominios de conocimiento

Adriana Suárez Sánchez*

Artículo recibido:

1 de abril de 2022

Artículo aceptado:

28 de julio de 2022

Artículo de investigación

RESUMEN

Los dominios son campos de conocimiento con una base de saber, objetivos y usuarios. Se encuentran en la realidad, como espacios cognitivos individuales o sociales. Están “ahí”, pero para ser identificados y comunicados requieren ser explicitados. Desde la perspectiva bibliotecológica, los dominios de conocimiento han sido organizados temáticamente mediante listas de temas, tesauros y clasificaciones; sin embargo, en las últimas décadas las tecnologías de la información han aportado nuevas herramientas que están transformando la actividad. A partir de lo señalado anteriormente, el objetivo del presente artículo es analizar los alcances de las ontologías terminológicas en la

* Instituto de Investigaciones Bibliotecológicas y de la Información, Universidad Nacional Autónoma de México, México asuarez@iibi.unam.mx

organización temática de dominios de conocimiento. La metodología de investigación aplicada fue exploratoria-cualitativa y se fundamentó en dos técnicas de estudio: hermenéutica del discurso y análisis de casos. Se encontró que las ontologías terminológicas son un sistema de organización de conocimiento caracterizado por elementos conceptuales, categoriales, atributivos y relacionales que les permiten representar de manera semántica campos del conocimiento. Se concluye que son sistemas con grandes ventajas en la gestión temática de dominios: contribuyen en la clarificación terminológica y conceptual de los tópicos, posibilitan el entendimiento consensuado entre los expertos, son bases de datos comprensibles para las máquinas y pueden actuar como lenguajes documentales.

Palabras clave: Ontologías terminológicas; Organización del conocimiento; Dominios de conocimiento

Terminological ontologies in subject organization of knowledge domains

Adriana Suárez Sánchez

ABSTRACT

Domains are fields of knowledge with a cognitive base, objectives, and users. They are found, as individual or social cognitive spaces. They are “there”, but to be identified and communicated they need to be made explicit. According with the librarian perspective, knowledge domains have been organized by subject through topics lists, thesauri, and classifications; however, in recent decades information technologies have provided new tools that are transforming the activity. Based on the above mentioned, the objective of this article is to analyze the scope of terminological ontologies in the subject organization of knowledge domains. The research methodology applied was analytical-qualitative and it was based on two study techniques: discourse hermeneutics and case analysis. It was found that terminological ontologies are knowledge organization systems characterized by conceptual, categorical, attributive, and relational elements that allow them to represent knowledge domains in a semantic way. It is concluded that they are systems with great advantages in subject domain management. They are

machine-understandable databases which contribute to the terminological and conceptual clarification of topics, enable consensual understanding among experts, and can act as documentary languages.

Keywords: Terminological ontologies; Knowledge organization; Knowledge domains

INTRODUCCIÓN

En el desarrollo de la sociedad, la organización temática de los dominios de conocimiento ha sido una actividad fundamental que ha tenido gran incidencia en la estructuración de la ciencia, la investigación disciplinar, la formación del currículo académico, la difusión del saber y la recuperación de la información.

El conocimiento es un cúmulo inconmensurable que solemos fragmentar en aras de hacerlo accesible (Berger y Luckmann, 2011). En este orden de ideas, la sociedad ha desarrollado herramientas particulares y generales que lo explicitan y segmentan. Desde una perspectiva particular, los profesionales y expertos de los diversos campos del saber han construido glosarios, listas de temas y diccionarios. Desde una perspectiva más general, la bibliotecología, ciencia encargada de la gestión temática del universo del saber, ha desarrollado robustos vocabularios controlados y clasificaciones (Rubin, 2010).

Durante casi todo el siglo XX, el ordenamiento temático de los dominios de conocimiento siguió un paradigma fundamentado en diccionarios, encabezamientos de materia, clasificaciones bibliográficas y tesauros; empero, tras la introducción y consecuente expansión de la web, el mundo de la información cambió en volumen, formato y recursos.

Como resultado, en los tiempos recientes las tecnologías para la representación y organización del conocimiento han traído consigo herramientas que están renovando la gestión temática de los dominios de conocimiento. Remito a Peña Vera (2011) quien asume que las alternativas para simbolizar y ordenar el conocimiento se han ido estructurando a través del tiempo, tomando en cuenta las necesidades y/o demandas informativas de cada momento histórico, y en la actualidad nos encontramos en una etapa caracterizada por modernos sistemas asociados a la organización temática de dominios de conocimiento, entre los cuales es posible mencionar las taxonomías digitales, los mapas de tópicos y las ontologías terminológicas. Moreiro (2018) coincide

al respecto cuando establece que la organización temática del conocimiento ha cambiado debido a las tecnologías de la información y en tal contexto han aparecido nuevas estructuras como las taxonomías digitales y las ontologías.

De entre los sistemas de organización de conocimiento (SOC) antes señalados, en la última década las ontologías terminológicas han llamado la atención de los bibliotecólogos y otros profesionales de la información debido a las supuestas ventajas que ofrecen en la organización temática de dominios. Pese a ello, son pocas las investigaciones que, desde una perspectiva bibliotecológica, analizan el tema.

A partir de la aseveración antes esbozada, la presente investigación tiene por objetivo: Analizar los alcances de las ontologías terminológicas en la organización temática de dominios de conocimiento.

Tomando en cuenta el contexto teórico que autores como King y Reinold (2008), Lacasta, Nogueras-Iso y Zarazaga Soria (2010) y Stuart (2016) han esbozado sobre las ontologías, los supuestos de los que partió la presente investigación asumieron que:

1. El modelado semántico y gráfico que ofrecen las ontologías terminológicas posibilita una organización temática de dominios compleja que responde a la tendencia establecida en modelos bibliotecológicos como FRBR - *Requisitos funcionales para registros bibliográficos* (1998) y FRSAD - *Requisitos funcionales para datos de autoridad de materia. Un modelo conceptual* (2010).
2. Pese a sus ventajas en la organización temática de dominios de conocimiento, la implementación de las ontologías terminológicas es reciente y su crecimiento es lento, debido a la complejidad de su construcción.
3. En el contexto de la América hispanohablante, el diseño e implementación de ontologías terminológicas en la organización temática de dominios es un asunto incipiente que requiere mayor investigación.

METODOLOGÍA

Para cumplir con el objetivo establecido en la investigación se siguió una metodología descriptiva-analítica.

En cuanto a las técnicas de investigación, se contemplaron dos procedimientos:

1. *Hermenéutica del texto*: a partir de un conjunto de documentos teóricos sobre el tema, se realizó un proceso interpretativo para determinar

el alcance, las ventajas y desventajas que las ontologías terminológicas detentan en la organización temática de dominios.

2. *Análisis de casos*: se revisaron casos de ontologías terminológicas en la organización temática de dominios. La obtención de los casos se realizó en el Descubridor de Información (DI) de la Dirección General de Bibliotecas y Sistemas de Información (2022) de la Universidad Nacional Autónoma de México, que incluye más de cien bases de datos bibliográficos: Scopus, Web of Science, Library and Information Science Source, entre otras. Los descriptores empleados en la recuperación de los casos fueron: “Ontologías terminológicas *and* Organización del conocimiento,” “Ontologías terminológicas y Dominios,” “Terminological ontologies *and* Knowledge organization,” y “Terminological ontologies *and* Domains.” Tales descriptores sirvieron como elementos de búsqueda en los campos de título, temas y resumen. A partir de la búsqueda realizada se recuperaron 36 casos de ontologías terminológicas en funciones de organización temática de dominios.¹

El análisis de los recursos seleccionados (teóricos y casos) se llevó a cabo con el programa MaxQDA que permitió la gestión de etiquetas sobre: la definición, la estructura, las características, las ventajas, las desventajas y las aplicaciones de las ontologías. Tras el trabajo desarrollado, los hallazgos de la investigación se presentan en la sección *Resultados*, que se presenta a continuación.

RESULTADOS

La organización temática del conocimiento

El conocimiento es un componente esencial en el desarrollo de la sociedad y los individuos. En su estado más esencial, como conocimiento tácito, se encuentra en la mente de las personas, formando estructuras cognitivas. Mientras que como elemento social, conocimiento explícito, emerge de la mente humana, produciendo el conocimiento compartido que se transmite entre sujetos y de generación a generación (Berger y Luckmann, 2011). El conocimiento explícito se manifiesta en segmentos de información transmitidos entre las personas o bien en información registrada en fuentes documentales (libros, revistas, videos, audios, carteles, infografías, sitios web, etcétera).

1 Los 36 casos analizados se comparten en el Anexo 1. Ontologías terminológicas en la organización temática de dominios de conocimiento.

El conocimiento está “ahí”, en la mente humana y, como tal, es imposible de aprehender. Ante tal inconveniente, los seres humanos creamos simbolizaciones y arreglos de éste. Primero, como recursos de información que no son sino meros fragmentos del conocimiento. Segundo, creamos estructuras representadoras que posibilitan su manipulación y comunicación. Teniendo en cuenta lo señalado anteriormente, la organización temática del conocimiento entraña tres grandes propósitos:

- *Propósito epistemológico*: construimos un arreglo temático del conocimiento para darle orden y sentido al universo cognitivo que la humanidad ha construido (Bates, 2010).
- *Perspectiva cognitivo-social*: el arreglo temático del conocimiento responde a un fin social-cultural. Los seres humanos se mueven en “redes sistémicas de interacción en las que es preciso articular informaciones y conocimientos” (Peña Vera, 2011: 15).
- *Perspectiva de la recuperación*: la organización temática del conocimiento es necesaria en todo intento posterior de obtener información de modo preciso y efectivo (Bates, 2010).

En este orden de ideas, pensadores como Aristóteles, Porfirio, Plinio el viejo, San Agustín de Hipona, Isidoro de Sevilla, Conrad von Gesner, Augusto Comte, Jacques Charles Brunet, Melvin Dewey, Shiyali Ramamrita Ranganathan, Henry Evelyn Bliss, etcétera, en sus respectivos tiempos esbozaron esquematizaciones temáticas del universo del conocimiento (San Segundo Manuel, 1996; Stockwell, 2001; Abbas, 2010). Y, de igual modo, los profesionales de la información en cada tiempo y espacio han formulado divisiones temáticas del conocimiento:

- Calímaco, en la Biblioteca de Alejandría, dividió el saber en seis clases: filosofía, jurisprudencia, historia, oratoria, poética y escritos misceláneos (Witty, 1958).
- En la Edad Media la división del conocimiento estuvo asociada al Trivium (gramática, retórica y dialéctica) y al Quadrivium (aritmética, geometría, astronomía y música) (San Segundo Manuel, 1996).
- En el siglo XVIII, la estructura del conocimiento se amplió con la aparición de nuevas disciplinas: mecánica, industria, economía, etcétera (Galende Diaz, 1996).
- Finalmente, el siglo XX registró gran importancia en la estructuración temática del conocimiento, pues desde la perspectiva bibliotecológica aparecieron los sistemas de clasificación y los vocabularios controlados (Rubin, 2010; Abbas, 2010).

El universo de conocimiento es un todo; un conglomerado amplio que, con fines simbolizadores y comunicativos, suele fragmentarse en secciones temáticas (*Figura 1*):



Figura 1. División de algunas áreas de conocimiento

Fuente: elaboración propia, 2022, con información de *Sistema de clasificación decimal Dewey* (2000)

Dichas secciones han sido llamadas de modos distintos: ciencias, disciplinas, campos de saber, dominios, etcétera. Entre tales denominaciones, destaca la noción de dominio, entendido como un espacio de análisis para la determinación temática de un universo cognitivo y el desarrollo de sistemas de organización del conocimiento.

En esencia, un dominio de conocimiento es un:

Ámbito del saber o de la experiencia humana que se caracteriza por su especialización, o por la posibilidad de ser delimitado en función de las prácticas o los intereses propios de una comunidad de usuarios [...] se constituye así en un universo autónomo y más o menos autosuficiente, que se apoya en una estructura de conceptos relacionados entre sí bajo una perspectiva común (Barité, 2015: 64).

El dominio es un espacio de saber que constituye el pilar de cualquier sistema de organización del conocimiento. Pese a ello, en el ámbito bibliotecario persiste una equiparación errónea entre las nociones de dominio y disciplina. En este rubro tenemos que el dominio “parte de un enfoque social y contextual que encauza y delimita las necesidades, la búsqueda y la recuperación eficaz de información, de la comunidad que compone tal dominio” (Tirador Ramos, 2010: 50), mientras que la disciplina es un campo de saber asociado

con la investigación y la formación académica. Una disciplina puede ser la física o la química, mientras que el dominio, de naturaleza más incluyente, puede ser una disciplina, una colección de vinos de cierta casa productora, un catálogo de productos comerciales, una colección de pinturas albergada en un museo, un proceso, etcétera.

Ontologías terminológicas: definición y estructura

Las ontologías son un sistema para la organización del conocimiento gestado en las ciencias computacionales hacia finales de la década de 1980.

Gruber (2008: 1) señala que:

En el contexto de la computación y las ciencias de la información, una ontología define un conjunto de primitivas representacionales con las cuales se modela un dominio de conocimiento o discurso. Las primitivas representacionales son típicamente clases (o conjuntos), atributos (o propiedades) y relaciones (o asociaciones entre los miembros de las clases).

Guarino, Oberle y Staab (2009: 2) indican que las ontologías:

Son un medio para modelar formalmente la estructura de un campo, es decir, las entidades y relaciones significativas que surgen de la observación [...] Se analizan las entidades relevantes y se organizan mediante conceptos y asociaciones que, respectivamente, se representan mediante predicados unarios y binarios.

Es común que a las ontologías se las asuma como una explicitación, un catálogo del mundo, una taxonomía o un sistema relacional; no obstante, son simbolizaciones del conocimiento que se hacen explícitas tanto para los humanos como para los agentes computacionales mediante elementos conceptuales, categoriales, atributivos y relacionales.

Por mucho tiempo, las ontologías sirvieron para modelar el conocimiento desde la perspectiva informática; sin embargo, a mediados de la década de 1990 se expandieron hacia otras disciplinas que organizan el conocimiento, la información y aún los recursos de información (García Marco, 2008).

Insertas en la bibliotecología, las ontologías fueron agregadas al amplio abanico de los sistemas de organización del conocimiento (Hodge, 2000), rubro en el que conviven con los encabezamientos de materia, las clasificaciones bibliográficas, los tesauros, las taxonomías y los mapas de tópicos.

En el ámbito bibliotecario, las ontologías aportan a la representación, organización y recuperación del conocimiento y, entre tales instrumentos, persisten dos grandes tipos: ontologías axiomatizadas y ontologías terminológicas

(Stuart, 2016). Las ontologías axiomatizadas tienen como objetivo el modelado del conocimiento en agentes artificiales, mientras que las ontologías terminológicas son útiles para organizar de modo eficiente y preciso áreas de conocimiento.

Las ontologías terminológicas “especifican los términos que se utilizan para representar el conocimiento en un dominio o discurso” (Van Heijst, Scheiberg y Wielinga, 1997: 192) y se utilizan ampliamente en el contexto de la indexación, la clasificación de recursos y la recuperación de información con miras a proporcionar formas estandarizadas para describir el contenido de los documentos y mejorar los sistemas de búsqueda (Lacasta, Nogueras-Iso y Zarazaga-Soria, 2010). Comparten rasgos con otros vocabularios controlados, desarrollados para su uso en un dominio, disciplina o comunidad de práctica (Hodge, 2000); no obstante, se diferencian de éstos en complejidad estructural, explicitación semántica y despliegue visual.

Tomando como base lo anterior, en la organización temática de un dominio (Hodge, 2000; Zeng, 2008):

- Los encabezamientos de materia incluyen el término:
Huracanes
- Las clasificaciones y taxonomías engloban el término y lo colocan en una estructura jerárquica:
Fenómenos naturales meteorológicos
Huracanes
- Los tesauros contemplan el término y establecen sus relaciones semánticas básicas:
Huracanes
BT Fenómenos naturales meteorológicos
NT Huracán Agnes
RT Marejadas
- Por último, las ontologías terminológicas incorporan el término en una estructura conceptual, jerárquica, relacional y gráfica como sigue (Figura 2):

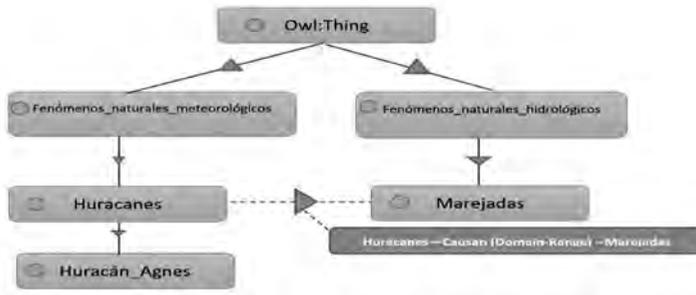


Figura 2. Ontología terminológica creada en Protégé 5.5.0. (Stanford University, 2022)

Fuente: elaboración propia, 2022

Como se aprecia en la figura anterior, mediante una ontología terminológica es posible simbolizar y organizar temáticamente un dominio de conocimiento. Asimismo, es posible establecer que “Huracanes” es una clase de “Fenómeno natural meteorológico,” que “Huracán Agnes” es un tipo de “Huracán” y que los “Huracanes” causan “Marejadas”. Aunado a ello, superando la semántica de otros lenguajes documentales, las ontologías terminológicas están en posibilidades de integrar la definición del término: “Huracanes” como “gigantescos torbellinos atmosféricos que se originan en las zonas de calma tropicales y se desplazan hacia latitudes superiores” y definir cualquier propiedad del término, por ejemplo, sobre “Huracanes” el atributo “Velocidad del viento: +117 km/h” (Figura 3):

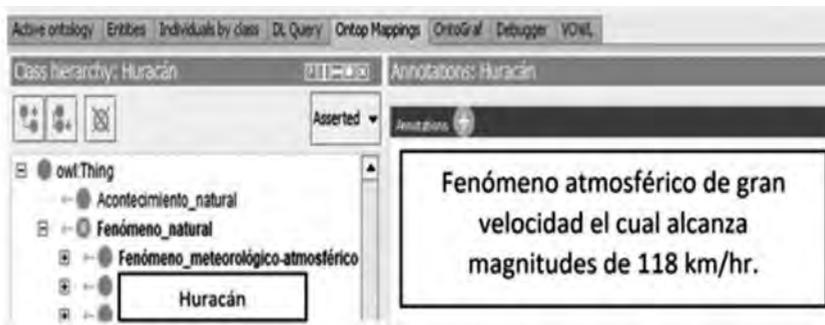


Figura 3. Explicitación Concepto-Anotación conceptual-atributiva en ontología de “Fenómenos naturales” creada en Protégé 5.5.0. (Stanford University, 2022)
Fuente: elaboración propia, 2022

Se encontró que las ontologías terminológicas son sistemas de organización del conocimiento con una estructura compleja, que ha sido definida como semántica (Figura 4):

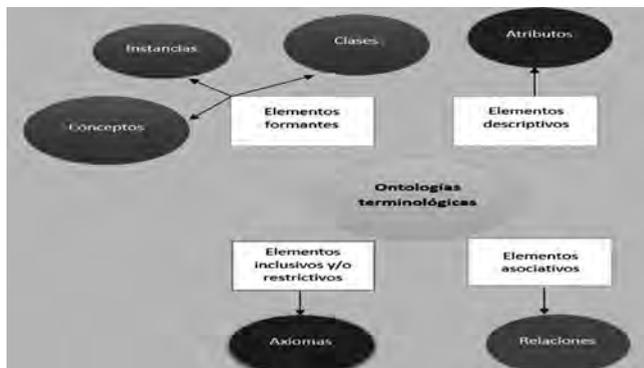


Figura 4. Estructura de las ontologías terminológicas
Fuente: elaboración propia, 2022

Desde sus elementos formantes, tenemos:

- **Conceptos:** son unidades de pensamiento (DIN: 2342, 2011. *Vocabulary of Terminology*, 2011) que se concretan en etiquetas lingüísticas (términos) y representan las entidades del dominio: Fenómenos naturales > “Huracán”, “Arcoíris,” etc.
- **Clases:** consisten en el orden jerárquico asignado a los conceptos a partir de sus rasgos en común (Reitz, 2004). Posibilitan representar un dominio de forma categorial, por ejemplo: Fenómenos naturales > “Fenómenos naturales hidrológicos”, “Fenómenos naturales biológicos”, etc.
- **Instancias:** son cada una de las entidades particulares de un dominio, por ejemplo: Ciclón tropical > “Batsirai”, “Dumako”, etc.

Desde sus elementos descriptivos, incluyen:

- **Atributos:** son características de una entidad, por ejemplo: Cirrocúmulos > “Altura de la base: 6-12 km”, “Temperatura de la base: -20/- 60 °C”, etc.

Desde sus elementos asociativos, integran:

- **Relaciones:** son vinculaciones significativas entre los elementos conceptuales. Emplean tripletas del Marco de Descripción de Recursos - RDF, constituidas bajo la sintaxis sujeto-predicado-objeto. A partir de tal esquema, es posible construir cualquier tipo de relación entre elementos conceptuales, por ejemplo: “Olas oscilatorias” causan “Ciclones tropicales”, “Cirrocúmulos” es un tipo de “Nube”, etc.

Desde sus elementos restrictivos y/o inclusivos, tenemos:

- **Axiomas:** son proposiciones lógicas que cumplen la función de establecer limitaciones. Las ontologías terminológicas registran niveles axiomáticos ligeros, asociados principalmente a clases que se señalan como separadas.

Según su estructura, las ontologías terminológicas son sistemas conceptuales, jerárquicos, atributivos, relacionales y gráficos que, en la organización temática de dominios, pueden funcionar de tres formas:

- Como un vocabulario controlado que, mediante términos, registra las nociones o entidades del dominio.
- Como una plantilla temática, segmentada en nodos, que de manera visual representa y organiza el dominio, con miras a explicitarlo en interfaces digitales locales o interfaces disponibles en la web.
- Como una base de datos que aporta información sobre el dominio, sus entidades, sus atributos y sus relaciones.

Se resume que, en la organización temática de dominios, por ejemplo “Fenómenos naturales”, las ontologías terminológicas ofrecen normalización terminológica, control sinónimo, clarificación conceptual, estructuración jerárquica, definición atributiva y entramado relacional semántico (*Figura 5*):

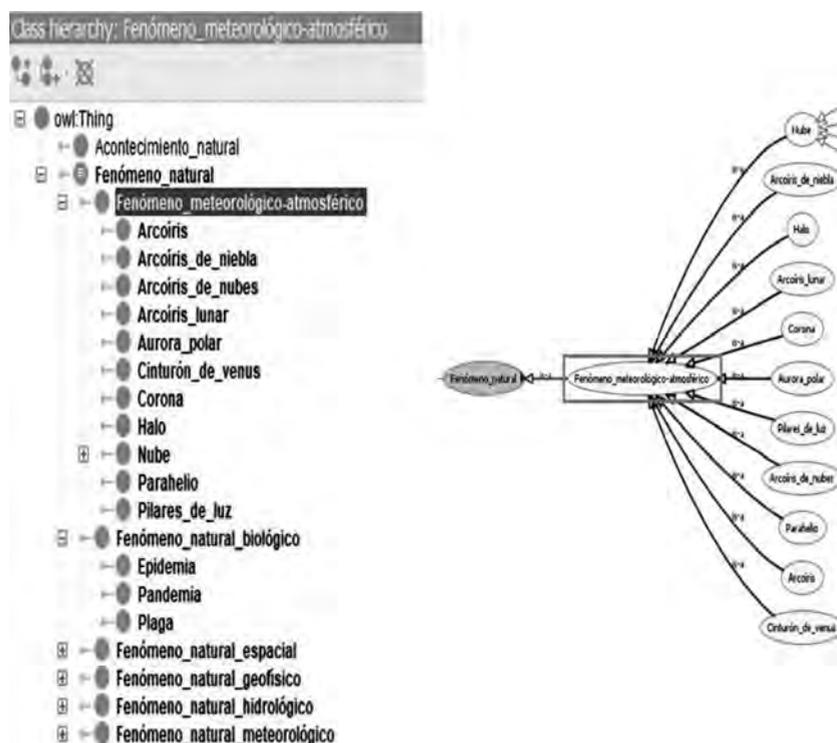


Figura 5. Ontología terminológica creada en Protégé 5.5.0. (Stanford University, 2022)
Fuente: elaboración propia, 2022

Aunado a ello, sus ventajas también están asociadas a su estructuración gráfica y desplegable en la que los usuarios pueden tener acceso al dominio (modelado gráficamente) como fuente complementaria de aprendizaje sobre el

área. Como se asumió en el primer supuesto de la investigación, las ontologías terminológicas se adhieren a una tendencia organizadora derivada de dos modelos conceptuales bibliotecológicos: desde la perspectiva de FRBR, se promulga por la construcción de sistemas fundamentados en entidades-atributos y relaciones; desde la perspectiva de FRISAD, se busca la construcción de sistemas terminológicamente robustos, con estructura jerárquica bien definida, relacionales y gráficamente explorables.

Ontologías terminológicas: tendencias y perspectivas en la organización temática de dominios de conocimiento

Como se señaló en el apartado anterior, las ontologías terminológicas ofrecen grandes beneficios en la organización temática de dominios. Pese a ello, la investigación detectó que su implementación es incipiente. A partir del análisis de casos se encontraron los siguientes resultados (*Figura 6*):

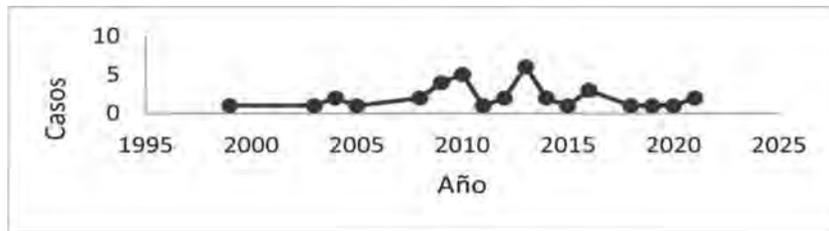


Figura 6. Años de publicación de los casos analizados
 Fuente: elaboración propia, 2022

Como puede verse en la gráfica, las ontologías terminológicas empezaron a tener presencia en la organización temática de dominios de conocimiento a principios del siglo XXI. Desde entonces, dicha tendencia se ha mantenido con algunos picos de incremento, registrados especialmente en la última década. Tal hecho coincide con lo señalado por García Marco (2008: 125) quien menciona que “el despegue de las ontologías en el campo de las ciencias de la documentación aparece como intensamente conectado con el movimiento general para abordar el problema terminológico-conceptual dentro de la informática y, concretamente, dentro de internet a través de la propuesta de la Web Semántica.”

En lo que a los dominios de las ontologías “refiere, tanto la teoría como los casos analizados confirmaron que tratan campos especializados, tanto de naturaleza cognitiva (“Neurociencias”, “Genética”) como tangible (“Vasijas de cerámica china”, “Artefactos de iluminación”).

Los dominios detectados en los casos analizados fueron (Figura 7):



Figura 7. Dominios tratados en los casos analizados
Fuente: elaboración propia, 2022

Se observó que los dominios médicos evidencian gran uso de ontologías terminológicas, ya sea como bases de datos, vocabularios controlados, mapas visuales o sistemas exploratorios para la recuperación de recursos de información.

En lo tocante a las instituciones que están desarrollando tales ontologías, persisten brechas geográficas en su diseño e implementación. La región europea (Dinamarca, Francia, Inglaterra, Países Bajos e Italia) y la zona angloparlante de América (Canadá y Estados Unidos) registran grandes avances. Algunos países asiáticos (India, China) también evidencian fuerte presencia en su construcción. En América, Brasil registra una producción teórica consistente, mientras que, como se pensó en el supuesto número tres del trabajo, el estudio de tales sistemas se observa inicial en la región hispano hablante de América Latina y requiere mayor investigación teórica y aplicada.

En este mismo rubro, se identificó que las instituciones creadoras de ontologías suelen ser entidades académicas (Universidad de Montreal, Universidad Politécnica de Valencia, etc.) y centros de investigación especializados (Centro de Oxford para la Diabetes, Centro de Investigación Atómica Bhabha, etc.).

Los casos analizados constatan que en la construcción de ontologías confluyen conocimientos multidisciplinares:

- Áreas especializadas en el dominio, por ejemplo: Departamento de Ciencias Médicas de la Universidad de Arkansas, Departamento de Neurología de la Universidad de Buffalo.

- Áreas expertas en el manejo terminológico, conceptual y categorial de los dominios y la gestión temática del conocimiento, por ejemplo: Departamento de Ciencias Bibliotecarias y Sistemas de Información de la Universidad de Atenas, Departamento de Lenguajes y Sistemas Informáticos de la Universidad Jaime I.
- Áreas con experiencia en cómputo y sistemas de información, por ejemplo: Departamento de Ingeniería de la Universidad de Florida, Instituto de Tecnología de la Universidad de California.

Se confirma el segundo supuesto del estudio: la tendencia de implementación de las ontologías terminológicas en el modelado temático de dominios es reciente y de crecimiento lento, en vista de que su construcción es una labor multidisciplinaria e implica conocimientos especializados sobre estándares y tecnologías. Consecuentemente, en la medida que una institución requiere un sistema semántico para la organización temática del dominio y conjunta los recursos (humanos, tecnológicos y económicos) para su creación, los proyectos emergen.

En cuanto a los alcances de las ontologías terminológicas en la organización temática de dominios de conocimiento, la investigación identificó que responden a los siguientes objetivos:

- Explorar métodos tecnológicos para la organización temática de dominios.
- Asentar el uso apropiado de la terminología.
- Determinar modelos conceptuales compartidos.
- Posibilitar, mediante modelos conceptuales, la comunicación temática interna y externa del dominio.
- Mejorar los procesos de información.
- Auxiliar, como una función extendida, en la organización temática de los recursos de información que los dominios producen.
- Posibilitar la búsqueda semántica de información y recursos de información.
- Proveer marcos temáticos de dominios poco tratados o recientes.
- Actuar como mapas de aprendizaje, a partir de sus rasgos gráficos-temáticos.
- Integrar los vocabularios controlados en los preceptos de la web semántica.

Por último, el estudio detectó que en la construcción de ontologías terminológicas persisten diversas metodologías. Algunas son propuestas individuales (Gruninger y Fox, 1995; Uschold y King, 1995; Noy y McGuinness, 2001) mientras que otras son propuestas vinculadas a proyectos (KACTUS, METHONTOLOGY, SENSUS, On-To-Knowledge). Pese a la diversidad metodológica, en el desarrollo de las herramientas persisten tres estándares:

- Marco de Descripción de Recursos (RDF): es un modelo para detallar recursos en la web (Fensel, Hendler y Lieberman, 2005). Las clases, atributos y relaciones que consideran las ontologías son estructuradas en sentencias sujeto-predicado-objeto, por ejemplo: Mareas - son causadas por - Fuerzas gravitacionales.
- Lenguaje de Ontologías Web (OWL): es empleado para publicar y compartir los datos que forman una ontología (Kashyap, Bussler y Moran, 2008). En agregación con RDF, posibilita la comprensión de estructuras significativas para las máquinas.
- Sistemas para la Organización del Conocimiento Simple (SKOS): es una propuesta del W3C para la estructuración de vocabularios controlados que respondan a los lineamientos de la web semántica (World Wide Web Consortium, 2012). Se observa como el eslabón perdido entre las aplicaciones de la web semántica y las herramientas para la organización temática tradicionales de la biblioteca.

CONCLUSIONES

La organización temática del conocimiento por mucho tiempo ha mantenido una preferencia por el uso de etiquetas lingüísticas como instrumentos para la representación y organización temática del universo del saber. No obstante, en las últimas décadas dicha tendencia empieza a confluir con una segunda propuesta que retoma las etiquetas lingüísticas, pero agrega mayor semanticidad relacional y estructuras gráficas.

En este tono, los resultados de la presente investigación coinciden con lo señalado por Lacasta, Nogueras-Iso y Zarazaga-Soria (2010: *viii*), quienes señalan que “el uso de simples vocabularios controlados se ha ido desplazando progresivamente por el uso de modelos de conocimiento [...] los modelos de conocimiento almacenados en papel (taxonomías, tesauros) de las bibliotecas y otras instituciones se han informatizado y transformado en modelos ontológicos más formales para proporcionar un mayor nivel de semántica.”

Actualmente, contamos con múltiples opciones para la explicitación temática de los dominios de conocimiento. En un primer nivel podemos construir listados de temas y encabezamientos de materia cuyas funciones son el asentamiento terminológico de las entidades de un dominio y el control sinonímico. En un segundo nivel contamos con tesauros y taxonomías que operan a nivel terminológico y relacional, especialmente jerárquico. Finalmente, en un tercer nivel, se nos ofrece la posibilidad de construir ontologías que contemplan elementos terminológicos, conceptuales, atributivos, relacionales y gráficos.

Si bien todas las opciones antes mencionadas son viables en la organización temática de los dominios de conocimiento, las ontologías terminológicas se observan sumamente útiles cuando trabajamos con dominios altamente especializados, con terminología y conceptualizaciones precisas que buscan asignar atributos a los conceptos y establecer cualquier tipo de relación entre los elementos conceptuales que las componen; por lo demás, son herramientas asociadas al tratamiento temático (representación, organización y recuperación) de la información en múltiples contextos digitales locales y del ciberespacio.

En la organización temática de dominios de conocimiento, las posibilidades tecnológicas que integran las ontologías son una gran ventaja, pues a diferencia de los encabezamientos de materia y tesauros que consistían en listados alfabéticos con algunos tipos de relaciones, éstas generan estructuras desplegables y navegables. La simbolización de dominios que posibilitan las ontologías favorece la comprensión del dominio, el uso correcto de su terminología, el consenso conceptual y el mapeo nodal del área.

Por último, se concluye que las tecnologías y la web han ampliado las posibilidades de representación y organización temática de dominios de conocimiento. En este contexto, la tendencia de las ontologías terminológicas es prometedora en los años próximos y en esta actividad los bibliotecólogos tienen una función primordial. En la planeación y diseño de ontologías se requieren equipos multidisciplinares y de forma particular se necesitan expertos con dos grandes capacidades: experiencia en el desarrollo de sistemas de organización del conocimiento y expertos capaces de detectar la estructura temática de los dominios de conocimiento (Murdock, Buckner y Allen, 2012). Los profesionales de la bibliotecología reúnen ambos requisitos, pues en el desarrollo de la ciencia bibliotecaria han participado en la planeación y diseño de variados sistemas de organización del conocimiento.

REFERENCIAS

- Barité, Mario Guido. 2015. *Diccionario de organización del conocimiento: clasificación, indización, terminología*. Montevideo: CSIC.
- Bates, Marcia J. (ed.). 2010. *Encyclopedia of library and information science*. Boca Raton: CRC.
- Berger, Peter L., y Thomas Luckmann. 2011. *La construcción social de la realidad*. Buenos Aires: Amorrortu.
- DIN: 2342:2011 *Vocabulary of terminology*. 2011. Berlin: Swedish Standards Institute.
- Dirección General de Bibliotecas y Servicios Digitales de Información, Universidad Nacional Autónoma de México. 2022. *Descubridor de información*. Consultado 14 de marzo de 2022.
<https://www.bidi.unam.mx/index.php>

- Fensel, Dieter, James A. Hendler, y Henry Lieberman. 2005. *Spinning the semantic web: bringing the World Wide Web to its full potential*. London: Massachusetts Institute of Technology.
- Galende Díaz, Juan Carlos. 1996. "Las bibliotecas de los humanistas y el Renacimiento". *Revista General de Información y Documentación*, 6 (2): 91-123
- García Marco, Francisco Javier. 2008. "El nacimiento y el despegue de la investigación moderna sobre ontologías." *Anuario ThinkEPI* 2: 122-125.
- Gruber, Thomas. 2008. *Ontology*. Consultado 17 de febrero de 2022.
<http://tomgruber.org/writing/ontology-definition-2007.htm>
- Gruninger Michael, y Mark Stephen Fox. 1995. "Methodology for the design and evaluation of ontologies". En *Workshop on basic ontological issues in knowledge sharing held in conjunction with IJCAI-95*. Montreal: Knowledge Sharing.
- Guarino, Nicola, Daniel Oberle, y Steffen Staab. 2009. *What is an ontology?* Consultado 22 de junio de 2022.
https://www.researchgate.net/publication/226279556_What_Is_an_Ontology
- 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.
- International Federation of Library Associations (IFLA). 1998. *Requisitos funcionales para registros bibliográficos (FRBR)*. Munich: Saur Verlag.
- _____. 2010. *Requisitos funcionales para datos de autoridad de materia. Un modelo conceptual (FRSAD)*. Washington: IFLA.
- Kashyap, Vipul, Christoph Bussler, y Matthew Moran. 2008. *The semantic web*. Heidelberg: Springer-Verlag.
- King, Brandy E., y Kathy Reinold. 2008. *Finding the concept, not just the word*. Oxford, United Kingdom: Chandos Publishing.
- Lacasta, Javier, Javier Noguera-Iso, y F. Javier Zarazaga-Soria. 2010. *Terminological ontologies: design, management and practical applications*. New York: Springer.
- Moreiro, José Antonio. 2018. "Adaptación de los vocabularios documentales al ambiente digital en red: léxico, significado y relaciones semánticas." *Informação & Sociedade* 28 (1): 35-46.
- Murdock, Jaimie, Cameron Buckner, y Colin Allen. 2012. *Containing the semantic explosion*. Consultado 14 de junio de 2021.
<https://www.semanticscholar.org/paper/Containing-the-Semantic-Explosion-MurdockAllen.d91b5d3617209b179f3002b2856a3f5a109327b7>
- Noy, Natalya F., y Deborah L. McGuinness. 2001. *Ontology development 101: a guide to create your first ontology*. Consultado 1 de noviembre de 2021.
https://protege.stanford.edu/publications/ontology_development/ontology101.pdf
- Peña Vera, Tania. 2011. *Organización y representación del conocimiento: incidencia de las tecnologías de la información y comunicación*. Buenos Aires: Alfagrama.
- Reitz, Joan M., (ed.). 2004. *Dictionary for library and information science*. Westport: Libraries Unlimited.
- Rubin, Richard. 2010. *Foundations of library and information science*. New York: Neal-Schuman.
- San Segundo Manuel, Rosa. 1996. *Sistemas de organización del conocimiento: la organización del conocimiento en las bibliotecas españolas*. Madrid: Universidad Carlos III de Madrid: Boletín Oficial del Estado.

- Sistema de clasificación decimal Dewey*. 2000. Bogotá: Rojas Eberhard.
- Stanford University. 2022. *Protégé*. Consultado 22 de junio de 2022.
<https://protege.stanford.edu/products.php>
- Stockwell, Foster. 2001. *History of information storage and retrieval*. Del Norte, Jefferson California: McFarland & Company.
- Stuart, David. 2016. *Practical ontologies: for information professionals*. Chicago: Neal-Schuman.
- Tirador Ramos, Janet. 2010. “El dominio y su implicación para la gestión de la información.” *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 24 (5): 49-60.
- Uschold Mike, y Martin King. 1995. “Towards a methodology for building ontologies”. En *Workshop on basic ontological issues in knowledge sharing held in conjunction with IJCAI-95*. Montreal: Knowledge Sharing.
- Van Heijst, G., A. Scheiberg, y B. Wielinga, 1997. “Using explicit ontologies in KBS development.” *International Journal of Human-Computer Studies* 45: 183-292.
- Witty, Francis. 1958. “The Pínakes of Callimachus”. *Library Quarterly*, 28 (2): 132-136.
- World Wide Web Consortium (W3C). 2012. *SKOS: simple knowledge organization system*. Consultado 15 de junio de 2021.
<https://www.w3.org/2004/02/skos/>
- Zeng, Marcia Lei. 2008. “Knowledge organization systems (KOS).” *Knowledge Organization* 35 (2-3): 160-182.

Para citar este texto:

- Suárez Sánchez, Adriana. 2022. “Ontologías terminológicas en la organización temática de dominios de conocimiento”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 89-113.
<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58630>

Anexo 1

Casos revisados

Ontologías terminológicas en la organización temática de dominios de conocimiento

#	Referencia	Dominio	Institución creadora
1	Periñan-Pascual, Carlos, y Francisco Arcas-Túnez. 2014. "La ingeniería del conocimiento en el dominio legal: la construcción de una ontología satélite en FunGramKB." <i>Signos: estudios de lingüística</i> 47 (84): 113-139. http://dx.doi.org/10.4067/S0718-09342014000100006	Derecho	Universidad Politécnica de Valencia Universidad Católica de San Antonio
2	Dancette, Jeanne, y Rosa Castrillón. 2009. "Un diccionario enciclopédico y ontológico: el campo de la migración transnacional." <i>Onomázein: revista de lingüística, filología y traducción</i> 20: 65-86.	Migración	Universidad de Montreal
3	Lotte Weilgaard, Christensen, y Madsen Bodil Nistrup. 2020. "A Danish terminological ontology of incident management in the field of disaster management." <i>Journal of Contingencies and Crisis Management</i> 28 (4): 466-478. http://dx.doi.org/10.1111/1468-5973.12334	Gestión de desastres	University of Southern Denmark / Department of Management, Society and Communication, Copenhagen Business School
4	Varela Vila, Tamara. 2013. "Corpora as a source of biomedical information: building a technological knowledge base." <i>Procedia - Social and Behavioral Sciences</i> 95: 630-636. http://dx.doi.org/10.1016/j.sbspro.2013.10.691	Galactosemia	Universidad de Vigo
5	Almeida, Bruno, Rute Costa, y Christophe Roche. 2019. "The names of lighting artefacts: extraction and representation of Portuguese and Spanish terms in the archaeology of Al-Andalus." <i>Revue TAL</i> 60 (3): 113-137.	Artefactos de iluminación- Arqueología	Centro de Lingüística de la Universida de NOVA de Lisboa
6	Ritschel, Bernd, Friederike Borchert, Günther Neher, Susane Schildbach, Gregor Kneitschel, Toshihiko Iyemori, Akiyo Koyama, et al. 2013. "Integration of ESPAS, IUGONET and ISDC: connection of domain and terminological ontologies." En EGU General Assembly Conference Abstracts.	Geofísica	General Assembly European Geosciences Union

7	Jimeno Yepes, Antonio, Ernesto Jiménez-Ruiz, Rafael Berlanga-Llavori, y Dietrich Rebholz-Schuhmann. 2009. "Reuse of terminological resources for efficient ontological engineering in life sciences." <i>BMC Bioinformatics</i> 10 (4): 1-13. https://doi.org/10.1186/1471-2105-10-S10-S4	Ciencias de la vida	European Bioinformatics Institute/Departamento de Lenguajes y Sistemas Informáticos, Universidad Jaime I
8	Nisheva-Pavlova, Maria, y Pavel Pavlov. 2011. "Ontology-based search and document retrieval in a digital library with folk songs." <i>Information Services and Use</i> 31 (3-4): 157-166. https://doi.org/10.3233/ISU-2012-0645	Canciones folklóricas búlgaras	Faculty of Mathematics and Informatics, Sofia University
9	Ziamba, Lukasz, Camilo Cornejo, y Howard W. Beck. 2011. "A water conservation digital library using ontologies." <i>The Electronic Library</i> 29 (2): 200-211. https://doi.org/10.1108/02640471111125168	Conservación del agua en Florida	Engineering Department, University of Florida
10	Kelso, Janet, Johann Visagie, Gregory Theiler, Alan Christoffels, Soraya Bardiën, Damian Smedley, Darren Otgaar, et al. 2003. "EVOC: a controlled vocabulary for unifying gene expression data." <i>Genome Research</i> 13 (6a): 1222-1230.	Genética	University of the Western Cape/ National Bioinformatics Institute/ Oxford Center for Diabetes
11	Baazaoui Zghal, Hajer, y Antonio Moreno. 2014. "A system for information retrieval in a medical digital library based on modular ontologies and query reformulation." <i>Multimedia Tools and Applications</i> 72 (3): 2393-2412. https://doi.org/10.1007/s11042-013-1527-4	Medicina	Manouba University, Riadi-GDL Laboratory University Rovira i Vigil ITAKA Research Group
12	Das, Subhashis, y Sayon Roy. 2016. "Faceted ontological model for brain tumor study." <i>Knowledge Organization</i> 43 (1): 3-12. https://doi.org/10.5771/0943-7444-2016-1-3	Tumores cerebrales	University of Trento ICT Doctoral School Documentation Research and Training, Indian Statistical Institute
13	Deokathey, Sangeeta, Arashanipalai Neelameghan, y Vijai Kumar. 2010. "A method for developing a domain ontology: a case study for a multidisciplinary subject." <i>Knowledge Organization</i> 37 (3): 173-184. http://dx.doi.org/10.1111/1468-5973.12334rg/10.5771/0943-7444-2010-3-173	Sistemas accionados por aceleradores	Scientific Information Resource, Bhabha Atomic Research Centre /Documentation Research and Training, Indian Statistical Institute/ Knowledge Management Group, Bhabha Atomic Research Centre

14	Arenas Grisales, Sandra Patricia, Ana María Tangarife Patino, Wilmar Arley Martínez Márquez, y Jaime Alberto Bornacelly Castro. 2009. "Ontology based system of documental information about forced displacement in Colombia." <i>Revista Interamericana de Bibliotecología</i> 32 (1): 129-152.	Desplazamiento forzado en Colombia	Escuela Interamericana de Bibliotecología, Universidad de Antioquia Instituto de Estudios Políticos, Universidad de Antioquia
15	Kyriaki-Manessi, Daphne, y Markos Dendrinou. 2014. "Developing ontology for the university archives: the domain of technological education." <i>Procedia - Social and Behavioral Sciences</i> 147: 349-359.	Educación tecnológica	Technological Educational, Institute of Athens Department of Library Science and Information Systems, Institute of Athens
16	Jensen, Mark, Alexander P Cox, Naveed Chaudhry, Marcus Ng, Donat Sule, William Duncan, Patrick Ray, et al. 2013. "The neurological disease ontology." <i>Journal of Biomedical Semantics</i> 4 (1): 42. https://doi.org/10.1186/2041-1480-4-42	Enfermedades neurológicas	Department of Philosophy, University at Buffalo Department of Neurology, University at Buffalo Department of Oral Diagnosis sciences, University at Buffalo
17	Padmavathi, T., y Madaiah Krishnamurthy. 2016. "Ontology for the domain of food science." <i>SRELS Journal of Information Management</i> 53 (5): 409-417. https://doi.org/10.17821/srels/2016/v53i5/89230	Ciencias de los alimentos	Food Science and Technology Information Services, CSIR-Central Food Technological Research Institute/ Documentation Research & Training Centre, Indian Statistical Institute
18	Zambach, Sine, y Bodil Nistrup Madsen. 2009. "Applying terminological methods and description logic for creating and implementing an ontology on inhibition." <i>Proceedings of the International Conference on Knowledge Engineering and Ontology Development</i> : 452-455. https://doi.org/10.5220/0002304904520455	Inhibición enzimática	Roskilde University Copenhagen Business School
19	De Keizer, Nicolette F., Ronald Cornet, Johanna Schonk, y A. Abu-Hanna. 1999. "Analysis and design of an ontology for intensive care diagnoses." <i>Methods of Information in Medicine</i> 38 (2): 102-12. https://doi.org/10.1267/METH99020102	Cuidados médicos intensivos	Department of Medical Informatics, Academic Medical Center of Amsterdam

20	Gangemi, Aldo, Domenico M. Pisanelli, y Geri Steve. 1999. "An overview of the ONIONS project: applying ontologies to the integration of medical terminologies." <i>Data & Knowledge Engineering</i> 31 (2): 183-220. https://doi.org/10.1016/S0169-023X(99)00023-3	Medicina	University of Bologna University of L'Aquila
21	Almeida, Bruno, y Rute Costa. 2021. "OntoAndalus: an ontology of islamic artefacts for terminological purposes". <i>Semantic Web</i> 12 (2): 295-311	Arqueología islámica	Centro de Lingüística Universidade NOVA de Lisboa
22	Raskin, Robert G., y Michael J. Pan. 2005. "Knowledge representation in the semantic web for earth and environmental terminology (SWEET)" <i>Computers & Geosciences</i> 31 (9): 1119-1125. https://doi.org/10.1016/j.cageo.2004.12.004	Ciencias de la tierra y medio ambiente	California Institute of Technology
23	Charlet, Jean, Bruno Bachimont, y Marie-Christine Jaulent. 2006. "Building medical ontologies by terminology extraction from texts: an experiment for the intensive care units." <i>Computers in Biology and Medicine</i> 36 (7-8): 857-870. https://doi.org/10.1016/j.combiomed.2005.04.012	Cuidados médicos intensivos	Université Paris Descartes
24	Gianmaria, Ajani, Guido Boella, Luigi Di Caro, y Livio Robaldo. 2017. "The european legal taxonomy syllabus: a multi-lingual, multi-level ontology framework to untangle the web of european legal terminology." <i>Applied Ontology</i> 11 (4): 325:375.	Derecho	Università degli Studi di Torino Università del Piemonte Oriental University of Luxembourg
25	Hogan, William R., Josh Hanna, Eric Joseph, y Mathias Brochhausen. 2013. "Towards a consistent and scientifically accurate drug ontology." <i>Information Technology and Nanotechnology</i> 1060: 68-73. https://doi.org/10.1101/gr.985203	Fármacos	University of Arkansas for Medical Sciences
26	Bug, William J., Giorgio A. Ascoli, Jeffrey S. Grethe, Amarnath Gupta, Christine Fennema-Notestine, Angela R. Laird, Stephen D. Larson, et al. 2008. "The NIFSTD and BIRNLex vocabularies: building comprehensive ontologies for neuroscience." <i>Neuroinformatics</i> 6 (3): 175-94. https://doi.org/10.1007/s12021-008-9032-z	Neurociencias	University of California Stanford University University of Texas George Mason University Yale University

27	Carvalho, Sara, Christopher Roche, y Rute Costa. 2015. "Ontologies for terminological purposes: the EndoTerm Project." En 11th International Conference on Terminology and Artificial Intelligence. Granada, España: CEUR Workshop Proceedings.	Endometriosis	Department of Linguistics, Universidade NOVA de Lisboa Condillac Research Group, Université de Savoie
28	Zhou, Xuezhong, Zhaohui Wu, Aining Yin, Lancheng Wu, Weiyu Fan, y Ruen Zhang. 2004. "Ontology development for unified traditional chinese medical language system" <i>Artificial Intelligence in Medicine</i> 32 (1): 15-27. https://doi.org/10.1016/j.artmed.2004.01.014	Medicina tradicional china	Zhejiang University Institute of Information on Traditional Chinese Medicine
29	Wei, Tong, Christophe Roche, y Maria Papadopoulou. 2021. "Using ISO and semantic web standard for building a multilingual terminology e-Dictionary: a use case of Chinese ceramic bases." <i>Journal of Information Science</i> 1: 1-16. https://doi.org/10.1177/01655515211022185 org/10.1177/01655515211022185 horg/10.1177/01655515211022185 htorg/10.1177/01655515211022185	Vasijas de cerámica china	Peking University Condillac Research Group, Université de Savoie Liaocheng University
30	Rubin, Daniel L. 2008. "Creating and curating a terminology for radiology: ontology modeling and analysis" <i>Journal of Digital Imaging</i> 21 (4): 355-362. https://doi.org/10.1007/s10278-007-9073-0	Radiología	Stanford University Department of Radiology, Stanford Medical Center
31	Ceusters, Werner, y Barry Smith. 2010. "Foundations for a realist ontology of mental disease" <i>Journal of Biomedical Semantics</i> 1 (1): 10. https://doi.org/10.1186/2041-1480-1-10	Enfermedades mentales	Ontology Research Group, Center of Excellence in Bioinformatics and Life Sciences
32	Nixdorf, D. R., M. T. Drangsholt, D. A. Ettlin, C. Gaul, R. De Leeuw, P. Svensson, J. M. Zakrzewska, A. De Laat, y W. Ceusters. 2012. "Classifying orofacial pains: a new proposal of taxonomy based on ontology." <i>Journal of Oral Rehabilitation</i> 39: 161-169. https://doi.org/10.1111/j.1365-2842.2011.02247	Dolores orofaciales	University of Minnesota University of Washington University of Zurich

33	Sahoo, Satya, Samden D. Lhatoo, Deepak K. Gupta, Licong Cui, Meng Zhao, Catherine Jayapandian, Alireza Bozorgi, y Guo-Qiang Zhang. 2013. "Epilepsy and seizure ontology: towards an epilepsy informatics infrastructure for clinical research and patient care." <i>Journal of the American Medical Informatics Association</i> 21 (1): 82-73. https://doi.org/10.1136/amiajnl-2013-001696	Epilepsia	University of Ohio
34	Babcock, Shane, John Beverley, Lindsay G. Cowell, y Barry Smith. 2021. "The infectious disease ontology in the age of COVID-19." <i>Journal of Biomedical Semantics</i> 12 (13): 1-20.	Enfermedades infecciosas	Department of Philosophy, Niagara University National Center for Ontological Research, University at Buffalo
35	Yoder, Matthew J, Katja István Mikó, Matthew Seltmann, Matthew A. Bertone, y Andrew R. Deans. 2010. "A gross anatomy ontology for hymenoptera." <i>PLoS One</i> 5 (12): e15991. https://doi.org/10.1371/journal.pone.0015991	Anatomía de Himenópteros	North Carolina State University
36	Walls, Ramona L., Balaji Athreya, Laurel Cooper, Justin Elser, Maria A. Gandolfo, Pankaj Jaiswal, Christopher J. Mungall, et al. 2012. "Ontologies as integrative tools for plant science." <i>American Journal of Botany</i> 99 (8): 1263-1275. https://doi.org/10.3732/ajb.1200222	Botánica	Oregon State University Cornell University/ Lawrence Berkeley National Laboratory University of Freiburg University at Buffalo

Políticas de información en las bibliotecas del Congreso General de los Estados Unidos Mexicanos

Arturo Vargas Huerta*

Artículo recibido:
23 de febrero de 2022

Artículo aceptado:
28 de julio de 2022

Artículo de investigación

RESUMEN

El artículo presenta las políticas de información de las bibliotecas del H. Congreso de la Unión de México. Realiza un diagnóstico sobre sus áreas, servicios, colecciones y personal por medio de entrevistas a los responsables de las bibliotecas en su lugar de trabajo. Interroga si en cada una de las áreas, servicios, manejo de las colecciones y selección del personal se tiene una política de información que les indique como se llevan a cabo las actividades. El instrumento es un cuestionario que se elaboró con base en las *Directrices para bibliotecas parlamentarias* creadas por la IFLA. Es una investigación documental y de campo. La muestra que se estudió son las tres bibliotecas que pertenecen al sis-

* Posgrado en Bibliotecología y Estudios de la Información, Instituto de Investigaciones Bibliotecológicas y de la Información, Universidad Nacional Autónoma de México, México
arvargash@gmail.com

tema. Compara lo que dice la IFLA sobre las directrices que deben tener este tipo de bibliotecas, con las políticas que se encontraron en la investigación documental y de campo. Los resultados obtenidos demuestran que sí cuentan con políticas de información explícitas en los niveles *macro*, *mezzo* y *micro*. Concluye que es necesario que los responsables de las bibliotecas elaboren, implementen, evalúen y mejoren políticas de información que les ayuden a alcanzar sus objetivos y metas.

Palabras clave: Políticas de Información; Biblioteca del Congreso - México

Information policy in the libraries of the General Congress of the United Mexican States

Arturo Vargas Huerta

ABSTRACT

The article presents information policies of the libraries of the Honorable Congress of the Mexican Union. It makes a diagnosis about its areas like services, collections, and human resources through interviews with those responsible for the libraries at their workplace. It asks whether in each of the areas: services, collection management and human resources selection if they have an information policy that indicates them how they should be carried out these activities. The instrument is a questionnaire elaborated based on IFLA's *Parliamentary Library Guidelines*. It is a documental and field research. The sample studied are the three libraries that belong to the system. It compares what IFLA says about the guidelines the type of libraries ought to have, with the policies that were found in the documental and field research. The results obtained demonstrate that they do count with the explicit information policies at the *macro*, *mezzo* and *micro* levels. It is concluded that it is necessary that the responsible of the libraries elaborate, implement, evaluate, and improve information policies that help them to reach their objectives and goals.

Keywords: Information policy; Library of Congress – Mexico

INTRODUCCIÓN

Esta investigación identifica las políticas de información de las bibliotecas del H. Congreso de la Unión en cada una de sus áreas, servicios, colecciones y personal. Se determina de qué nivel son: institucional, organizacional o individual y de qué tipo: implícitas o explícitas. Se incluye además una propuesta de concepto de políticas de información bibliotecarias.

Al analizar las definiciones de políticas de información de Anglada (2014: 105), Braman (2011: 3), Case (2010: 1), Pasek (2015: 298), Jaeger *et al.* (2015: 178), Washington (2014: 320), Almada Navarro (2012: 9) y Sánchez Vanderkast (2012: 31) se puede decir que las políticas de información son: las leyes, regulaciones, directrices, lineamientos, o cualquier otra posición doctrinal, relacionadas con la creación, producción, distribución, el acceso, el uso y la evaluación de la información, durante el ciclo de vida de la información. Pueden ser elaboradas a nivel gubernamental (internacional, nacional o local), institucional u organizacional, y sirven para fijar los objetivos y alcanzar sus metas.

Al intentar implementar una política de información en un ámbito concreto, en este caso en una biblioteca, como lo mencionan Fernández Bajón y Martínez Montalvo (2006: 421), lo que se hace es constituir un conjunto de acciones que regulan los criterios generales, intenciones y objetivos que orientarán las decisiones y líneas de acción más inmediatas.

Las políticas de información bibliotecarias, de acuerdo con Umar y Bakare (2018), Demeka (2016: 164-165) y Witt (2016), deben considerar varios aspectos para su realización, comenzando por el tipo de institución a la que pertenecen las bibliotecas, sus necesidades de información y recreación, los distintos formatos que existen para albergar la información y las normas nacionales e internacionales para su uso.

Así que una definición para “políticas de información bibliotecarias” quedaría como sigue: “Son las diferentes normas, directrices, regulaciones o cualquier otro tipo de lineamientos con que cuenta una biblioteca para el manejo de la información desde su selección y adquisición o elaboración hasta su puesta a disposición para los usuarios, incluyendo su uso y evaluación. Ayudan a cumplir eficientemente los objetivos y metas de la biblioteca”.

Algunos ejemplos de políticas de información que se pueden aplicar en cada una de las áreas del ciclo de vida de la información dentro de la biblioteca son:

- a) Origen o creación de la información. Artículo 6° de la Constitución Política de los Estados Unidos Mexicanos, que establece que todos

- los ciudadanos tienen derecho a la información. Derechos de autor: el bibliotecario debe considerarlos para la adquisición de información y no violar los derechos morales y patrimoniales del autor.
- b) Selección y adquisición de la información. Ley General de Bibliotecas, los editores deben entregar una copia a la biblioteca del H. Congreso de la Unión, a la Biblioteca Nacional y la Biblioteca de México. Políticas de selección y adquisición: establecen los lineamientos que debe seguir el bibliotecario para la correcta elección, para aceptar donaciones y para la compra de información.
 - c) Procesamiento de la información. Procedimiento de procesos técnicos: establece la normatividad que se debe seguir para la organización de la información. Reglas de Catalogación Angloamericanas: establece cómo se debe realizar la catalogación descriptiva de la información, para su correcta organización.
 - d) Flujo de información. Lineamientos de servicios de información: indican cuáles son los medios a través de los que la información fluye a sus usuarios. Reglamento de biblioteca: señala al bibliotecario de forma precisa cuáles son las reglas que se deben seguir para la circulación de la información.
 - e) Acceso y uso de la información. Procedimiento de circulación de información: establece cuáles usuarios tienen derecho al uso y de qué manera tienen acceso a la información. Reglamento de biblioteca: indica al bibliotecario de forma precisa quiénes tienen derecho al acceso y el uso.
 - f) Evaluación de la información. Procedimiento de desarrollo de colecciones a través del cual, mediante la medición ofrecida por indicadores, se valora si determinado acervo permanece dentro de la colección o es posible su descarte. Políticas de descarte: una vez evaluadas las colecciones, se decide qué materiales se pueden separar para una posible donación a otras bibliotecas o para otro destino.

Las políticas de información bibliotecarias se elaboran en distintos niveles del sistema: a nivel institucional, organizacional y local. Referente a los tipos, se encuentran las explícitas o formales: están por escrito e implementadas, son conocidas y aplicadas por todo el personal. Las implícitas o informales: no están por escrito, su implementación se lleva de manera tradicional a través de una cultura de comunicación entre el personal.

En México el poder legislativo se deposita en un Congreso General con dos Cámaras: de Diputados y de Senadores. La Cámara de Diputados se integra con representantes de la Nación, electos en su totalidad cada tres años.

Consta de un total de 500 diputados. La Cámara de Senadores se integra por 128 senadores, se renueva en su totalidad cada seis años.

Para el ejercicio de sus funciones, cuentan en su organización con órganos políticos, técnicos y administrativos. Entre los que podemos mencionar: la mesa directiva, los grupos parlamentarios, comisiones, secretaría general, secretaría de servicios parlamentarios, contraloría interna entre otras. Además, constan de un conjunto de áreas que dan apoyo a las comisiones, coordinaciones, comités, secretarías, direcciones, institutos y dependencias camerales, por ejemplo, las relativas a la recopilación y sistematización de la información: las bibliotecas.

Este Sistema de Bibliotecas, de acuerdo con el reglamento publicado el 14 de mayo de 2019, está integrado por la Biblioteca del Congreso de la Unión (General), la Biblioteca Legislativa de San Lázaro y la Biblioteca Parlamentaria “Melchor Ocampo” del Senado de la República.

Su objetivo es proporcionar servicios de información bibliográfica y documental al público en general y a los legisladores en particular, con la intención de divulgar e informar sobre las actividades del Congreso.

METODOLOGÍA

Es una investigación de análisis documental y de campo. Para la investigación de campo se llevaron a cabo entrevistas estructuradas a los responsables de las bibliotecas *in situ*, durante los meses de noviembre y diciembre de 2019. La muestra está compuesta por las tres bibliotecas pertenecientes al sistema. El cuestionario que se utilizó (véase *Anexo 1*) está dividido en las siguientes áreas tomadas de las *Directrices para bibliotecas parlamentarias* de la Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas (IFLA), elaboradas por Cuninghame (2009):

1. Personal
2. Desarrollo de Colecciones y Presupuesto
3. Organización Documental
4. Servicios de Información
5. Cooperación Bibliotecaria Institucional
6. Infraestructura (edificio, mobiliario y equipo)

PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

Durante la revisión documental se encontró que existen políticas de información explícitas con las que cuentan las bibliotecas del sistema, es decir, que se encuentran por escrito, están autorizadas y se considera que son conocidas por el personal (esto último se verifica con la investigación de campo, resultado de las entrevistas *in situ*).

En el nivel *macro* o institucional, es decir, las que elabora el Congreso General como institución encargada de crear leyes, decretos y reglamentos, entre otros, se encontraron las siguientes políticas de información bibliotecarias (*Tabla 1*).

Políticas de información bibliotecarias	Descripción	Características	Dirigida a:	Elabora / Aprueba
<i>Ley Orgánica del Congreso General de los Estados Unidos Mexicanos; Sección Segunda De la Secretaría de Servicios Parlamentarios, Artículo 49 (2021)</i>	1. f) Servicios de Biblioteca, que comprende los de: acervo de libros, hemeroteca, videoteca, multimedia, museografía, e informática parlamentaria. 2. Cada uno de los Servicios establecidos en el párrafo anterior se constituye en una Dirección, la cual se estructura con las oficinas que se requieran, conforme a lo que se disponga en el Estatuto.	Ley constitucional que establece la institucionalidad de las bibliotecas del H. Congreso de la Unión, así como los servicios que deben incluir.	Biblioteca Legislativa, Biblioteca General y Biblioteca del Senado de la República.	Cámara de Diputados, LXV Legislatura
<i>Decreto Depósito Legal (1991)</i>	Dispone la obligación de los editores y productores de materiales bibliográficos y documentales, de entregar ejemplares de sus obras a la Biblioteca Nacional y a la Biblioteca del Congreso de la Unión.	Este decreto quedó abrogado por el Decreto por el que se expide la Ley General de Bibliotecas (2021) donde se detalla, en el Capítulo X, el tipo de publicaciones recopiladas, así como agrega a la Biblioteca de México como depositaria.	Biblioteca Nacional, Biblioteca del Congreso de la Unión y Biblioteca de México.	Cámara de Diputados, LXV Legislatura/Presidencia de la República.

<i>Estatuto de la Organización Técnica y Administrativa y del Servicio de Carrera</i> (2000)	Su objetivo es formar funcionarios de alta calidad que apoyen profesional y eficazmente el cumplimiento de las atribuciones y funciones de la Cámara.	Documento que establece que el personal debe tener estudios profesionales relacionados con el área que desempeña, así como garantizar su permanencia y desarrollo profesional dentro de la Cámara de Diputados.	Personal administrativo, operativo y de confianza de la Cámara de diputados.	Cámara de Diputados
<i>Estatuto del Servicio Civil de Carrera de la Cámara Alta</i> (2002)	Su objetivo es establecer las bases para la planeación, organización, operación, desarrollo, formación y capacitación, profesionalización, control y evaluación del Servicio Civil de Carrera de la Cámara de Senadores.	Documento que establece que el personal debe tener estudios profesionales relacionados con el área que desempeña, así como garantizar su permanencia y desarrollo profesional dentro de la Cámara de Senadores.	Personal administrativo, operativo y de confianza del Senado de la República.	Senado de la República

Tabla 1. Políticas de información bibliotecarias explícitas a nivel *macro*/institucional
Elaboración propia, con base en la Ley Orgánica del Congreso General, Decreto de Depósito Legal y Estatutos de Servicio Civil de Carrera

En el nivel *mezzo* u organizacional, está el reglamento que elaboró el H. Congreso de la Unión, pero esta vez en calidad de ordenamiento organizacional (*Tabla 2*).

Políticas de información bibliotecarias	Descripción	Características	Dirigida a:	Elabora / Aprueba
<i>Reglamento del Sistema de Bibliotecas del Congreso de la Unión</i> (2019)	Su objetivo es normar el funcionamiento del Sistema de Bibliotecas del Congreso de la Unión. Tiene el propósito de ejecutar las directrices bibliotecológicas para los servicios, así como coordinar, normar y supervisar las actividades bibliotecarias y establecer los vínculos de cooperación con los Congresos Estatales, así como con otros sistemas bibliotecarios internacionales.	Establece las bases de operación de las bibliotecas pertenecientes al Sistema de Bibliotecas del Congreso de la Unión, así como el perfil y funciones del director del sistema.	Biblioteca Legislativa, Biblioteca General y Biblioteca del Senado de la República.	Cámara de Diputados, LXIV Legislatura

Tabla 2. Políticas de información bibliotecarias explícitas a nivel *mezzo*/organizacional

En el nivel *micro* o local, se encuentran las elaboradas o destinadas para las bibliotecas (*Tabla 3*).

Políticas de información bibliotecarias	Descripción	Características	Dirigido a:	Elabora/ Aprueba
<i>Manual General de Organización de la Cámara de Diputados</i> (2009)	Su objetivo es contar con un instrumento administrativo de consulta, que les permita conocer... su estructura orgánica y los objetivos y funciones asignados a cada una de las Unidades Administrativas que la conforman.	Documento que establece la organización y funciones de las áreas y servicios administrativos de la Cámara de Diputados.	Áreas, servicios y personal de la Cámara de Diputados.	Cámara de Diputados, LX Legislatura
<i>Lineamientos para la prestación de los servicios bibliotecarios y documentales de la Cámara de Diputados</i> (2009)	Tienen por objeto regular la prestación de los servicios bibliotecarios y documentales que son responsabilidad de la Dirección General del Centro de Documentación, Información y Análisis.	Documento que menciona las colecciones, servicios y disposiciones para el uso y préstamo del acervo.	Biblioteca Legislativa y General de la Cámara de Diputados. Áreas de Desarrollo de Colecciones y Servicios de Información.	Cámara de Diputados, LX Legislatura
<i>Manual de procedimientos de los Servicios de Información y Análisis de la Cámara de Diputados</i> (2008)	Ofrece información sobre la forma en que proporciona servicios especializados relacionados con los recursos documentales y la elaboración de análisis, consultas e investigaciones documentales relacionadas con el trabajo legislativo.	Incluye los procedimientos para la selección, clasificación y ordenamiento de material documental; para el otorgamiento de información especializada y para el préstamo de material documental.	Biblioteca Legislativa y General de la Cámara de Diputados. Áreas de Desarrollo de Colecciones y Servicios de Información.	Cámara de Diputados, LX Legislatura
<i>Manual de procedimientos administrativos en materia de adquisición, catalogación, préstamo y divulgación del acervo bibliográfico</i> (2013)	Establece las políticas y secuencia de actividades para la administración de los recursos bibliográficos y documentales del acervo patrimonial y su divulgación, ofrece servicios de información a los usuarios internos y externos.	Instrumento normativo que establece los procedimientos necesarios para enriquecer y controlar el acervo bibliográfico, atender solicitudes de información y divulgación del acervo.	Biblioteca "Melchor Ocampo" del Senado de la República. Áreas de Desarrollo de Colecciones, Análisis documental, Servicios de información y divulgación.	Senado de la República, LXII Legislatura, Secretaría General de Servicios Administrativos, Unidad de Modernización Administrativa.

Tabla 3. Políticas de información bibliotecarias explícitas a nivel *micro*/individual
Elaboración propia, con base en los Lineamientos, Manuales de Organización y Procedimientos del Sistema de Bibliotecas del Congreso General de México

A continuación, se presentan los resultados de las entrevistas de acuerdo con las áreas en que se dividió el cuestionario:

1. *Personal*: el personal responsable de las bibliotecas General y Legislativa tienen cargo de: responsable de biblioteca; en la biblioteca del Senado es: subdirector de biblioteca. Los tres nombramientos son de confianza. Los responsables de las bibliotecas General y Legislativa tienen una carrera distinta a bibliotecología. El responsable en la biblioteca del Senado es doctor en bibliotecología.

La cantidad de personal es la siguiente: en la biblioteca Legislativa, 23, la General, 21 y la del Senado, 9. La biblioteca Legislativa y la del Senado cuentan con dos personas con estudios de licenciatura en bibliotecología cada una, el resto del personal en las tres bibliotecas cuenta con estudios de bachillerato y estudios técnicos, los nombramientos son: Técnico/Operativo y Técnico/Especializado. Reciben 40 horas de capacitación al año, pero en áreas distintas a la bibliotecología.

La biblioteca General y la Legislativa dependen de la Secretaría de Servicios Parlamentarios y reportan a la Subdirección de Biblioteca. La biblioteca del Senado depende de la Secretaría General de Servicios Parlamentarios y reporta al Comité de Biblioteca y Asuntos Editoriales.

No presentaron políticas de información para esta área.

2. *Desarrollo de colecciones y presupuesto*: las tres bibliotecas cuentan con libros de temas en general, diarios de sesiones y boletines parlamentarios, gacetas parlamentarias, publicaciones oficiales de gobierno, legislación y jurisprudencia, diccionarios y enciclopedias, revistas académicas, periódicos y colecciones especiales. Sobre el presupuesto, la biblioteca General y la Legislativa no cuentan con uno; la forma en que adquieren sus materiales es a través del Depósito Legal. Con respecto al acervo, tienen un aproximado de 122 000 y 123 000 volúmenes respectivamente. A la biblioteca del Senado en 2019 le asignaron 50 000 pesos m.n.; adquiere sus materiales por medio de la donación y la compra; la cantidad aproximada de acervo es de 25 000.

La biblioteca del Senado presentó sus políticas de información para desarrollo de colecciones. Las de la Cámara Diputados no presentaron ninguna; a las tres les hace falta una política que incluya el presupuesto.

3. *Organización documental*: las bibliotecas de la Cámara de Diputados no realizan esta actividad, utilizan Janium para la consulta de su catálogo. En la biblioteca del Senado sí se realiza esta actividad. Cuentan con las siguientes

políticas: Formato MARC, las listas de encabezamientos de materia de Gloria Escamilla y tesauros especializados. Para la clasificación utilizan el sistema decimal de Dewey; además, cuentan con el *Manual de procedimientos administrativos en materia de adquisición, catalogación, préstamo y divulgación del acervo bibliográfico*, donde se especifica cómo se deben realizar los procesos técnicos y físicos del material. Las tres bibliotecas están trabajando para que se unifiquen sus catálogos en uno solo.

4. *Servicios de información*: las bibliotecas cuentan con los siguientes servicios: préstamo de libros en sala, externo e interbibliotecario; servicio de consulta; servicio de estudios o de investigación; consulta de catálogos; servicio de alerta y servicio de reprografía, además de contar con servicios en línea como: servicio de consulta, servicio de estudios o de investigación, consulta de catálogos, servicio de alerta, consulta a bases de datos en línea. En las bibliotecas General y del Senado la estantería es cerrada. La biblioteca Legislativa cuenta con estantería mixta.

La biblioteca del Senado presentó el *Manual de procedimientos administrativos en materia de adquisición, catalogación, préstamo y divulgación del acervo bibliográfico*. Las bibliotecas de la Cámara de Diputados no presentaron políticas de información.

5. *Cooperación bibliotecaria institucional*: las tres bibliotecas cuentan con el servicio de préstamo interbibliotecario y pertenecen a la Red de Bibliotecas del Centro Histórico de la Ciudad de México.

No presentaron políticas bibliotecarias para esta área.

6. *Infraestructura*: la biblioteca general necesita más espacio, pero las condiciones del edificio no le permiten crecer, debido a que es un edificio histórico. La biblioteca Legislativa también necesita más espacio, tienen una propuesta de ampliación; además necesitan más estantería, computadoras, impresoras, escáner y mejorar la conexión a la red Wifi. La biblioteca del Senado está esperando la respuesta del Comité de biblioteca para saber si le pueden asignar un espacio en la sede nueva del Senado, pues actualmente se encuentra en otra dirección postal; necesita estantería compacta y mejor ancho de banda.

No presentaron políticas de información bibliotecarias para esta área.

DISCUSIÓN

1. *Personal*. La IFLA considera que se debe realizar una correcta selección para encontrar al candidato idóneo que cumpla con los conocimientos profesionales y técnicos que amerite el puesto. Capacitar continuamente y contar con oportunidades de desarrollo profesional dentro de la institución.

Aunque el personal responsable de las bibliotecas no mencionó algún documento normativo, tienen las directrices llamadas *Estatutos del servicio civil de carrera*, que señalan lo siguiente: deben contar con personal profesional de carrera, expertos en las diferentes áreas para realizar de manera eficaz y eficiente sus tareas, deben capacitarse continuamente para que se desarrollen profesionalmente dentro del Congreso.

2. Desarrollo de colecciones y presupuesto. La IFLA señala que deben realizar una adecuada selección de los materiales que formarán el acervo de sus colecciones y deben incluir tanto material impreso como audiovisual y digital. Para la adquisición propone se haga por compra, donación o depósito legal. Tener un presupuesto para la adquisición de acervo, ya que ni el depósito legal, ni las donaciones son suficientes para integrar una colección adecuada con las necesidades de los legisladores. El presupuesto debe contemplar dinero para capacitación del personal y compra de mobiliario y equipo de cómputo, así como para arreglos o mejoras en el edificio de la biblioteca.

Las bibliotecas General y Legislativa no mencionaron políticas de información, pero cuentan con: *Lineamientos para la prestación de los servicios bibliotecarios y documentales* y *Manual de Procedimientos de los Servicios de Información y Análisis*; en el Senado, el *Manual de procedimientos administrativos en materia de adquisición, catalogación, préstamo y divulgación del acervo bibliográfico*, donde se indican las colecciones necesarias para prestar el servicio. Relativo al presupuesto, falta una directriz o regulación explícita que dicte la cantidad de presupuesto para la compra de material bibliohemerográfico, audiovisual y digital, así como una partida especial para la capacitación del personal en actividades bibliotecarias y para la compra de mobiliario y equipo de cómputo.

3. *Organización documental*. La IFLA recomienda que el personal que labora en esta área debe tener ciertas capacidades, tales como: capacidad de organizar y clasificar el material, con el fin de facilitar su almacenamiento y posterior recuperación. Solamente la biblioteca del Senado lleva a cabo esta actividad; en la entrevista presentaron su *Manual de procedimientos administrativos en*

materia de adquisición, catalogación, préstamo y divulgación del acervo bibliográfico; el personal encargado de esta área tiene estudios en bibliotecología.

4. *Servicios de información.* La IFLA propone los siguientes servicios, como mínimo: préstamo de libros, de CD-ROMs y de microfichas; servicios de documentación; servicios de estudios; sistemas de información parlamentaria; publicaciones, conservación, archivo; consultas individuales: en persona, por teléfono, correo electrónico, formulario en la página web, por escrito y mensajes de texto, además de servicios en línea.

Durante las entrevistas, únicamente la biblioteca del Senado presentó sus políticas de información; sin embargo, las bibliotecas de la Cámara de Diputados también cuentan con los lineamientos y manuales antes expuestos, en los que se establecen los servicios necesarios para atender las necesidades de sus usuarios, tanto tradicionales como en línea.

5. *Cooperación bibliotecaria institucional.* La IFLA recomienda que se realicen convenios de cooperación con bibliotecas de la misma especialidad.

No señalaron que contaran con alguna política de información, no obstante, en el *Reglamento del Sistema de Bibliotecas del Congreso de la Unión*, se especifica que deben establecer vínculos de cooperación con otras bibliotecas, incluyendo las bibliotecas de los Congresos de los Estados.

6. *Infraestructura.* La IFLA recomienda que deben contar con una buena infraestructura que sea capaz de aprovechar las ventajas de las TIC. El elemento clave para el éxito es el desarrollo de un plan estratégico que cuente con el respaldo y el apoyo de los líderes institucionales y políticos.

Hacen falta políticas de información explícitas para esta área; deben elaborar unas que incluyan aspectos relacionados con el espacio físico necesario en metros cuadrados por cada usuario, espacio para la estantería, áreas de oficina, así como mesas de estudio para los usuarios, entre otros elementos. Que satisfagan las condiciones óptimas de iluminación, ventilación y medio ambiente.

En resumen, las tres bibliotecas cuentan con directrices y regulaciones explícitas a nivel *micro* en: desarrollo de colecciones, servicios de información y préstamo interbibliotecario. En organización documental, solamente la biblioteca del Senado. La diferencia es que en las bibliotecas de la Cámara de Diputados las desconocen y por eso no las utilizan, en cambio en la biblioteca del Senado sí las conocen y son utilizadas.

CONCLUSIONES

Se identificaron las áreas, servicios, colecciones y personal de las bibliotecas del H. Congreso de la Unión. Las tres bibliotecas cumplen con lo que propone la IFLA en cuanto a: desarrollo de colecciones, servicios de información, organización documental, cooperación bibliotecaria institucional e infraestructura. Se identificaron sus políticas de información en los tres niveles de organización: organizacional, institucional y local o individual.

Se destacan los siguientes hallazgos: respecto al personal se encontró que, de los tres responsables de estas bibliotecas, solamente uno cuenta con estudios en bibliotecología. Relativo al presupuesto, únicamente la biblioteca del Senado tiene asignada una cantidad, lo que limita el desarrollo de las áreas de la biblioteca, sobre todo la compra de material bibliohemerográfico, la capacitación del personal y la creación de nuevos servicios.

En cuanto a infraestructura, la biblioteca Legislativa tiene un proyecto de remodelación o ampliación. La biblioteca General, aunque necesita más espacio, no puede modificar la estructura del edificio pues es un monumento histórico. La biblioteca “Melchor Ocampo” de la Cámara Alta, está buscando que se le otorgue un espacio dentro del edificio nuevo del Senado de la República, pues actualmente atiende en otro inmueble.

Respecto a las políticas de información explícitas a nivel *macro*, únicamente mencionaron el decreto de depósito legal; a nivel *mezzo*, no nombraron el *Reglamento del Sistema de Bibliotecas*, y a nivel *micro* sí cuentan con ellas en la Cámara de Diputados, pero los responsables de las bibliotecas las desconocen, motivo por el cual no las señalaron. La biblioteca del Senado sí presentó sus políticas, pero ambos casos deben agregar los aspectos de personal, presupuesto e infraestructura.

Finalmente, se recomienda que elaboren las políticas de información que les hacen falta, y que implementen, evalúen y mejoren las ya existentes. Esto les permitirá identificar las actividades no controladas y, al mismo tiempo, servirá para elevar la calidad de los servicios, proyectando áreas de oportunidad y proponiendo soluciones a los problemas cotidianos.

REFERENCIAS

- Almada Navarro, Elisa Margarita. 2012. “Bases teóricas para comprender las políticas de información”. En *La naturaleza objetiva y subjetiva de las políticas de información*, coordinado por Egbert J. Sánchez Vanderkast, 3-26. México: UNAM, IIBI.
- Anglada, Lluís. 2014. “De qué hablamos cuando hablamos de políticas de información”. *El Profesional de la Información* 23 (2): 105-112.

- Braman, Sandra. 2011. "Defining information policy". *Journal of Information Policy* 1: 1-5. doi: 10.1080/13614533.2015.1009126
- Cagnoli, Roberto V. 1996. *Administración de bibliotecas*. Buenos Aires: EB Publicaciones.
- Cantú Delgado, José Humberto. 2011. *Desarrollo de una cultura de calidad*, 4a. ed. México: McGraw-Hill.
- Case, Donald O. 2010. "A framework for information policies with examples for United States". *Library Philosophy and Practice* (September): 1-10.
- Chiavenato, Idalberto. 2006. *Introducción a la teoría general de la administración*, 17a. ed. México: McGraw-Hill.
- Colunga Dávila, Carlos. 1996. *La administración del tercer milenio: diccionario temático*. México: Panorama.
- Cunninghame, Keith. 2009. *Directrices para bibliotecas parlamentarias*. IFLA, De Gruyter Saur. Recuperado de: Directrices para bibliotecas parlamentarias31012013.doc (ifla.org)
- Drovetta, Mabel Susana. 1995. *Diccionario de administración y ciencias afines*. México: Limusa.
- Fernández Bajón, María Teresa, y Esperanza Martínez Montalvo. 2006. "Políticas de gestión de calidad en los centros de información y documentación". En *Seminario Hispano-Mexicano de investigación en bibliotecología y documentación en México y España*, compilado por Filiberto Felipe Martínez Arellano y Juan José Calva González, 417-429. México: UNAM, CUIB.
- Jaeger, Paul T., Ursula Gorham, Natalie Greene Taylor, y John C. Berlot. 2015. "Teaching information policy in the digital age; issues, strategies and innovation". *Journal of Education for Library and Information Science*, 56 (3, Summer, July): 175-189.
- Jones, Gareth R., y Jennifer M. George. 2010. *Administración contemporánea*, 2a. ed. México: McGraw-Hill.
- México. Cámara de Diputados, LV Legislatura. 1991. *Decreto de Depósito Legal*. Recuperado de: Biblioteca Legislativa - Cámara de Diputados del H. Congreso de la Unión.
- México. Cámara de Diputados, LVIII Legislatura. 2000. *Estatuto de la Organización Técnica y Administrativa y del Servicio de Carrera*. Recuperado de: Estatuto de la organización técnica y administrativa y del servicio de carrera de la cámara de diputados.
- México. Cámara de Diputados, LX Legislatura. 2008. *Manual de procedimientos de los Servicios de Información y Análisis de la Cámara de Diputados*. Recuperado de: compendio_02_dic2008.pdf (congreso.gob.mx)
- México. Cámara de Diputados, LX Legislatura. 2009. *Lineamientos para la prestación de los servicios bibliotecarios y documentales de la Cámara de Diputados*. Recuperado de: LIPRESByD_CD.pdf (diputados.gob.mx)
- México. Cámara de Diputados, LX Legislatura. 2009. *Manual General de Organización de la Cámara de Diputados*. Recuperado de: Manual General de Organización de la Cámara de Diputados. Gaceta Parlamentaria 27-03-2008
- México. Cámara de Diputados, LXIV Legislatura. 2019. *Reglamento del Sistema de Bibliotecas del Congreso de la Unión*. Recuperado de: http://www.diputados.gob.mx/sedia/Reg_SBCU_orig_14may19.pdf

- México. Cámara de Diputados, LXV Legislatura. 2021. *Ley Orgánica del Congreso General de los Estados Unidos Mexicanos*. Recuperado de: Ley Orgánica del Congreso General de los Estados Unidos Mexicanos (diputados.gob.mx)
- México. Cámara de Diputados, LXV Legislatura. *Decreto por el que se expide la Ley General de Bibliotecas*. Recuperado de: Ley General de Bibliotecas (diputados.gob.mx)
- México. Cámara de Senadores, LVIII Legislatura. 2002. *Estatuto del Servicio Civil de Carrera de la Cámara Alta*. Recuperado de: DOF - Diario Oficial de la Federación
- México. Cámara de Senadores, LXII Legislatura. 2013. *Manual de procedimientos administrativos en materia de adquisición, catalogación, préstamo y divulgación del acervo bibliográfico*.
- Pasek, Judith E. 2015. "Defining information policy: relating issues to the information cycle". *New Review of Academic Librarianship*, 21 (3): 286-303.
- Rosenberg, Jerry Martin. 2004. *Diccionario de administración y finanzas*. Barcelona: Océano: Centrum.
- Sánchez Vanderkast, Egbert J. 2012. "Algunas reflexiones metodológicas para el estudio sobre políticas de información". En *La naturaleza objetiva y subjetiva de las políticas de información*, coordinado por Egbert J. Sánchez Vanderkast, 33-48. México: UNAM, IIBI.
- Umar, Yila Musa, y Olatunji Azeez Bakare. 2018. "Assessment of collection development policy in academic libraries: case study of National Open University of Nigeria". *The Information Technologist: an International Journal of Information, and Communication Technology (ICT)*, 15 (1): 164-170.
- Washington, Anne L. 2014. "Government information policy in the era of big data". *Review of Policy Research*, 31 (4): 319-325.
- Whitten, Jeffrey y Lonnie D. Bentley. 2008. *Análisis de sistemas: diseño y métodos*, 7a. ed. México: McGraw-Hill.

Agradecimientos

Esta investigación se ha llevado a cabo gracias al apoyo y respaldo del Consejo Nacional de Ciencia y Tecnología (CONACYT) a través de una beca de posgrado. A los responsables de las bibliotecas del Congreso General de México por su participación en las entrevistas y al personal de las bibliotecas por su amabilidad.

Para citar este texto:

Vargas Huerta, Arturo. 2022. "Políticas de información en las bibliotecas del Congreso General de los Estados Unidos Mexicanos". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 115-134.
<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58585>

Anexo 1

Cuestionario para entrevistar a los responsables de las bibliotecas del Congreso General de los Estados Unidos Mexicanos

Objetivo: Identificar las directrices y regulaciones bibliotecarias con las que cuentan las bibliotecas del H. Congreso de la Unión, en qué servicios se aplican, de qué nivel y de qué tipo son.

I. Personal

1. ¿Cuántas personas laboran en su Biblioteca?
2. ¿Qué escolaridad tienen?
3. Según su catálogo de puestos, ¿cuál es el nombramiento del personal que labora en su Biblioteca?
4. ¿Qué necesidades de capacitación ha detectado entre su personal?
5. ¿Con qué frecuencia participa el personal de la Biblioteca en actividades de actualización y capacitación?
6. ¿Cuentan con algún documento que indique el perfil que debe cumplir el personal de la biblioteca? Si es así ¿Cómo se llama?

II. Desarrollo de Colecciones

7. ¿Qué colecciones conforman su acervo impreso?
8. ¿Qué colecciones conforman su acervo audiovisual y digital?
9. Si cuenta con bases de datos en línea ¿cuáles son sus nombres?
10. ¿Cuántos documentos tiene cada base de datos en línea?
11. ¿De cuántos títulos y ejemplares constituye la colección completa de la Biblioteca?
12. ¿De qué manera se incrementa su acervo?

13. ¿Cómo evalúan sus colecciones?
14. ¿Con qué periodicidad evalúan sus colecciones?
15. ¿Cuentan con algunos de los siguientes documentos para mantener un acervo acorde a las necesidades de su comunidad?
 - Programas de desarrollo de colecciones
 - Políticas de compra, canje o donación
 - Políticas de descarte
 - Ley o Decreto de Depósito Legal
16. ¿La Biblioteca cuenta con un presupuesto asignado?
17. ¿De dónde proviene el presupuesto?
18. ¿A cuánto asciende ese presupuesto o podría darme un estimado?
19. ¿De qué manera se ejerce ese presupuesto?
20. ¿Cuentan con un documento que especifique la cantidad de presupuesto que se debe asignar a la biblioteca? Si es así ¿Cómo se llama?

III. Organización documental (Procesos Técnicos)

21. Para realizar la catalogación o descripción bibliográfica del acervo, ¿qué normatividad utilizan?
22. ¿Qué formatos utilizan para la descripción catalográfica?
23. ¿Qué sistema de clasificación utilizan?
24. ¿Qué listas de encabezamientos de materia o tesauros utilizan para la asignación de temas?
25. ¿Qué sistema utilizan para la automatización de su biblioteca?
26. ¿El sistema que tienen les sirve para compartir registros?

27. ¿Cuentan con alguno de los siguientes documentos para realizar sus procesos técnicos?
- Manual de Catalogación
 - Procedimiento de catalogación
 - Políticas de catalogación

IV. Servicios de Información

28. ¿Cuáles son los servicios que ofrece la Biblioteca de manera tradicional?
29. ¿Cuáles son los servicios que se ofrecen en línea?
30. ¿De qué medios se vale para la difusión de los servicios?
31. ¿Su estantería es?
32. ¿Existe un Reglamento general aplicable en la Biblioteca?
33. ¿Cuál es la cantidad de usuarios internos y externos del Centro de Información?
34. ¿Cuántos usuarios atienden al año aproximadamente?
35. De cinco años a la fecha ¿han desarrollado nuevos servicios?
36. ¿Cuentan con políticas, manuales, reglamentos, guías u otros documentos normativos para los siguientes servicios?
- Catálogo Automatizado
 - Servicio al Público
 - Uso de Estantería
 - Uso de las computadoras (en caso de que las tenga)
37. Si están en formato impreso ¿Me podría regalar copia de esos documentos?
38. ¿Cómo determinaron que se tenían que realizar?
39. ¿Cuentan con políticas, manuales, reglamentos, guías u otros documentos normativos para fomentar el uso de los siguientes servicios bibliotecarios?

- Cursos y talleres sobre el uso de la biblioteca
- Visitas guiadas
- Elaboración de bibliografías y citas
- Cursos y talleres sobre bases de datos

V. Cooperación bibliotecaria institucional

40. ¿Proporcionan el servicio de Préstamo Interbibliotecario?
41. ¿Con cuáles bibliotecas tienen ese convenio?
42. ¿Bajo qué reglas o normatividad prestan el servicio de Préstamo Interbibliotecario?
43. Si no tienen Préstamo Interbibliotecario ¿cómo hacen para intercambiar información con otras bibliotecas en caso de que un usuario@ necesite un libro con el que no cuente?
44. ¿Pertenece a alguna red de cooperación bibliotecaria?
45. ¿Cuál es su nombre o denominación?
46. ¿Quiénes la integran?
47. ¿Qué beneficios han encontrado al pertenecer a esa red?
48. ¿Qué documentos normativos han resultado de esta red?
49. ¿Qué trabajos o proyectos se han desarrollado en esta red?

VI. Infraestructura (Edificio, mobiliario, equipo)

50. ¿Qué necesidades de infraestructura en cuanto al edificio han detectado?
51. ¿Qué necesidades de mobiliario han detectado?
52. ¿Qué necesidades en cuanto a equipo de cómputo y telecomunicaciones han detectado?

53. ¿Qué necesidades en cuanto a ventilación e iluminación han detectado?
54. ¿A través de qué medio detectaron esas necesidades de infraestructura?
- Diagnóstico de la biblioteca
 - Encuesta a usuarios
 - Buzón de quejas y sugerencias
 - Solicitudes expresas de los usuarios
 - Ejemplos de infraestructura de otras bibliotecas
55. ¿Cómo han resuelto esas necesidades de infraestructura?
- Proyecto de edificio nuevo para la biblioteca
 - Plan para remodelación y/o modernización de las instalaciones
 - Programa de compra de mobiliario, equipo de cómputo y telecomunicaciones
56. Si se llevó a cabo algún plan, proyecto o programa, ¿quién aportó los recursos?

Políticas editoriales de revistas de Comunicación en acceso abierto indexadas en Scopus

Bexi Perdomo*
Oscar Alberto Morales**

Artículo recibido:
18 de mayo de 2022
Artículo aceptado:
19 de agosto de 2022

Artículo de investigación

RESUMEN

Los autores de artículos necesitan información acerca de ciertas políticas editoriales (PE) de las revistas para elegir la que mejor responda a sus necesidades. *Objetivo:* caracterizar las tendencias en PE en las revistas de acceso abierto (AA) catalogadas en el área de Ciencias Sociales (CCSS) y en la subárea de Comunicación, en Scopus. *Métodos:* ésta fue una investigación descriptiva y transversal, apoyada en la bibliometría en la que se analizaron 77 revistas. *Resultados:* se observa el predominio de revistas con periodicidad semestral y del idioma inglés como lengua considerada por 90% de las

- * Centro de Investigación de la Creatividad, Universidad de Ciencias y Artes de América Latina, Perú
bjperdomo@crear.ucal.edu.pe
- * Facultad de Odontología, Universidad de Los Andes, Venezuela
geode.ula@gmail.com

revistas. Si bien la tendencia es normar la extensión de los artículos, ésta suele variar según género publicado y se mide en caracteres con espacios, palabras y páginas. En atención al cargo por procesamiento para AA, 55,8% no lo exige. *Conclusión:* las PE de estas revistas de AA en Scopus se traducen en ofrecer posibilidades de visibilidad a los autores y fortalecer la investigación de forma colaborativa.

Palabras clave: Acceso abierto; Políticas editoriales; Comunicación; Publicación científica

Editorial policies of Communications in Open Access journals indexed in Scopus

Bexi Perdomo and Oscar Alberto Morales

ABSTRACT

Authors need of information about the editorial policies (EP) of journals to choose the one that best meets their needs. *Aim:* To characterize the trends in the EP in the open access (OA) journals cataloged in the area of Social Sciences (CCSS) and in the sub-area Communication in Scopus. *Methods:* This research was descriptive, cross-sectional, and supported by bibliometrics for analyzing 77 journals. *Results:* it was observed the predominance of journals published every six months, and that English is the language considered by 90% of the journals for publishing. Although the trend is to standardize the length of the articles, there are variations by gender, and it is measured in characters with spaces, words, and pages. Regarding the processing charge for OA, 55.8% do not require it. *Conclusions:* It is concluded that the EP of these OA journals in Scopus translates into offering possibilities of visibility to authors and reinforce collaborative research.

Keywords: Open Access; Editorial policies; Communication; Scientific publication

INTRODUCCIÓN

La publicación de un artículo científico supone múltiples tareas, incluyendo la elección de la revista idónea considerando el género discursivo, la audiencia y el nivel de especialización. Esta elección puede hacerse incluso desde que se inicia la escritura del artículo. En un contexto con criterios de evaluación cada vez más rigurosos para los investigadores, para tomar esta decisión es fundamental considerar algunos aspectos editoriales como la visibilidad, la periodicidad, la indexación y el costo de publicación (Pérez, Herrera y Ramírez 2020). Para ello, la bibliometría ofrece evidencia científica estadística de la actividad científica desde diferentes perspectivas (Alcalá-Albert y Parra-González 2021), permitiendo la valoración de la producción científica de áreas de conocimiento, autores, instituciones, revistas, bases de datos, países, regiones y continentes (González, Cardentey, y Izquierdo 2017; Madrid et al. 2017). En la disciplina de la comunicación, se han realizado algunos análisis bibliométricos de la producción científica (De-Filippo 2013; Arroyave-Cabrera y González-Pardo 2022), pero no se han publicado estudios que orienten a los autores sobre las tendencias de las revistas en materia de políticas editoriales (PE), que les permitan tomar decisiones con base en la evidencia.

Las PE de las revistas científicas definen sus criterios editoriales y describen el proceso que siguen para la publicación, desde la recepción de los manuscritos hasta su divulgación. Entre los aspectos considerados, están: los géneros discursivos que publica (tipos de artículos: artículo de investigación, revisiones, ensayos, reseñas, entre otros), idiomas, periodicidad, normas para citación y referencias, y las políticas para el acceso abierto (AA). Estudios previos han evidenciado que algunas de estas PE afectan las métricas de las revistas, autores e instituciones, entre otras razones porque el idioma y la extensión están asociados a la cantidad de citas recibidas (Hasan y Breunig 2021; Henshall 2018). Esto puede explicar el interés incipiente por el estudio de algunos de los aspectos incluidos en las PE, como la periodicidad, la autoría y el idioma (Felipe et al. 2020; Silva y Blattmann 2018), así como los tipos de publicaciones (Tippu y Ryan 2022) y la migración al AA (Bojo-Canales y Melero 2021).

La publicación en AA es una manera de divulgar la información científica sin costo o restricciones para los lectores (Araiza Ramírez y Díaz 2019). Esta política es producto de una serie de iniciativas y acuerdos internacionales orientados a la democratización del conocimiento científico. Sin embargo, un importante número de revistas se mantiene con acceso restringido o híbrido (con artículos en AA, bajo ciertas condiciones). En algunas revistas, es discrecional para los autores si el artículo se publica en AA, para lo cual deben pagar un monto por el procesamiento de sus manuscritos una vez que han sido evaluados por los pares y han sido aceptados para su publicación.

Los estudios bibliométricos están dirigidos predominantemente a la producción científica sobre áreas o temas particulares (Alcalá-Albert y Parra-González 2021; De-Filippo 2013; Arroyave-Cabrera y González-Pardo 2022; Lopera-Pérez et al. 2021) y analizan la productividad de las revistas por países, regiones o temáticas (Estrada-Cuzcano y Alhuay-Quispe 2020). No obstante, sus resultados también podrían servir para sistematizar las PE de las revistas a partir de los artículos que publican. Este aspecto no ha recibido la atención necesaria de parte de los investigadores, pese a su pertinencia. La revisión de la literatura sugiere la necesidad de realizar estudios que muestren las tendencias en PE actuales de las revistas de AA en diferentes áreas de conocimiento, incluyendo la de comunicación.

El área de comunicación, la cual figura en Scopus como una sub área de las Ciencias Sociales (CCSS), abriga un importante número de revistas científicas en esta base de datos (427 en el listado disponible a inicios de marzo de 2022), de las cuales 93 se identificaron como de AA. No obstante, no se ha publicado un estudio que analice la información acerca de las PE para el envío, el procesamiento y la publicación de los artículos, que les permita a los autores conocer mejor lo que las revistas esperan de ellos y las opciones que ofrecen. En este sentido se planteó esta investigación, cuyo objetivo fue caracterizar las tendencias en PE en revistas de AA etiquetadas en la sub área de comunicaciones en Scopus. De esta forma, se espera ofrecer a los autores evidencia científica relevante y útil para la preparación de sus manuscritos y para la elección de revistas de acceso abierto.

METODOLOGÍA

Se realizó un estudio descriptivo y transversal, empleando herramientas de la bibliometría, similar al de Miguel, Tannuri y Cabrini (2016). Se analizaron las revistas de AA indexadas en Scopus, catalogadas en el área de CCSS y en la sub área de comunicación. Se consideraron las revistas cuyas métricas estuvieran disponibles en el momento de la recolección y análisis de la información (marzo de 2022). Inicialmente, se identificaron 93 revistas. De éstas, se descartaron 16 que no presentaban información completa de sus métricas (por ejemplo, número de documentos publicados y citas), por lo que la muestra final fue de 77 revistas.

Para el procesamiento de la información se creó una ficha de observación en Microsoft Excel a partir de la operacionalización de la variable PE (ver *Tabla 1*). Ésta se delimitó en las dimensiones de interés para el estudio: periodicidad, idiomas, géneros publicados, autores, extensión, normas de estilo y

cargos por APC (*Article Processing Charges*, por sus siglas en inglés). La operacionalización y la ficha fueron validados, sin observaciones, por tres investigadores con experiencia en estudios bibliométricos. El análisis se llevó a cabo con la ayuda del *software* IBM SPSS (versión 26).

Variable	Definición conceptual	Dimensión	Definición de la dimensión	Indicadores
Política editorial	Acuerdos de la revista que marcan las pautas a las que los autores se sujetan para que sus manuscritos sean procesados para su publicación.	Periodicidad	Frecuencia con que se publica la revista.	<ul style="list-style-type: none"> • Anual • Semestral • Cuatrimestral • Trimestral • Bimestral • Mensual • Publicación continua • Otras
		Géneros discursivos publicados	Hace alusión a los tipos de artículos que se publican por ser de interés para la comunidad académica a la cual se dirige.	<ul style="list-style-type: none"> • Artículo de investigación • Revisiones de literatura o estudios teóricos • Revisiones sistemáticas (y meta-análisis) • Estudios de caso • Reseñas • Cartas al editor • Ensayos • Otros (editoriales, entrevistas, por ejemplo)
		Idiomas	Lenguas en las que se publican los manuscritos.	<ul style="list-style-type: none"> • Inglés • Español • Portugués • Francés • Otros
		Autores	Normas editoriales relacionadas con la coautoría por artículo y frecuencia de publicación, por autor, en la revista.	<ul style="list-style-type: none"> • Número máximo de autores por artículo. • Tiempo de espera para enviar una nueva propuesta
		Extensión	Espacio que ocupará el manuscrito. Puede medirse en caracteres, palabras o páginas.	<ul style="list-style-type: none"> • Mínimo y máximo de caracteres, palabras o páginas del artículo

		Estilo de referencias	Normas o sistemas para presentar las citas y las referencias.	<ul style="list-style-type: none"> • APA • Harvard • Chicago • Vancouver • IEEE • Otros
		Cobro de APC	Monto que paga el autor para cubrir los cargos para publicación de su artículo en AA.	<ul style="list-style-type: none"> • Monto de APC

Tabla 1. Operacionalización de la variable de estudio

RESULTADOS

Se analizaron 77 revistas científicas de comunicación. Éstas se ubican en los cinco continentes: Europa (n = 55), América (n = 16), Asia (n = 4), África (n = 1) y Oceanía (n = 1). Para el momento de la búsqueda, se habían publicado 7 336 documentos con 11 862 citaciones entre 2017 y 2022.

Periodicidad

Los resultados se aprecian en la *Tabla 2*, la cual muestra la frecuencia (fi) y el porcentaje para cada periodicidad observada. La mayoría de las revistas (51,9%) se publica dos veces al año. Pocas de ellas asumen la modalidad de publicación continua; es decir, publicar inmediatamente el artículo una vez aprobado por los pares, sin esperar a que se haya completado el total de artículos contemplado para el número.

Periodicidad	fi	%
Anual	1	1,3
Cuatrimstral	5	6,5
Mensual	1	1,3
Publicación Continua	10	13,0
Semestral	40	51,9
Trimestral	20	26,0
Total	77	100

Tabla 2. Periodicidad de las revistas analizadas

Se analizó la periodicidad en función de los diferentes países donde se editan. Como se observa en la *Tabla 3*, predominan las revistas que se publican semestralmente, las cuales en su mayoría provienen de España, seguidas de las que aparecen trimestralmente, las cuales no mostraron una tendencia marcada por país.

País	Periodicidad						País	%
	Semestra	Mensual	Trimestral	Cuatrimestral	Anual	Continua		
Alemania	2	0	1	0	0	0	3	3.9
Austria	1	0	0	0	0	0	1	1.3
Brasil	2	0	2	1	0	1	6	7.8
Bulgaria	0	0	0	0	0	1	1	1.3
Canadá	1	0	1	0	0	0	2	2.6
Chile	0	0	0	1	0	0	1	1.3
China	0	0	1	0	0	0	1	1.3
Colombia	1	0	0	0	0	1	2	2.6
Corea del Sur	1	0	0	0	0	0	1	1.3
Croacia	3	0	0	0	0	0	3	3.9
EEUU	2	0	1	0	0	1	4	5.2
Eslovaquia	0	0	1	0	0	0	1	1.3
España	13	0	2	0	0	1	16	20.8
Finlandia	0	0	0	1	0	0	1	1.3
Hungría	0	0	1	0	0	0	1	1.3
Italia	1	0	1	0	0	0	2	2.6
Lituania	2	0	1	0	0	1	4	5.2
Malasia	2	0	0	0	0	0	2	2.6
Nueva Zelanda	1	0	0	0	0	0	1	1.3
Países Bajos	0	0	0	0	0	1	1	1.3
Perú	1	0	0	0	0	0	1	1.3
Polonia	1	0	2	0	0	1	4	5.2
Portugal	3	1	1	0	0	0	5	6.5
Reino Unido	2	0	2	1	0	1	6	7.8
Rep. Checa	0	0	1	0	0	0	1	1.3
Rusia	0	0	1	1	0	0	2	2.6
Serbia	0	0	0	0	1	0	1	1.3
Sudáfrica	0	0	0	0	0	1	1	1.3

Suiza	1	0	1	0	0	0	2	2.6
Total	40	1	20	5	1	10	77	100

Tabla 3. Periodicidad de las revistas en los diferentes países

Géneros publicados

Se publica una amplia gama de géneros discursivos (Tabla 4). El artículo de investigación, tradicionalmente estructurado en el formato IMRD (Introducción, Métodos, Resultados, Discusión) o su variación IMRDC (Introducción, Métodos, Resultados, Discusión, Conclusiones) está en casi el 100% de las revistas, seguido de las reseñas de libros (45,5%). La categoría “otros” comprende entrevistas, editoriales (por invitación) y otras publicaciones no arbitradas.

Género	fi	%
Artículo de investigación	75	97,4
Cartas al editor	1	1,3
Ensayo	12	15,6
Estudios de caso	5	6,5
Reseña de libros (videos y obras)	35	45,5
Revisión de la literatura	18	23,4
Revisión sistemática	10	13
Otros	57	74

Tabla 4. Géneros discursivos que publica la revista

Idioma

Casi el 50% de las revistas publica sólo en inglés (Tabla 5), pero si a éstas se suman las que publican en otros idiomas e inglés, la presencia de la *lingua franca* supera el 90%.

Idioma	fi	%
Español	4	5,2
Español y otro	5	6,5
Inglés	37	48,1
Inglés y otro	13	16,9
Plurilingüe	16	20,8
Portugués	1	1,3

Portugués y otro	1	1,3
Total	77	100,0

Tabla 5. Idiomas en que publican las revistas analizadas

Aquellas cuya lengua nacional no es el inglés (por ejemplo, portugués, español, francés) lo incluyen como una opción. Aun así, algunas exigen presentar la versión traducida al inglés una vez superada la revisión de pares y haber sido aceptado en la lengua original. Todas las revistas plurilingües incluyen este idioma como una opción.

Autores

Pocas revistas limitan la cantidad de autores (Tabla 6), pero todas exigen explicar cuál fue la participación de cada uno.

Autores	fi	%
Hasta dos	1	1.3
Hasta tres	2	2.6
Hasta cuatro	2	2.6
No indica	71	92.2
Sin límite	1	1.3
Total	77	100

Tabla 6. Indicaciones para cantidad de autores por artículo

En cuanto al tiempo que deben esperar los autores que han publicado para volver a someter un manuscrito a evaluación en la misma revista, la mayoría (87%) no plantea esta limitante (ver Tabla 7). De las pocas que tienen esta política, la mayoría no excede el año de espera entre la publicación de un artículo y la presentación de un nuevo manuscrito.

Espera para publicar	fi	%
Un año	6	7.8
Entre uno y dos años	4	5.2
No indica	67	87
Total	77	100

Tabla 7. Tiempo de espera para someter nuevos manuscritos en la misma revista

Extensión

Las revistas suelen indicar de forma explícita la extensión, discriminada según el género discursivo publicado (artículo de investigación, revisiones de literatura, ensayos, otros). De las 77 analizadas, 68 (88,3 %) presentan una política asociada a la extensión de las propuestas que presenten los autores (en caracteres, palabras o páginas); sólo nueve (11,7 %) no prescriben la extensión.

Más del 97% de las revistas publican artículos de investigación, el género al que se le permite la mayor extensión. En consecuencia, se tomó éste como referencia para presentar los rangos de extensión. Cuando se norma en función de caracteres, todas lo hacen contando “caracteres con espacios” (25 000 a 60 000). Por otra parte, en cuanto a palabras, el rango va desde las 4 000 hasta las 12 000 (siendo 8 000 el límite máximo más frecuente). Finalmente, algunas norman la extensión en función de las páginas (10-32) con diferentes indicaciones tipográficas (interlineado, tipo y tamaño de fuente, por ejemplo). Estos rangos comprenden desde el título hasta los anexos. Algunas de estas revistas presentan sus propias plantillas (*templates*, en inglés) para la organización y presentación de la información (por ejemplo, la revista *Media and Communication*).

Estilo para referencias

Esta dimensión hace referencia a las normas para la presentación de las citas en el texto y las referencias al final del manuscrito. Se observó variedad en las normas exigidas (*Tabla 8*), aunque predominan las normas de la American Psychological Association (APA). La categoría “otro” incluye aquellas revistas que hacen adaptaciones de alguna norma o sistema de referencias para crear su propio estilo y las que asumen otros sistemas de referencias como ISO 690 o GOST, por ejemplo.

Nor. - Ref.	fi	%
ABNT	3	3,9
APA	35	45,5
Chicago	5	6,5
Harvard	8	10,4
IEEE	1	1,3
MLA	1	1,3

Vancouver	1	1,3
Otros	22	28,6
Libre	1	1,3
Total	77	100

Tabla 8. Normas para referencias exigidas por las revistas

Cargos por procesamiento de los manuscritos (APC)

La mayoría de estas revistas (55,8 %), señalan que no cobran APC (ver *Tabla 9*), mientras que otras sólo hacen referencia a sus patrocinadores o a la iniciativa de acceso abierto de Budapest (*Budapest Open Access Initiative*), sin especificar si el acceso abierto será asumido por los autores. Algunas revistas que no cobran APC dan el crédito del AA a sus patrocinadores (generalmente universidades y entes gubernamentales asociados a la producción de ciencia y tecnología).

APC	fi	%
Sí	16	22.1
No	43	55.8
No indica	17	22.1
Total	77	100

Tabla 9. Cobro de APC para el acceso abierto

DISCUSIÓN

La investigación perseguía caracterizar las tendencias en PE en revistas de AA etiquetadas en la sub área de comunicación en Scopus, particularmente las relacionadas con: periodicidad, géneros discursivos publicados, idiomas, autoría, extensión, estilo para referencias y cobro de APC. Esas tendencias y algunas necesidades de investigación se discuten a continuación.

Aun cuando publiquen sólo en formato digital, las revistas prefieren la periodicidad semestral y trimestral. Sería interesante investigar si éstas han mantenido esa periodicidad y cantidad de artículos publicados en cada número o si las han incrementado con el tiempo. De igual forma, sería oportuno conocer las razones para mantenerse en periodicidad semestral, pues algunas revistas prestigiosas evitan expandir la cantidad de números y volúmenes por miedo a perder calidad o ser confundidas con revistas depredadoras que luego de incrementarse exponencialmente en unos pocos años han sido excluidas de las bases de datos por temas de ética y calidad (Zhang 2021). Ahora

bien, como señala Zhang (2021), si el crecimiento es consistente y no dramático, puede haber un efecto positivo en el factor de impacto.

En cuanto a la publicación continua, ésta ha resuelto un problema para autores y lectores, ya que una vez aceptado el artículo por los pares revisores está disponible en corto tiempo en formato *preprint* o en el formato ordinario permanente (Cruz 2008). Sin embargo, aún son pocas las revistas que ofrecen esta modalidad de publicación en AA en la sub área de comunicación. En estos casos, el autor debe “esperar desde que escribe el artículo hasta que se publica, lo cual hace que se vea afectado en sus procesos de evaluación, promociones, reconocimientos y compensaciones, a la vez que el artículo pierde vigencia” (Araiza, Ramírez y Díaz 2019, 203). Por ello, las revistas de publicación continua son de mayor conveniencia para quienes deben presentar evidencias de publicación a sus universidades o financiadores y quienes deben entregar un artículo publicado como parte de sus requisitos en estudios doctorales (Cruz 2008), y se perfila como una necesidad en las CCSS.

Conocer los géneros discursivos que se publican en una comunidad académica es importante, ya que existe evidencia acerca de la necesidad de su conocimiento y dominio para una exitosa incorporación en éstas (Perdomo 2021). Se observó la tendencia a publicar en diferentes géneros, con predominio del artículo de investigación. Este resultado es cónsono con quienes lo han reportado como el género fundamental para la difusión del conocimiento científico (Sánchez 2018; Meza 2018), el más prestigioso en el ámbito científico internacional (Swales 1990; Knorr-Cetina 1981) y el medio clave para legitimar el conocimiento y las disciplinas (Berkenkotter y Huckin 1995), mediante la comunicación de los hallazgos de sus investigaciones; entre otra razones, por “la inmediatez, la accesibilidad y la actualidad que conlleva en términos de resultados, estados de la cuestión, discusión y posibilidad de contacto con sus autores” (Sánchez 2018, 18). Debido a su importancia, este género discursivo ha sido estudiado en diferentes disciplinas, incluida la comunicación (Perdomo 2021).

Las reseñas de libros son frecuentes, aun cuando son en su mayoría por invitación y no arbitradas. De hecho, al parecer éstas han dado origen a nuevos géneros, pues se observa que se publican reseñas de películas, lo cual es bastante razonable considerando que, en el ámbito de las comunicaciones, este tipo de obras es de interés para la comunidad disciplinar.

Finalmente, las revisiones de literatura y los estudios teóricos se presentan como un género aparte, dejando el artículo original como formato para la presentación de revisiones sistemáticas. No obstante, por no ser materia directa de la investigación, no se profundizó en este hallazgo. Sería oportuno estudiar en detalle los géneros producidos en la comunidad de las comunicaciones y

sus tipos de investigación, ya que hay poca evidencia científica al respecto (Perdomo 2021).

En atención al idioma, los resultados parecen indicar, como se ha señalado en estudios previos, que se impulsa a los autores a orientar sus esfuerzos para publicar en inglés (Araiza, Ramírez y Díaz 2019) pues este idioma se observó en las PE de más de 90% de las revistas. Esta masificación de la publicación en inglés en las revistas de comunicación es consistente con los hallazgos de Silva y Blattmann (2018), quienes hallaron que 88% de las revistas brasileras de Economía publicaban en inglés, además de portugués y otros idiomas. También se alinean con un estudio previo en el que se encontró que en revistas de AA la mayoría publicaba artículos en inglés, incluso en revistas de países que no eran de habla inglesa (Fukuzawa 2017) y otro que reportó que las PE en materia de idioma en 239 revistas de Economía indizadas en la Web of Science (WoS) planteaban el inglés como lengua para la publicación (Henshall 2018).

Podría haber diferentes motivaciones para esta PE. En primer lugar, la visibilidad que ofrece al autor publicar en el idioma considerado la *lingua franca* de la ciencia para cubrir la necesidad de agilizar la difusión del conocimiento científico (Henshall 2018). Por otra parte, hay evidencia de que los artículos publicados en inglés presentan mayor número de citas (Fukuzawa 2017). Se sugieren estudios que verifiquen si, ante esta PE, los autores efectivamente están publicando más en inglés que en otros idiomas en las revistas que así lo permiten. De igual forma, sería oportuno conocer las PE en cuanto a la variante de inglés que los editores exigen en estas revistas, como lo hizo Henshall (2018) para otras disciplinas.

La mayoría de las revistas no limita la cantidad de autores ni el tiempo para volver a publicar. En consecuencia, es posible suponer que los editores reconocen la importancia para los autores del trabajo en equipo y del tiempo para publicar. Todas las revistas presentan políticas en cuanto a la declaración de autoría, las cuales se concretan en el envío o con formatos diseñados para tal fin. Se sugiere abordar este tema desde distintos enfoques, ya que en la literatura se observa preocupación por situaciones irregulares por conductas anti éticas y errores conceptuales entre la colaboración científica y la autoría propiamente dicha (Hilário, Cabriní y Chaves 2018). Además, se recomienda analizar políticas para disminuir prácticas que permitan cuestionar la autoría, como lo han expuesto Supak-Smolcic y Simundic (2015) y estudiar si existe un programa de cumplimiento que lo garantice.

La extensión de los artículos varía entre los diferentes géneros, siendo el artículo de investigación al que se le permite la mayor extensión. El hecho de que en algunas revistas ésta tiene un rango muy amplio (por ejemplo,

algunas que aceptan artículos de hasta 12 000 palabras) puede ser un indicador de que los editores reconocen la evidencia que señala que la extensión de los artículos afecta la citación de los mismos en las CCSS (Henshall 2018; Lyu y Wolfram 2018). Aun cuando en algunos casos esta preferencia es incluso inconsciente, como lo observaron Hasan y Breunig (2021); sería oportuno investigar al respecto.

Las normas de la APA son las más frecuentes para citas y referencias, aunque en las PE se observaron diferencias para citas dentro del texto y referencias. Algunas revistas plantean el uso de la misma norma para ambas. Por ejemplo, la revista *Comunicar*, que exige APA o la *Revista Historia y Comunicación Social* que se ciñe a las normas Harvard. Otras, por su parte, presentan normas particulares híbridas para las citas según algún sistema autor-fecha y una norma diferente para las referencias (por ejemplo, la revista *BiD*). Por otra parte, algunas revistas permiten libre elección del sistema de referencias siempre y cuando su uso sea consistente, para lo cual recomiendan gestores automáticos de referencias. Un ejemplo de esta política se aprecia en la revista *Publications*. Serían útiles nuevas investigaciones que profundicen en las razones por las que algunos editores mezclan normas y estilos para la presentación de citas y referencias, lo cual puede resultar incómodo para los autores. Estos hallazgos alertan a los investigadores en comunicación a familiarizarse con otros sistemas de referencia además de APA y trabajar con gestores como Mendeley, Zotero y similares.

Finalmente, en atención al APC, en las revistas sujetas totalmente a AA, la tendencia es a no exigirlo, ya que la mayoría de éstas cuentan con patrocinadores. Es importante que los autores interesados en publicar en AA y que no cuenten con apoyo financiero para cubrir el APC, seleccionen aquellas que son totalmente AA y no las híbridas o con APC opcional, las cuales en algunos casos tienen cuotas tan altas que superan el pago por suscripción (Araiza, Ramírez y Díaz 2019).

CONCLUSIONES Y CONSIDERACIONES FINALES

Este estudio ha sido una primera aproximación a la descripción de las políticas editoriales de revistas de AA indexadas en Scopus. En consecuencia, se requieren más investigaciones que den cuenta a los autores acerca de las opciones con que cuentan para publicar sus estudios en estas revistas. A partir del análisis realizado, se observó, en líneas generales, que las revistas de AA indexadas en Scopus en la sub área de comunicación tienden a ofrecer amplias posibilidades de visibilidad a los autores y a fortalecer la investigación

de forma colaborativa. De igual forma, se aprecia que la extensión se mantiene alineada con lo que otros autores han observado para las CCSS. Conocer esta información será de utilidad para los autores que desean publicar sus artículos en AA, ya que permite que se haga la elección en función de las necesidades y limitaciones de los autores (por ejemplo, idioma, urgencia de publicar y falta de presupuesto para APC). Los autores cuentan con un importante número de revistas de AA que se ajustan a diferentes necesidades, incluida la exoneración del pago por APC, lo cual es una excelente opción para aquellos investigadores que realizan sus estudios sin contar con el financiamiento de su institución de afiliación o entes externos a ésta.

Conflictos de interés: Los autores declaran que no existen conflictos de interés asociados a la presente investigación.

REFERENCIAS

- Alcalá-Albert, Gregorio Jesús, y María Elena Parra-González. 2021. "Bibliometric Analysis of Scientific Production on Nursing Research in the Web of Science." *Education Sciences* 11 (9).
<https://doi.org/10.3390/educsci11090455>
- Araiza, Verónica, María Esther Ramírez, y Alma Silvia Díaz. 2019. "El Open Access a debate: entre el pago por publicar y la apertura radical sostenible." *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 33 (80): 195.
<https://doi.org/10.22201/iibi.24488321xe.2019.80.58039>
- Arroyave-Cabrera, Jesús, y Rafael González-Pardo. 2022. "Investigación Bibliométrica de Comunicación en Revistas Científicas en América Latina (2009-2018)." *Comunicar* 30 (70): 81-91.
<https://orcid.org/0000-0002-5781-0456>
- Berkenkotter, Carol, y Thomas Huckin. 1995. *Genre Knowledge in Disciplinary Communication: Cognition/ Culture/ Power*. NJ: Lawrence Erlbaum Associates, Publishers.
- Bojo-Canales, C., y R. Melero. 2021. "Open Access Editorial Policies of SciELO Health Sciences Journals." *Journal of Information Science*, 1-14.
<https://doi.org/10.1177/01655515211015135>
- Cruz, Edison. 2008. "Continuous Publication." *Revista Estomatología* 336 (7659): 1450.
<https://doi.org/10.1136/bmj.a491>
- De-Filippo, Daniela. 2013. "La producción científica española en comunicación en WOS." *Comunicar* 21 (41): 25-34.
<http://dx.doi.org/10.3916/C41-2013-02>
- Estrada-Cuzcano, Alonso, y Joel Alhuay-Quispe. 2020. "A Bibliometric Approach to Revista de Comunicación (Perú) since 2002 to 2019." *Revista de Comunicación* 19 (2): 111-24.
<https://doi.org/10.26441/RC19.2-2020-A6>

- Felipe, Jimena Beltrão, y Taíse Cruz Silva. 2020. “Análise de Políticas Editoriais de Periódicos Científicos Nacionais: Contribuições Para o Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas.” *Perspectivas Em Ciência Da Informação* 25 (3): 1-28.
<https://doi.org/10.1590/1981-5344/3301>
- Fukuzawa, Naomi. 2017. “Characteristics of Papers Published in Journals: An Analysis of Open Access Journals, Country of Publication, and Languages Used.” *Scientometrics* 112 (2): 1007–23.
<https://doi.org/10.1007/s11192-017-2414-y>
- González, Raidel, Juan Cardentey, y Yuleivis Izquierdo. 2017. “Estudio Bibliométrico de la Revista de Ciencias Médicas de Pinar del Río. 2013-2015.” *Revista de Ciencias Médicas de Pinar Del Río* 21 (1): 100-109.
<http://scielo.sld.cu/pdf/rpr/v21n1/rpr131>
- Hasan, Syed, y Robert Breunig. 2021. “Article Length and Citation Outcomes.” *Scientometrics* 126 (9): 7583–7608.
<https://doi.org/10.1007/s11192-021-04083-x>
- Henshall, Ann C. 2018. “English Language Policies in Scientific Journals: Signs of Change in the Field of Economics.” *Journal of English for Academic Purposes* 36: 26-36.
<https://doi.org/10.1016/j.jeap.2018.08.001>
- Hilário, Carla Mara, Maria Cabrini, y José Augusto Chaves. 2018. “Aspectos Éticos da Coautoria em Publicações Científicas.” *Em Questão* 24 (2): 12.
<https://doi.org/10.19132/1808-5245242.12-36>
- Knorr-Cetina, Karin. 1981. *The Manufacture of Knowledge*. Oxford: Pergamon.
- Lopera-Pérez, Marisol, Alexander Maz-Machado, María José Madrid, y Astrid Cuida. 2021. “Bibliometric Analysis of the International Scientific Production on Environmental Education.” *Journal of Baltic Science Education* 20 (3): 428-42.
<https://doi.org/10.33225/jbse/21.20.428>
- Lyu, Peng-hui, y Dietmar Wolfram. 2018. “Do Longer Articles Gather More Citations? Article Length and Scholarly Impact among Top Biomedical Journals.” *Proceedings of the Association for Information Science and Technology* 55 (1): 319-326.
<https://doi.org/10.1002/pra2.2018.14505501035>
- Madrid, María José, Noelia Jiménez-Fanjul, Carmen León-Mantero, y Alexander Maz-Machado. 2017. “Revistas Brasileñas de Educación en SCOPUS: Un Análisis Bibliométrico.” *Biblios* 67: 30–41.
<https://doi.org/10.5195/biblios.2017.344>
- Meza, Paulina. 2018. “Estrategias de Posicionamiento del Autor en Artículos de Investigación de Ciencias Sociales, Humanidades e Ingeniería: Novatos versus Expertos.” *Información Tecnológica* 29 (2): 3–18.
<https://doi.org/10.4067/s0718-07642018000200003>
- Miguel, Sandra, Ely Tannuri, y Maria Cabrini. 2016. “Scientific Production on Open Access: A Worldwide Bibliometric Analysis in the Academic and Scientific Context.” *Publications* 4 (1).
<https://doi.org/10.3390/publications4010001>
- Perdomo, B. 2021. “Análisis Retórico-Discursivo del Género Artículo de Investigación en Español en el Ámbito de la Comunicación.” *Nueva Revista del Pacífico* 74 (e-0719-5176): 170-195.
<http://dx.doi.org/10.4067/S0719-5176202100>

- Pérez, M., F. Herrera, y J. Ramírez. 2020. "El Artículo de Investigación Científica en la Era del Acceso Abierto: Consideraciones e Implicaciones." *Publicacions Semestral Herreriana* 2 (1): 30-36.
<https://repository.uaeh.edu.mx/revistas/ind>
- Sánchez, Alexander Arbey. 2018. "Consideraciones Sobre el Artículo Científico (Ac): Una Aproximación Desde el Análisis de Género y el Posicionamiento." *Lingüística y Literatura* 39 (73): 17-36.
<https://doi.org/10.17533/udea.lyl.n73a01>
- Silva, Angelita, y Ursula Blattmann. 2018. "Revistas de Arquivologia: Estudo Bibliométrico." *ÁGORA: Revista Do Curso de Arquivologia Da UFSC* 28 (56): 146-63.
<https://agora.emnuvens.com.br/ra/article/view/721>
- Supak-Smolcic, Vesna, y Ana-Maria Simundic. 2015. "Biochemia Medica's Editorial Policy on Authorship." *Biochemia Medica* 25 (3): 320-323.
<https://doi.org/10.11613/BM.2015.032>
- Swales, John. 1990. *Genre Analysis: English in Academic Research Settings*. Glasgow: Cambridge University Press.
- Típu, S., y J. Ryan. 2022. "Are Business and Management Journals Anti-Replication? An Analysis of Editorial Policies." *Management Research Review* 45 (1): 101-117.
<https://doi.org/10.1108/MRR-01-2021-0050>
- Zhang, Tenghao. 2021. "Will the Increase in Publication Volumes 'Dilute' Prestigious Journals' Impact Factors? A Trend Analysis of the FT50 Journals." *Scientometrics* 126 (1): 863-69.
<https://doi.org/10.1007/s11192-020-03736-7>

Para citar este texto:

- Perdomo, Bexi, y Oscar Alberto Morales. 2022. "Políticas editoriales de revistas de Comunicación en acceso abierto indexadas en Scopus". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 135-151.
<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58650>

Producción de libros académicos en Colombia entre 2013 y 2019: un avance para estudios comparados

Esteban Giraldo-González*

Elea Giménez-Toledo**

Juan Felipe Córdoba-Restrepo***

Artículo recibido:

7 de abril de 2022

Artículo aceptado:

24 de agosto de 2022

Artículo de investigación

RESUMEN

El artículo tiene como objetivo identificar y analizar la producción de libros académicos en Colombia entre 2013 y 2019. Para hacerlo, toma como fuente la base de datos de ISBN gestionada por la Cámara Colombiana del Libro y selecciona dentro del universo editorial general, a través de materias, la producción propiamente académica. Así, con la información exigida por la organización gremial del libro en el país y suministrada por los mismos editores, se presentan criterios que permiten caracterizar la producción de libros académicos en Colombia dentro del periodo.

* Dpto. de Biblioteconomía y Documentación, Universidad Carlos III de Madrid, España
esteban.giraldo@alumnos.uc3m.es

** Centro de Ciencias Humanas y Sociales, Consejo Superior de Investigaciones Científicas (CSIC) de España
elea.gimenez@cchs.csic.es

*** Editorial, Universidad del Rosario, Bogotá, Colombia
juan.cordoba@urosario.edu.co

Es la primera caracterización de este tipo que se hace en el país. Resulta llamativo que, según la metodología utilizada, el grupo de libros académicos constituya el 67% de la producción total de libros registrada. Las materias en las que más se publican libros académicos son Sociología con el 15%, Economía con el 12% y Derecho con el 10%. Con respecto al origen de las publicaciones, se debe reconocer el papel fundamental que juegan las editoriales universitarias. El grupo de libros académicos es además analizado por si se trata o no de traducciones. Este aporte es el piloto para consolidar los criterios para hacer posibles estudios comparados de la producción de libros académicos en países de habla española y portuguesa.

Palabras clave: Libros académicos; Editoriales académicas; Comunicación pública de la ciencia

Production of academic books in Colombia between 2013 and 2019: a step forward in comparative studies
Esteban Giraldo-González, Elea Giménez-Toledo and Juan Felipe Córdoba-Restrepo

ABSTRACT

This paper aims to identify and analyze academic books production in Colombia between 2013 and 2019. For doing this takes as resource the ISBN database managed by the Colombian Book Chamber and selects the academic production properly within the general publishing universe and through the subjects. Thus, based on the information required by the book trade organization in Colombia and supplied by the publishers themselves, the proposed criteria will allow characterizing the production of books in the country. It is the first characterization of this type done in Colombia. It is remarkable that according to the used methodology, the group of academic books makes up 67% of the total recorded book production. The subjects in which more academic books are published are Sociology (15%), Economics (12%), and Law (10%). Regarding the origin of the publications, it should be recognized that university publishers play a fundamental role. The group of academic books is also analyzed to determine whether these are translations or not. This contribution is the pilot study to consolidate

criteria for carrying out comparative studies about the production of academic books in Spanish and Portuguese-speaking countries.

Keywords: Academic books; Academic publishers; Disseminating science

INTRODUCCIÓN

Identificar la producción de libros académicos en Colombia entre 2013 y 2019 a partir de la información contenida en las bases de datos de ISBN gestionada por la Cámara Colombiana del Libro constituye un insumo para entender lo que el país publica bajo esta amplia categoría. Es uno de los primeros pasos para adquirir un conocimiento detallado del panorama editorial académico en el país, condición imprescindible para justipreciar el valor de la edición académica en el universo general de la edición y construir análisis que permitan una interlocución informada y argumentada sobre las políticas de medición y evaluación que, dentro del sistema de ciencia y tecnología nacional, pesan sobre los libros. Este esfuerzo, además, es un de piloto para, con la información del ISBN de otros países, construir modelos de análisis y presentación de la información que hagan posibles análisis comparados entre países, regiones y temáticas. En ese sentido, hace parte la consolidación de una cartografía de la edición iberoamericana, proyecto que está en curso (García-Valencia 2021; Giraldo-González et al. 2021) y del cual este artículo es uno de sus primeros productos.

Para lograr la identificación fue necesario implementar una metodología de análisis de datos, a partir de las materias académicas con las que los agentes editoriales registran los libros al momento de solicitar el ISBN. Se entiende como agente editorial la persona natural o jurídica que ocasional o periódicamente edita uno o más títulos (Cerlalc 2008). Concretamente, se utilizan los metadatos de materia, seleccionando aquellas materias que se inscriben dentro de lo académico/científico y excluyendo el resto.

Para hacerlo establece la producción editorial general del país, año a año, desde 2013 hasta 2019, según agentes editoriales. Después, compara la producción de libros propiamente académicos con la producción total, discriminándolas por año y por tipo de agente editor. Pasa a presentar los resultados analizándolos desde el punto de vista de las materias en las que se publicaron libros académicos en el país en el periodo elegido. Posteriormente, indica el número de libros que son traducciones dentro de la producción académica.

Estos resultados son un desarrollo de los estudios adelantados por el Grupo de investigación sobre Libro Académico (ILIA) del Consejo Superior de Investigaciones Científicas de España, referente a la valoración de la producción académica, esta vez en el contexto colombiano. Tiene, pues, antecedentes metodológicos en trabajos como “La edición académica española a partir de sus metadatos” (Giménez-Toledo y Mañana-Rodríguez 2016) y el informe *La edición universitaria española. Análisis de la producción editorial de libros* (Giménez Toledo 2019), entre otros. Además se suma, desde una perspectiva más general, a los estudios particulares sobre el prestigio de las editoriales universitarias, particularmente “Calidad y prestigio de la edición académica. El caso colombiano” (Giménez Toledo et al. 2021) y el estudio análogo para Brasil. Atendiendo al estado de la cuestión de los estudios editoriales en Colombia —un campo que si bien es incipiente se encuentra en desarrollo—, entra en franca conversación, concentrándose en el campo de la producción de libros académicos, con los trabajos “El libro en Colombia: Entre la sostenida concentración y la lenta marcha hacia la independencia (2000-2019)” (Marín-Colorado 2020), “Edición en Colombia (1970-1990). Del boom de la industria gráfica a la diversificación de la industria editorial” (Marín-Colorado 2018), y los estudios seminales *El libro en Colombia. Situación y perspectivas* (Arango 1991) y *Caracterización del sector editorial en Colombia. Una aproximación al mundo editorial en el país 2013* (Lado B 2016).

El artículo es un análisis descriptivo que no se contenta con presentar escuetamente los resultados, sino que propone algunas hipótesis en torno a los motivos que explican el panorama que se presenta y arriesga posibles líneas de investigación que surgen de los resultados obtenidos. Pretende ser, pues, el punto de partida para futuras investigaciones, tanto a nivel colombiano como internacional.

METODOLOGÍA

Se tuvo como fuente la base de datos de registros ISBN que administra la Cámara Colombiana del Libro. La elección de acudir al ISBN obedece a que es la base de datos más amplia y universalmente reconocida por los editores para registrar su producción. Se debe aclarar, no obstante, que un registro no corresponde necesariamente a un título en la medida que se solicitan y se expiden ISBN dependiendo del formato del libro: físico, digital o de impresión bajo demanda. Esto significa que un solo título puede tener hasta tres ISBN, uno por cada uno de sus formatos. El tratamiento de los datos fue llevado a cabo en el Cerlalc, que es una de las instituciones vinculada a la investigación de la que surge este trabajo.

Dentro del universo general de ISBN, para consolidar el grupo de libros académicos se llevó a cabo el siguiente proceso: se excluyeron los libros registrados por Autores/Editores, ya que no son editores institucionalmente respaldados, por lo tanto, son editores accidentales y se trata de autopublicación. Una vez realizada esa selección, se excluyeron, a partir de la clasificación temática Dewey —que es la que sistemática e históricamente ha utilizado la Cámara Colombiana del Libro y, por lo tanto, los editores en el país—, las siguientes materias consideradas como no académicas: 010–Bibliografías, 030–Enciclopedias y libros de hechos, 050–Revistas, diarios y publicaciones seriadas, 060–Asociaciones, organizaciones y museos, 070–Medios de comunicación, periodismo y publicación, 080–Colecciones generales, 090–Manuscritos y libros raros, 130–Fenómenos paranormales, 220–La Biblia, 240–Moral cristiana y teología piadosa, 250–Órdenes cristianas e Iglesia local, 260–Organización cristiana, trabajo social y adoración, 290–Otras religiones, 700–Las artes, bellas artes y artes decorativas, 730–Artes plásticas, Escultura, 740–Dibujo y artes decorativas, 750–Pintura y pinturas, 760–Artes gráficas, arte de grabar y grabados, 770–Fotografía y fotografías, 790–Artes recreativas y de la actuación, 800–Literatura y retórica, 810–Literatura americana en inglés, 820–Literatura inglesa e inglesa antigua, 830–Literatura de las lenguas germánicas, 840–Literatura de las lenguas romances, 850–Literaturas italiana, rumana y retorromana, 860–Literaturas española y portuguesa, 870–Literaturas latinas e italianas, 880–Literaturas griegas clásicas y modernas, 890–Literaturas de otras lenguas, 910–Geografía y viajes, 920–Biografía.

Las materias sí incluidas están consignadas en el Anexo 1. Allí, además, para poder tener un grupo de materias operativo para la presentación de resultados, una vez que se llevaron a cabo estas exclusiones, se indica cómo los registros obtenidos fueron reclasificados temáticamente de acuerdo con las materias utilizadas por el estudio de ILIA para el sector académico español (Giménez Toledo 2019, 12), esto para satisfacer la necesidad de una agrupación temática sintética que permita, en el futuro, comparaciones entre países (ver Anexo 1).

Una vez extraídos los metadatos de materia incorporados por los editores para cada libro publicado, siguiendo la clasificación Dewey, se procedió a agruparlos por grandes materias, atendiendo al cuadro de clasificación mostrado. Así agrupados, se procedió a describir los resultados obtenidos a partir de la información que requiere el ISBN.

RESULTADOS

Producción editorial general

Un primer hallazgo que llama la atención es la tendencia positiva del crecimiento de las solicitudes de ISBN para libros académicos en el país en todo el periodo. Sin embargo, es necesario señalar que éste no es constante año a año. En los siete años del estudio hubo un crecimiento de más del 24%. Si se compara el año de mayor producción, 2018, con el de menor producción, 2013, el crecimiento alcanza el 30,4%. Estos resultados se presentan de manera discriminada en la *Tabla 1*.

También es destacable el crecimiento interanual entre 2017 y 2018, que sólo en ese año fue del 13%. Este incremento y el posterior descenso del año siguiente obedecen a un comportamiento atípico de la producción de entidades privadas no editoriales ese año, concretamente, de la Caja de Compensación Familiar de Caldas, a la que le fueron expedidos 1 000 registros. Esta particularidad, en todo caso, no modifica el hecho relevante: a lo largo del tiempo la producción editorial en Colombia mantiene un crecimiento sostenido.

	2013	2014	2015	2016	2017	2018	2019	Total
Editorial comercial	7.469	6.646	7.103	7.407	7.014	6.647	7.135	49.421
Editorial universitaria	2.795	2.973	3.628	3.535	4.691	5.805	4.663	28.090
Entidad privada no editorial	2.561	2.817	2.852	2.740	2.763	4.046	3.535	21.314
Entidad pública	943	1.313	1.576	1.639	1.379	1.410	1.540	9.800
Autores/editores	1.966	1.865	2.174	2.079	2.411	2.670	2.766	15.931
Otros	312	286	288	436	227	289	283	2.121
Total	16.046	15.900	17.621	17.836	18.485	20.867	19.922	126.677

Tabla 1. ISBN expedidos por tipo de agente editorial

Fuente: Elaboración propia

La distribución por tipo de agente editorial, según los criterios utilizados por el Cerlalc, también muestra constantes significativas.

El Cerlalc clasifica a los agentes que solicitan ISBN en las siguientes categorías, las cuales se establecen a partir de criterios como su naturaleza jurídica y su actividad principal:

Editoriales comerciales: empresas cuya actividad principal es la edición de libros en cualquier formato y sustrato disponible.

Editoriales universitarias: instituciones de educación superior que editan libros. Dicha actividad puede ser llevada a cabo por una facultad, departamento o dependencia especializada.

Entidad privada no editorial: cualquier tipo de empresa o entidad de carácter privado que no tiene como actividad principal la edición de libros, pero que ocasionalmente lo hace.

Entidades públicas: cualquier institución pública que no tiene como actividad principal la edición de libros, pero que ocasionalmente lo hacen.

Autores-editores: autores que publican sus propias obras.

Otros agentes editores: impresores, distribuidores, editoriales de publicaciones periódicas y demás agentes no identificados (Cerlalc 2020, 19).

Así, en Colombia, para el periodo estudiado, son las editoriales comerciales las que de manera consistente lideran la expedición de ISBN. En promedio, en el periodo observado, constituyen el 39% de la producción total de libros en el país. Están seguidas por el grupo de editoriales universitarias con un 22% de la cantidad total y por las entidades privadas no editoriales, con un 17%. Si se suman ambos grupos, resulta que tienen un porcentaje igual al de las editoriales comerciales, lo que da a entender que la labor de publicación en el país no está mayoritariamente orientada por el afán de lucro. También llama la atención que el grupo de autores/editores, esto es, editores accidentales, tengan una participación promedio del 13%, con un crecimiento sostenido, lo que indica que la práctica de la autoedición es cada vez más común. Las entidades públicas y otro tipo de agentes editores completan el panorama con un 8% y 2% en promedio respectivamente. La *Figura 1* expresa discriminadamente estos hallazgos.

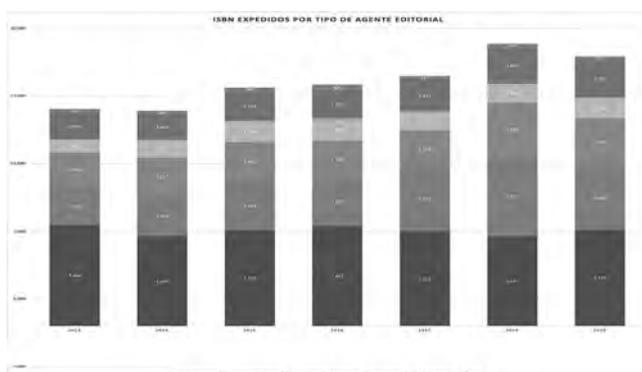


Figura 1. ISBN expedidos por tipo de agente editorial

Fuente: Elaboración propia

Como se indicó en la metodología, si se sustraen los autores/editores, en tanto que por definición no son empresas o instituciones dedicadas constantemente a la edición y publicación de libros, la tendencia al crecimiento en las solicitudes de ISBN a lo largo de los siete años del estudio se mantiene. Y resulta que, ya atendiendo al mundo de las editoriales formales, como queda detalladamente descrito en la *Tabla 2*, la proporción promedio en todo el periodo de la participación de las editoriales comerciales sube a un 45% frente a un 25% de las editoriales universitarias y un 19% de las entidades privadas no editoriales, mientras que las entidades públicas y otro tipo de editores componen el 11% restante.

	2013	2014	2015	2016	2017	2018	2019	Total
Editorial comercial	7.469	6.646	7.103	7.407	7.014	6.647	7.135	49.421
Editorial universitaria	2.795	2.973	3.628	3.535	4.691	5.805	4.663	28.090
Entidad privada no editorial	2.561	2.817	2.852	2.740	2.763	4.046	3.535	21.314
Entidad pública	943	1.313	1.576	1.639	1.379	1.410	1.540	9.800
Otros	312	286	288	436	227	289	283	2.121
Total	14.080	14.035	15.447	15.757	16.074	18.197	17.156	110.746

Tabla 2. ISBN expedidos a solicitud de editoriales por tipo de agente editorial excluyendo autores/editores
Fuente: Elaboración propia

En este punto cabe destacar que si bien las editoriales comerciales siguen siendo el grupo que más ISBN solicita, a lo largo del tiempo han perdido su participación en detrimento de las editoriales universitarias, las entidades privadas no editoriales y las entidades públicas. De representar el 53% de toda la producción formal en el país en 2013, pasaron a representar el 42% en 2019. En consecuencia, el grupo de las editoriales universitarias pasó del 20 al 27% de participación entre 2013 y 2019, el de las entidades privadas no editoriales del 18 al 21%, y las entidades públicas del 7 al 9%. En este punto quizás valga la pena señalar que éste es un rasgo característico de Colombia en comparación con el conjunto de América Latina y con varios de los países del subcontinente considerados individualmente. Si se toma a América Latina en conjunto, excluyendo a los autores-editores, la participación de las solicitudes de ISBN de las editoriales comerciales en el periodo ronda el 60% año a año, y en Argentina, Bolivia, Brasil, Chile, Honduras, México, Panamá, Paraguay, Perú y Uruguay, éste es igual, cuando no superior, a dicho 60% (Liber 2021).

Esto muestra un debilitamiento progresivo de las editoriales comerciales en Colombia en lo que respecta a su participación en el conjunto de los ISBN expedidos para libros relacionados con materias académicas. El incremento sostenido de la producción de las editoriales universitarias y, por lo tanto, su ganancia en la proporción dentro del universo de los ISBN relacionados con materias académicas, coincide con la presión para publicar que el sistema de ciencia y tecnología nacional les impone a los académicos de las universidades a través de las convocatorias para la clasificación de los grupos de investigación (Colciencias 2013; 2018).

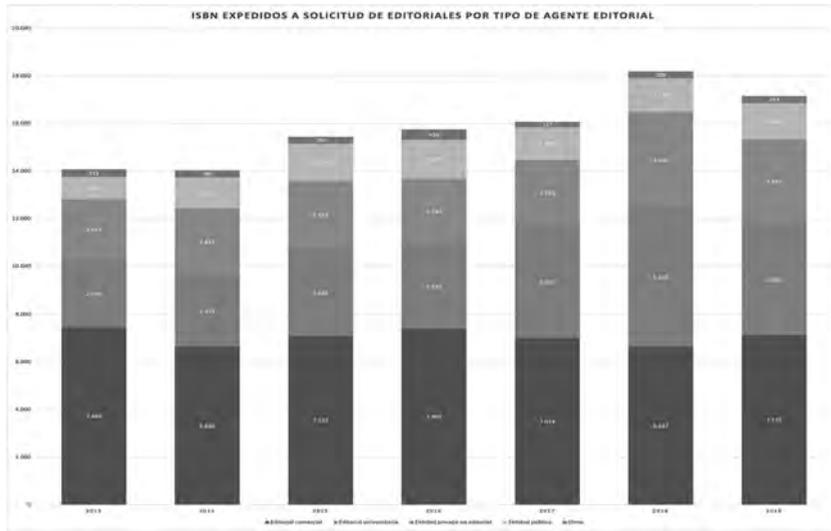


Figura 2. ISBN expedidos a solicitud de editoriales por tipo de agente editorial
Fuente: Elaboración propia

Un aspecto que es oportuno resaltar es el incremento en los grupos distintos al de la edición comercial. Por ejemplo, y de acuerdo con los datos expresados en la *Figura 2*, si se compara el año de menor producción, 2013, con el de mayor producción, 2018, dentro de los mismos grupos, resulta significativo que las editoriales universitarias tuvieron un incremento de 108%, el de las entidades privadas no editoriales uno de 58%, y el de las entidades públicas uno de 50%. Esto muestra el creciente dinamismo de los agentes editoriales cuyo motor no es el afán de lucro y, en la misma dirección, que la oferta de libros académicos no pasa, necesariamente, por las exigencias del mercado.

Así, desde la perspectiva de la cantidad de ISBN solicitados, es válido preguntarse por una posible y próxima recomposición del peso de los agentes editoriales en el universo de lo que se publica en el país, y en la que los editores

comerciales cederán su primacía en el terreno propiamente académico. Con las bases que funda esta metodología, en la agregación de datos por años, podría corroborarse esta tendencia y, a partir de su análisis desprender reflexiones que aborden políticas de preservación de la producción académica que en principio no se sustentan en el afán de lucro y, por lo tanto, podrían requerir de apoyos institucionales y estatales, o maneras no comerciales de financiación.

Producción de libros académicos

Con el objetivo de realizar un análisis más detallado del sector de la edición académica, se tomó el conjunto de registros clasificados en materias académicas, conforme al cuadro ya presentado. Al aplicar la metodología, esto es: al excluir dentro del universo total de los ISBN solicitados en el periodo analizado la producción de autores/editores y a partir de ahí determinar los ISBN asociados a materias académicas, se encuentran los siguientes resultados.

ISBN relacionados con materias académicas dentro de la producción de editoriales

La tendencia al crecimiento de la producción se mantiene a lo largo del tiempo, aun cuando entre los años 2018 y 2019 haya un decrecimiento cercano al 5%. Entre 2013 y 2019 se pasa de 9.174 a 11.512 números ISBN expedidos para libros de materias académicas, lo que significa un crecimiento en el periodo del 25%. En todo el periodo, fueron expedidos 73.772 ISBN de libros académicos. Los registros expresados año a año según el tipo de editorial se presentan en la *Tabla 3*.

	2013	2014	2015	2016	2017	2018	2019	Total
Editorial comercial	3.947	3.007	3.429	3.809	3.660	3.212	3.279	24.343
Editorial universitaria	2.475	2.636	3.192	3.112	4.152	5.086	4.077	24.730
Entidad privada no editorial	1.827	2.117	2.163	2.143	2.067	2.503	2.797	15.617
Entidad pública	748	988	1.174	1.325	1.115	1.108	1.219	7.677
Otros	177	215	217	333	140	183	140	1.405
Total	9.174	8.963	10.175	10.722	11.134	12.092	11.512	73.772

Tabla 3. ISBN relacionados con materias académicas dentro de la producción de editoriales

Fuente: Elaboración propia

Se marca, además, un cambio evidente entre el universo general y el universo constituido propiamente por ISBN relacionados con materias académicas. En los últimos tres años del periodo analizado, las editoriales universitarias son los agentes editores que más obtienen ISBN relacionados con materias académicas, en detrimento de las editoriales comerciales. Así, para el total del periodo, son las universitarias el grupo que más libros académicos produjo con el 33% total de los ISBN. El sector comercial retrocedió del 43% en 2013 al 28% en 2019. Llama también la atención el incremento sostenido de las entidades privadas no editoriales y las entidades públicas, con 21 y 10% de la producción total, respectivamente, datos expresados en la *Figura 3*, y que muestran el peso de la edición institucional en lo académico, con todos los significados que esto arrastra, concretamente el que por su origen no es necesariamente una producción académicamente validada —en términos de evaluación de pares, por ejemplo—.

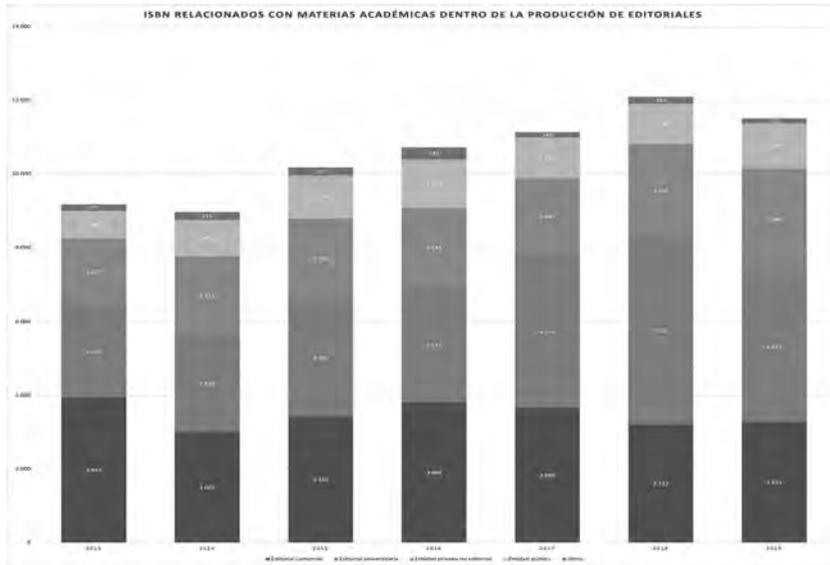


Figura 3. ISBN relacionados con materias académicas dentro de la producción de editoriales
Fuente: Elaboración propia

Las razones de este cambio pueden estar relacionadas con el proceso de profesionalización y consolidación de las editoriales universitarias en los últimos años, aupadas —como ya se ha señalado— por los requisitos del sistema de medición de ciencia y tecnología del país, que exigen a los investigadores validar su producción, entre otras maneras, a través de la producción de libros. Este hecho, positivo, debe suscitar, no obstante, una reflexión de fondo: se

debe estudiar si ese crecimiento de la producción universitaria ha sido natural, orgánico, muestra de un verdadero proceso de maduración de las editoriales y los lectores académicos o si ha respondido, más bien, a una dinámica artificial en la que se han creado infraestructuras editoriales y publicado libros en cantidades cada vez mayores para cumplir con los requisitos de las autoridades que evalúan el desempeño de las universidades y los investigadores a ellas adscritos. Esto es: habría que preguntarse si la producción de libros académicos ha crecido de manera significativa a espaldas de la calidad de lo que se publica y la cantidad de lectores de los textos en sí mismos, con el objetivo no de divulgar la ciencia sino para que, fundamentalmente, las universidades puedan clasificar sus grupos de investigación y, a través de ellos, alcanzar los registros de sus programas, certificaciones de calidad, acceso a los escasos recursos de investigación y mejoría en los ránquines que las clasifican. Este estudio puede convertirse, en el futuro, en un insumo para responder estas preguntas.

Esto refuerza, además, la intuición según la cual la producción de libros académicos en Colombia no está motivada por el afán de lucro, por las necesidades manifiestas de un mercado en crecimiento de lectores de contenidos académicos. Más allá de la constatación objetiva de este hecho, vale la pena que otras investigaciones se pregunten las razones que explican semejante condición y las implicaciones institucionales, intelectuales, políticas y económicas de ese panorama. Nuevamente, este estudio y su sostenimiento en el tiempo puede ser un punto de partida para este tipo de análisis.

Proporción de los ISBN relacionados con materias académicas dentro de la producción de editoriales

Como señala la *Figura 4*, surge en este punto una característica permanente: en el universo general de producción de libros en Colombia, la mayoría son libros a los que los agentes editores relacionan con materias académicas. El año que menos representación tuvo este grupo fue 2014, con el 64% de la producción. El que más fue 2017 con el 69%. En todo el periodo se consolidó un promedio de 67%.

Si bien las cifras resultan consistentes en términos objetivos, los resultados deben matizarse a la luz de otra constatación igual de evidente, producto del análisis que se presenta en este estudio: para avanzar con un estudio que permita el análisis y la discusión, sería necesario revisar con más detalle la clasificación solicitada en los ISBN, ya que todo parece indicar que no siempre la calidad de los datos aportados es la adecuada. Es usual, por ejemplo, que los libros didácticos y los que tienen propósitos educativos, de cualquier

materia, terminen siendo clasificados en la materia Educación, por su intención de enseñar tal o cual tema, más no por su contenido, que puede ser de cualquier otra materia, o que, por ejemplo, libros de autoayuda y guías prácticas de bienestar terminen clasificadas en las materias de Filosofía o Psicología. Incorrecciones de este tipo indican la necesidad de formar a los editores e insistir en los procesos de calidad de la información que se da al solicitar el registro del libro.

Sin embargo, es un hecho incontrovertible que la mayoría de los ISBN expedidos en el periodo de estudio están relacionados con materias de disciplinas académicas.

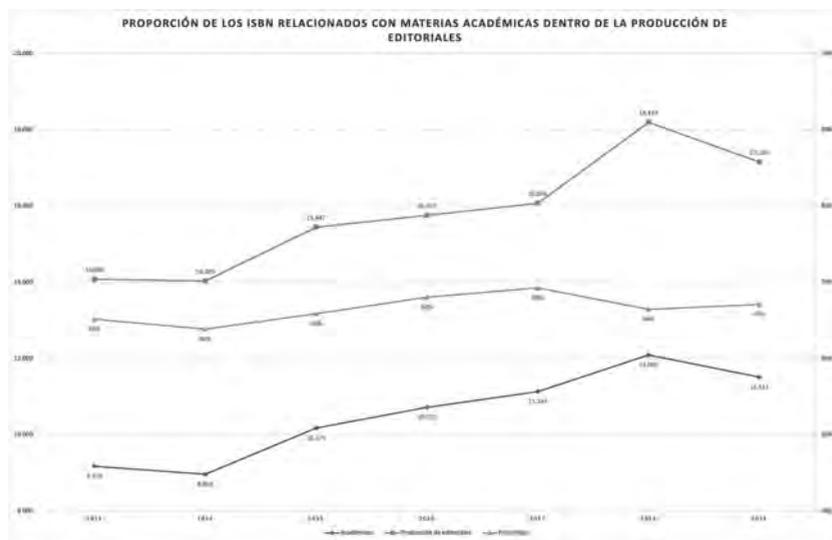


Figura 4. Proporción de los ISBN relacionados con materias académicas dentro de la producción total
Fuente: Elaboración propia

Agentes editoriales que solicitan ISBN relacionados con materias académicas

En este apartado se expresa la concentración de la producción académica en los agentes editores que se identifican en el registro ISBN. Es un punto muy relevante, en la medida que caracteriza la cantidad de editoriales que soportan la producción académica en el país. Nuevamente es el grupo de editoriales universitarias el que tiene mayor importancia relativa dentro de la producción académica.

Como se detalla en la *Tabla 4*, en un universo en el que 1.148 agentes editores en promedio al año solicitan ISBN relacionados con materias académicas, 154 editoriales, el grupo de las universitarias, obtuvo 24.730 de los

73.772 ISBN relacionados con materias académicas expedidos en el periodo 2013-2019. Esto quiere decir que sólo 13,4% de los agentes editores, las universitarias, publican el 33,5% de toda la producción académica.

En contraste, las entidades privadas no editoriales, que en promedio al año son 609 agentes editores, el 53% de todos los agentes editores, obtuvieron 15.617 ISBN relacionados con materias académicas, el 21,2% en el periodo 2013-2019.

En el caso de las editoriales comerciales, en el periodo del estudio, en promedio 226 editoriales, el 19,7% entre los agentes editoriales, obtuvieron 24.343 ISBN relacionados con materias académicas, el 33% del total.

	Editorial comercial	Editorial universitaria	Entidad privada no editorial	Entidad pública	Otros
2013	217	155	567	122	13
2014	207	148	592	136	15
2015	227	147	571	159	22
2016	241	145	589	132	35
2017	242	162	640	116	23
2018	212	163	620	128	25
2019	234	159	683	164	26

Tabla 4. Agentes editoriales que solicitan ISBN relacionados con materias académicas

Fuente: Elaboración propia

Visto de otra manera, en Colombia una editorial universitaria, entre 2013 y 2019, produjo en promedio 22,9 libros académicos, es decir: 3,27 libros académicos al año. Cada editorial comercial produjo en todo el periodo 15,41 títulos, lo que da 2,2 libros académicos por año. Cada entidad privada que publica libros académicos 3,66, lo que representa 0,52 libros académicos al año. Y cada entidad pública que edita libros académicos 8,02, lo que significa 1,14 libros académicos anualmente. El panorama, año a año, se presenta en la *Figura 5*.

Estos resultados muestran, en cualquier caso, una dispersión significativa en la producción de libros académicos. Es decir: que son muchos los agentes que publican, pero, en promedio, cada uno de ellos no publica muchos libros al año. Si se toma la totalidad de agentes, 8.037 registrados en los 7 años que se analizan, y se dividen entre los 73.772 ISBN relacionados con materias académicas expedidos por la Cámara Colombiana del Libro en el periodo, resulta que cada agente sólo publicó 9,18 libros en 7 años, esto es: 1,31 libros académicos al año.

Esta dispersión quizás soporte la hipótesis de una fragilidad de los agentes editoriales en tanto que empresas editoriales, sobre todo si se tiene en cuenta que estadísticamente los agentes editoriales publican menos de dos libros al año. Puede afirmarse que semejante dispersión de agentes editoriales habla de un alto grado de bibliodiversidad, al menos en lo referente a la cantidad de editores. Sin embargo, más allá de esa dispersión, un paso posterior de esta investigación permitirá identificar las infraestructuras editoriales sólidas y responsables de publicación de libros académicos en el país.

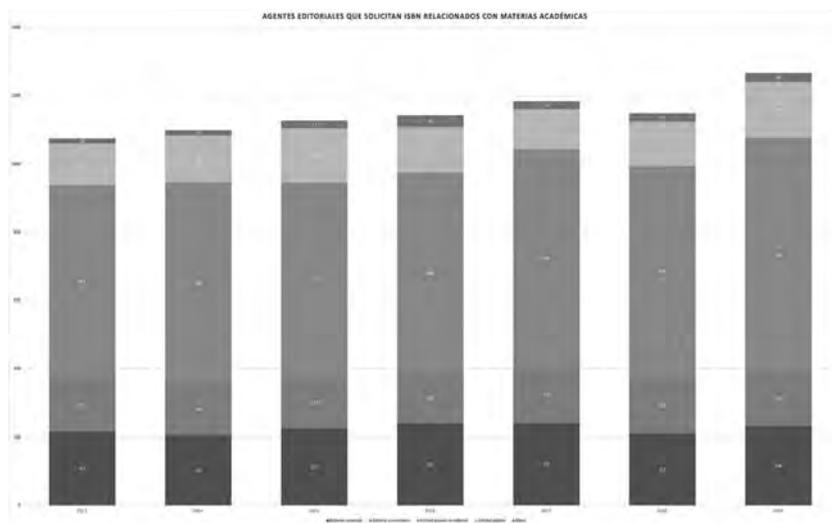


Figura 5. Agentes editoriales que solicitan ISBN relacionados con materias académicas

Fuente: Elaboración propia

ISBN relacionados con materias académicas por materias

La materia en la que más ISBN fueron expedidos fue Sociología con el 15% del total de los registros asociados a materias académicas en los 7 años analizados. En su orden le siguen Economía con el 12%, Derecho con el 10%, Educación con el 9%, Biblioteconomía y documentación con el 7%, Ciencias con el 7%, Ciencias políticas con el 6%, Filosofía con el 6%, Medicina con el 5%, Bellas artes con el 5%, Lingüística, literatura y filología con el 4%, Ingenierías con el 3%, Psicología con el 3%, y Agricultura, Arqueología e Informática con el 1% cada una. La *Tabla 5* expresa el panorama de manera ordenada.

Materia	Total entre 2013-2019	Porcentaje
Sociología	11.415	15%
Economía	8.750	12%
Derecho	7.423	10%
Educación	6.947	9%
Biblioteconomía y documentación	5.399	7%
Ciencias	5.344	7%
Ciencias políticas	4.597	6%
Filosofía	4.258	6%
Medicina	3.726	5%
Bellas artes	3.487	5%
Lingüística, literatura y filología	3.151	4%
Ingenierías	2.118	3%
Psicología	1.853	3%
Tecnología	1.316	2%
Historia	1.085	1%
Agricultura	1.073	1%
Arqueología	1.026	1%
Informática	804	1%
Total	73.772	100%

Tabla 5. ISBN relacionados con materias académicas por materias

Fuente: Elaboración propia

Año a año, materia a materia, como se presenta en la *Figura 6*, los resultados son los siguientes:

Más allá de los análisis sobre la fortaleza nacional en torno a tal o cual disciplina, midase como se mida, cabe resaltar la clara predominancia de materias de las ciencias sociales y las humanidades en la publicación de libros. Este patrón es coincidente con lo observado en otros estudios para otros países (Giménez-Toledo et al. 2016).

Esta información no es relevante sólo en términos cuantitativos: debería permitir el reconocimiento, a veces no necesariamente conseguido, de la valía que aún hoy tiene el formato libro en las Ciencias Sociales y las Humanidades, por fuera de la predilección de los investigadores, las instituciones y las agencias de evaluación por la publicación en revistas académicas.

La manera en que se presentan estos datos, por país, por año y por materia, permite, además, las comparaciones que se consideren pertinentes. Bien sea por disciplinas, por años, por países o cruzando, por ejemplo, la producción de libros y revistas en determinadas disciplinas y periodos.

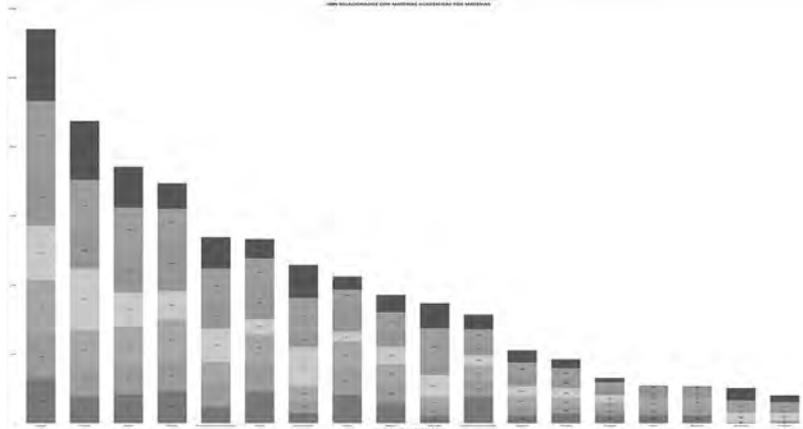


Figura 6. ISBN relacionados con materias académicas por materias
Fuente: Elaboración propia

ISBN solicitados a traducciones dentro de la producción académica

De los 73.772 asignados a libros relacionados con materias académicas, 4.392 indicaron que eran traducciones, el 6%. La Figura 7 expresa de manera detallada la trayectoria.

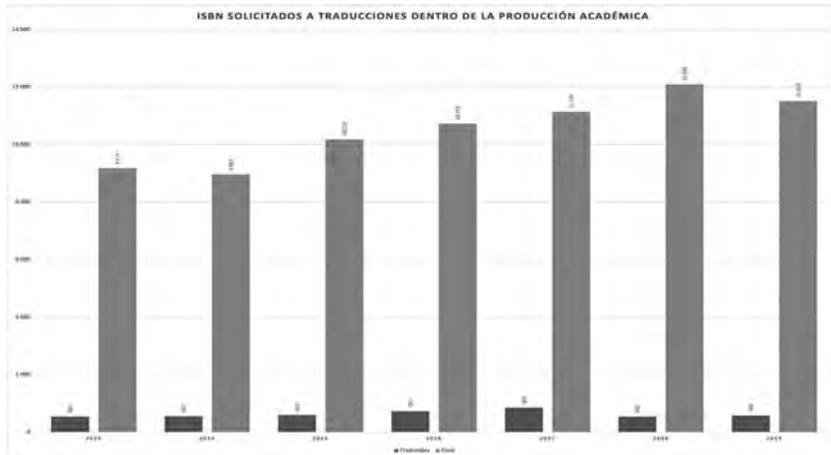


Figura 7. ISBN solicitados a traducciones dentro de la producción académica
Fuente: Elaboración propia

A lo largo del tiempo existe una estabilidad en la proporción de traducciones entre los libros académicos, oscila entre el 4% y el 8%, siendo el año que menos se tradujo 2018, con 542 libros, y siendo 2017 el año en el que más se asignaron ISBN a libros académicos traducidos, con 853.

Una exploración detallada de estos datos podrá dar cuenta de las influencias académicas en el país y los canales a través de los cuales se hacen posibles las traducciones. También es un punto de partida para establecer comparaciones con otros países en términos de la incidencia de otras tradiciones académicas, con otros idiomas, en el concierto editorial académico del país (por ejemplo, Dujovne 2017). Se presentan escuetamente los datos en este estudio, porque evidencia la potencia que la organización producto de la metodología ofrece para establecer análisis comparados entre países.

DISCUSIÓN, CONCLUSIONES Y FUTURO

A lo largo de este artículo, sección a sección, se han ido desgranando algunas consideraciones que podrían ser, cada una, objeto de análisis particulares y de discusiones relevantes. Se ha conformado con apenas indicar algunas de las hipótesis que los datos y la experiencia alumbran, como muestra de la efectividad que la metodología utilizada tiene para organizar la producción académica de un país y, desde allí, establecer comparaciones a lo largo del tiempo y entre países.

Uno de los puntos más preocupantes es, sin duda, la fiabilidad de los datos que los editores ingresan al solicitar el ISBN. De ahí se desprende la necesidad de insistir en la formación de aquellos que llevan a cabo esta labor. Y de establecer controles de calidad de los metadatos por parte de la Cámara Colombiana del Libro. Si estas páginas suscitan iniciativas en esa dirección estarían ya más que justificadas.

Por otro lado, los resultados muestran con claridad la importancia cada vez mayor de las editoriales universitarias. Y no sólo en el apartado específico de la edición de libros académicos, sino dentro del panorama de la edición en general. Es éste un llamado de atención a reconocer que, más allá de los grandes grupos editoriales y las iniciativas privadas, la producción de libros en Colombia depende al menos en una cuarta parte de agentes editoriales que publican sin que el ánimo de lucro sea su mayor interés. Es decir que, por ejemplo, la bibliodiversidad en Colombia, en gran medida, no depende del mercado sino de la labor las instituciones de educación superior. Por no hablar del aporte de las editoriales universitarias a la profesionalización de todo

el gremio editorial en Colombia. Y es un llamado de atención todavía más rotundo si se tiene en cuenta que la política del Ministerio de Ciencia, Tecnología e Innovación parece desestimular que las universidades publiquen libros a sus propios investigadores (MinCiencias 2021). Se espera que un estudio como éste aporte, desde la investigación académica, información y argumentos a los actores que, desde la edición académica, particularmente universitaria, deben implicarse en la evaluación y reformulación de estas políticas.

Por supuesto, no se puede soslayar la inquietud por el posible crecimiento artificial de esa producción, justamente porque tal aumento puede obedecer a que las universidades estén publicando más libros de los que, si hicieran un ejercicio de evaluación verdaderamente exigente, publicarían. Dicho en otros términos: los editores universitarios deberían preguntarse si el mandamiento del *publish or perish* ha hecho que los niveles de calidad del contenido que se publica hayan disminuido. Se trata, claro, de un cuestionamiento antipático que, por antipático, no debería dejar de ocupar un lugar privilegiado en las reflexiones de los autores, los evaluadores, los comités editoriales, las directivas y los editores universitarios (Uribe 2019). Y que de vuelta interpela a las autoridades nacionales en la medida en que, posiblemente, ha sido la política de reconocimiento de la producción académica del país la que haya llevado a esta condición.

Aparte de los comentarios que se han ido insertando en el texto, y que deben formar parte de la discusión en torno a la edición académica en Colombia, es importante apuntar que el siguiente paso en la investigación es la determinación del perfil de publicación de las editoriales para delimitar las *verdaderamente* académicas. Este primer análisis de metadatos permite acercarse a un subconjunto cercano a lo académico, pero es necesario un cribado más fino y un trabajo cualitativo —ya en curso— para avanzar en la caracterización de la edición académica, ahora desde la perspectiva de las editoriales, con mayor precisión.

Por último, se debe insistir, nuevamente, en el carácter seminal de estos resultados. No se pretende más que fundar un precedente para Colombia —ya consolidado en España con una metodología análoga—, que sea a su vez un piloto para la construcción de una cartografía de la edición académica iberoamericana. Estas líneas no son más que el intento de fundar un camino.

Agradecimientos

Este artículo no hubiera sido posible sin el concurso de diversos investigadores e instituciones. Agradecemos especialmente al grupo de investigación Cartografía de la edición académica iberoamericana, que está compuesto, además de los autores del artículo, por Sayri Karp de la Universidad de Guadalajara (México), José Diego González del Cerlalc, Alejandro Dujovne del Instituto de Desarrollo Económico y Social (Argentina) y Édgar Valencia de la Universidad Veracruzana (México). Nuestro reconocimiento y gratitud también van al Cerlalc, a la Universidad Santo Tomás (Colombia), a la Cámara Colombiana del Libro, a AseúC (Colombia), a Eulac, a la Red Alttexto (México), y a la Federación del Gremio de los Editores (España). El desarrollo de la investigación pudo hacerse gracias al apoyo de la convocatoria Capital Semilla de Investigación de la Universidad del Rosario (Colombia).

REFERENCIAS

- Arango, Juan Ignacio. 1991. *El libro en Colombia. Situación y perspectivas*. Bogotá: Cerlalc.
- Cerlalc. 2008. «Glosario de Términos». Cerlalc. 2008. <https://n9.cl/009p7>
- . 2020. *El espacio iberoamericano del libro 2020*. Bogotá: Cerlalc.
- Colciencias. 2013. «Convocatoria nacional para el reconocimiento y medición de grupos de investigación, desarrollo tecnológico y/o innovación y para el reconocimiento de investigadores del sistema nacional de ciencia, tecnología e innovación. 2013». <https://legadoweb.minciencias.gov.co/convocatoria/convocatoria-nacional-para-el-reconocimiento-y-medici-n-de-grupos-de-investigaci-n-des>
- . 2018. «Convocatoria nacional para el reconocimiento y medición de grupos de investigación, desarrollo tecnológico o de innovación y para el reconocimiento de investigadores del sistema nacional de ciencia, tecnología e innovación - SNCTEI 2018». <https://n9.cl/4ld8a>
- Dujovne, Alejandro. 2017. «Campo editorial y traducción». *Desarrollo Económico* 56 (220): 26. <https://n9.cl/qoyrx>
- Giménez Toledo, Elea. 2019. «La edición universitaria española. Análisis de la producción editorial de libros». <https://digital.csic.es/handle/10261/239787>
- Giménez-Toledo, Elea, y Jorge Mañana-Rodríguez. 2016. «La edición académica española a partir de sus metadatos». *Trama & Texturas*, no. 31: 125-31. <https://www.jstor.org/stable/26167553>
- Giménez-Toledo, Elea, Jorge Mañana-Rodríguez, Tim C. E. Engels, Peter Ingwersen, Janne Pölonen, Gunnar Sivertsen, Frederik T. Verleysen, y Alesia A. Zuccala. 2016. «Taking Scholarly Books into Account: Current Developments in Five European Countries». *Scientometrics* 107 (2): 685-99. <https://doi.org/10.1007/s11192-016-1886-5>

- Giménez-Toledo, Elea Giménez, Juan Felipe Córdoba Restrepo, Esteban Giraldo Giraldo-González, y Jorge Mañana Rodríguez. 2021. «Calidad y prestigio de la edición académica. El caso colombiano». *Signo y Pensamiento* 40 (78).
<https://doi.org/10.11144/Javeriana.syp.40-78.cpea>
- Giraldo-González, Esteban, Elea Giménez-Toledo, Juan Felipe Córdoba-Restrepo, Édgar García Valencia, José Diego González M, y Tatiana Morales Perdomo. 2021. «Tendencia Editorial UR Noviembre». Universidad del Rosario.
https://doi.org/10.12804/issne.2382-3135_10336.33150_teur
- Lado B. 2016. «Caracterización del sector editorial en Colombia». Investigación presentada a Grupo de Emprendimiento Cultural Ministerio de Cultura.
<https://economianaranja.gov.co/media/mhwbkiqe/a-7-8-1-caracterizaci%C3%B3n-del-sector-editorial-en-colombia.pdf>
- Liber, dir. 2021. *Cartografía de la edición académica iberoamericana*.
<https://www.youtube.com/watch?v=8ro-WZh0POs>
- Marín-Colorado, Paula Andrea. 2018. «Edición en Colombia (1970-1990):: del boom de la industria gráfica a la diversificación de la industria editorial». En *Lectores, Editores y Cultura Impresa En Colombia: Siglos XVI-XXI*, 1a. ed: 384-410. Bogotá: Universidad Jorge Tadeo Lozano.
- . 2020. «El libro en Colombia: Entre la sostenida concentración y la lenta marcha hacia la independencia (2000-2019)». *Amoxtli*, no. 5 (agosto): 39-58.
<https://revistas.uft.cl/index.php/amox/article/view/142>
- MinCiencias. 2021. «Convocatoria nacional para el reconocimiento y medición de grupos de investigación, desarrollo tecnológico o de innovación y para el reconocimiento de investigadores del sistema nacional de ciencia, tecnología e innovación - SNCTI 2021».
- Uribe, Alejandro. 2019. «Posibilidades de categorización de editoriales según su visibilidad e impacto. Una necesidad para la comunicación científica en Colombia e Iberoamérica. Propuesta metodológica». En *Sistemas de evaluación y edición universitaria*. Bogotá: Aseúc.
<https://n9.cl/uy9xns>

Para citar este texto:

- Giraldo-González, Esteban, Elea Giménez-Toledo, y Juan Felipe Córdoba-Restrepo. 2022. “Producción de libros académicos en Colombia entre 2013 y 2019: un avance para estudios comparados”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 153-176.
<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58633>

Anexo 1 Tabla de equivalencia de materias

Clasificación Dewey			Materia ILIA
Informática, conocimiento y sistemas			Biblioteconomía y Documentación
	4	Ciencia de los computadores	Informática
	5	Programación. Programas. Datos de computadores	Informática
	020	Bibliotecas y ciencias de la información	Biblioteconomía y Documentación
Filosofía y Psicología			Filosofía
	110	Metafísica	Filosofía
	120	Epistemología. Causalidad. Género humano	Filosofía
	140	Escuelas filosóficas específicas	Filosofía
	150	Psicología	Psicología
	160	Lógica	Filosofía
	170	Ética (Filosofía moral)	Filosofía
	180	Filosofía antigua. Medieval. Oriental	Filosofía
	190	Filosofía moderna occidental	Filosofía
Religión			Filosofía
	210	Filosofía y teoría de la religión	Filosofía
	230	Cristianismo y teología cristiana	Filosofía
	270	Historia del cristianismo y de la Iglesia cristiana	Historia
	280	Confesiones y sectas cristianas	Sociología
Ciencias sociales			Sociología
	310	Colecciones de estadística general	Sociología
	320	Ciencia política (Política y gobierno)	Ciencias políticas
	330	Economía	Economía
	340	Derecho	Derecho
	350	Administración pública y ciencia militar	Ciencias políticas
	360	Problemas y servicios sociales asociaciones	Sociología
	370	Educación	Educación
	380	Comercio, comunicaciones y transportes	Economía
	390	Costumbres, etiqueta, folclor	Sociología

Lenguas			Lingüística, literatura y filología
	410	Lingüística	Lingüística, literatura y filología
	420	Inglés e inglés antiguo	Lingüística, literatura y filología
	430	Lenguas germánicas Alemán	Lingüística, literatura y filología
	440	Lenguas romances Francés	Lingüística, literatura y filología
	450	Italiano. Rumano. Retorromano	Lingüística, literatura y filología
	460	Lenguas española y portuguesa	Lingüística, literatura y filología
	470	Lenguas itálicas Latín	Lingüística, literatura y filología
	480	Lenguas helénicas Griego clásico	Lingüística, literatura y filología
	490	Otras lenguas	Lingüística, literatura y filología
Ciencias naturales y Matemáticas			Ciencias
	510	Matemáticas	Ciencias
	520	Astronomía y ciencias afines	Ciencias
	530	Física	Ciencias
	540	Química y ciencias afines	Ciencias
	550	Ciencias de la tierra	Ciencias
	560	Paleontología Paleozoología	Ciencias
	570	Ciencias de la vida Biología	Ciencias
	580	Plantas	Ciencias
	590	Animales	Ciencias
Tecnología (Ciencias Aplicadas)			Tecnología
	610	Ciencias médicas Medicina	Medicina
	620	Ingeniería y operaciones afines	Ingenierías
	630	Agricultura y tecnologías relacionadas	Agricultura
	640	Economía doméstica y vida familiar	Economía
	650	Gerencia y servicios auxiliares	Economía
	660	Ingeniería Química	Ingenierías
	670	Manufactura	Ingenierías

	680	Manufactura para usos específicos	Ingenierías
	690	Construcción	Ingenierías
Las artes, bellas artes y artes decorativas			Bellas artes
	710	Urbanismo y arte paisajístico	Bellas artes
	720	Arquitectura	Bellas artes
	780	Música	Bellas artes
Geografía e historia			Historia
	930	Historia del mundo antiguo hasta a.C.	Arqueología y prehistoria
	940	Historia general de Europa	Historia
	950	Historia general de Asia Lejano Oriente	Historia
	960	Historia general de África	Historia
	970	Historia general de América del Norte	Historia
	980	Historia general de América del Sur	Historia
	990	Historia general de otras áreas	Historia

Uso de las habilidades digitales en el proceso de enseñanza-aprendizaje en ciencias de la información en un entorno virtual durante la pandemia por Covid 19

Eduardo Oliva-Cruz*
Adriana Mata-Puente*

Artículo recibido:
31 de marzo de 2022
Artículo aceptado:
5 de septiembre de 2022
Artículo de investigación

RESUMEN

El objetivo de este trabajo es analizar las habilidades que han desarrollado los docentes y estudiantes de la Facultad de Ciencias de la Información de la Universidad Autónoma de San Luis Potosí (UASLP) a partir del confinamiento por la contingencia sanitaria por Covid 19. Se analizan 130 estudiantes de esa facultad, 10 docentes y 17 materias impartidas durante el periodo de marzo de 2020 a junio de 2021. Se aplicaron encuestas que indagan sobre las habilidades tecnológicas y didácticas puestas en práctica durante 12 meses. En el caso de los estudiantes, además se indagó sobre su

- * Facultad de Ciencias de la Información, Universidad Autónoma de San Luis Potosí, S.L.P., México eduardo.oliva@uaslp.mx
- * Facultad de Ciencias de la Información, Universidad Autónoma de San Luis Potosí, S.L.P., México amata@uaslp.mx

situación económica y su acceso a internet. Se concluye que tanto docentes como estudiantes se han adaptado al proceso de enseñanza-aprendizaje en línea; las competencias digitales son adecuadas, pero no suficientes para alcanzar los objetivos del aprendizaje.

Palabras clave: Gestión de la Información - Educación en línea; Covid 19 y Educación; Habilidades Digitales; Educación a Distancia; Aprendizaje en Entornos Virtuales

Use of digital skills in the teaching-learning process in a virtual environment during the Covid 19 pandemic

Eduardo Oliva-Cruz and Adriana Mata-Puente

ABSTRACT

The objective of this work is to analyze the skills that the teachers and students of the Faculty of Information Sciences of the Autonomous University of San Luis Potosí (UASLP) have developed since the confinement due to the health contingency by Covid 19. 130 students of that faculty, 10 teachers and 17 subjects taught during the period from March 2020 to June 2021, surveys were applied that inquire about the technological and didactic skills put into practice for 12 months. In the case of students, the economic situation and internet accessibility were also inquired about. It was concluded that both teachers and students have adapted to the online teaching-learning process and that although their digital skills are adequate, this is not enough to achieve the learning objectives.

Keywords: Information Management - Online Education; Covid 19 and Education; Digital Skills; Distance Education; Learning in Virtual Environments

INTRODUCCIÓN

El estudio parte de la siguiente interrogante: ¿las habilidades digitales son suficientes para enseñar y aprender en un entorno virtual derivado

de una emergencia externa, o es necesario otro tipo de habilidades y/o conocimientos? Para analizar el problema, se parte del contexto en el que se encuentran inmersos los actores, desde la sociedad en su conjunto, hasta el espacio en el que desempeñan sus roles como docentes y estudiantes, y considerando los efectos que tuvo la pandemia del Covid 19 en el ámbito educativo y sus repercusiones en los modos de actuar de los actores estudiados.

Cabe señalar que, desde el inicio del confinamiento, la Universidad Autónoma de San Luis Potosí (UASLP) ha pasado por tres etapas de adaptación y planeación para la incorporación paulatina a las actividades de manera presencial. La primera de ellas fue en el semestre enero – junio 2020. El confinamiento inició en marzo de ese año, cuando tomó a todos por sorpresa. En ese momento se pensaba en un confinamiento de tres a cuatro semanas, por tanto, se podría regresar a las aulas para continuar el ciclo escolar como se venía realizando de manera presencial. La prolongación de la pandemia y la baja probabilidad del regreso a las aulas obligó a todos a buscar otras formas de interacción social. Por lo que, en esta etapa, la preocupación era dar continuidad al semestre: cada quien hizo lo que pudo, estableciendo la estrategia de comunicación a su alcance.

Una segunda etapa, correspondiente al semestre agosto – diciembre de 2020, se inicia con el diseño de actividades a distancia, sin fecha de retorno a las actividades presenciales. El inicio de un ciclo escolar y alumnos de nuevo ingreso a la universidad representaron un reto y, al mismo tiempo, una oportunidad para innovar; los docentes ya contaban con más herramientas para el diseño de actividades y mantener la comunicación con los estudiantes. En el caso de los estudiantes, ya estaban familiarizados con los recursos tecnológicos que se usarían para las clases, y la institución ya había implementado estrategias para facilitar la interacción entre alumnos y maestros.

La tercera etapa se refiere al semestre de enero a junio de 2021. Los docentes y los estudiantes ya se han ido adaptando a la nueva normalidad. El trabajo a distancia se ha ido mejorando, sobre todo por la experiencia del año 2020; también, algunos alumnos en su momento pidieron incorporarse a actividades de manera presencial; el protocolo del uso de laboratorios y otros espacios universitarios permite atender a los estudiantes garantizando su salud. En esta etapa se puede afirmar que se fortalecieron las actividades presenciales y a distancia en beneficio del proceso de enseñanza-aprendizaje en la Universidad.

En ese escenario surge la inquietud de reflexionar sobre las habilidades de los actores de ese proceso para lograr el objetivo de enseñanza-aprendizaje. Las estrategias implementadas van orientadas a adaptarse a las exigencias del mundo virtual, lo que supone cambios en el proceso educativo. Este

trabajo es un estudio exploratorio. Para contar con datos estadísticos sobre las habilidades digitales es necesario hacer un análisis más profundo de las estrategias implementadas por alumnos y docentes. La investigación es de corte cualitativo para describir las acciones implementadas por maestros y estudiantes en su proceso formativo.

El contenido del artículo se organiza en tres apartados. En el primero se señalan aquellos autores que han analizado el tema. Coinciden en señalar que “habilidades digitales” es un término más específico para describir el uso que se hace de la tecnología en las actividades educativas; involucra el dominio de la habilidad tecnológica hasta llegar al uso que se le da en el contexto educativo. El siguiente apartado describe la metodología empleada de corte cualitativo para indagar el sentir de los actores durante el proceso de cambio y adaptación de la educación, de su modalidad presencial a la virtual, y sus diferentes variantes en función del dominio de los agentes involucrados como: disponibilidad tecnológica, accesibilidad, estrategias didácticas adaptadas a la modalidad virtual, el uso de diferentes herramientas tecnológicas y el logro de los aprendizajes esperados. Finalmente, se presentan los resultados, donde se analizan las acciones que cada actor puso en práctica para cumplir con sus objetivos; si de estudiantes se trata, están aquellos que se adaptaron fácilmente, hasta llegar al otro extremo, donde decidieron pausar sus estudios. En el caso de los docentes, están quienes simplemente dejaron que la vida siguiera su rumbo, hasta llegar a aquellos docentes que rápidamente se involucraron en la educación virtual, poniendo todo su empeño en resolver los inconvenientes que el cambio abrupto les presentaba. Se llega a la conclusión que el camino no fue fácil, requirió de esfuerzo, dedicación y adaptación a las exigencias que demandaba el trabajar a distancia.

LAS HABILIDADES DIGITALES

Desde la década de los años 80 del siglo XX, las TICs se han integrado a las actividades cotidianas de las personas en general. Entendiendo éstas como todas aquellas “tecnologías de la comunicación, incluidas Internet, redes inalámbricas, teléfonos celulares, computadoras, software, middleware, videoconferencias, redes sociales y otras aplicaciones y servicios de medios que permiten a los usuarios acceder, recuperar, almacenar, transmitir y manipular información en forma digital” (World Business Council for Sustainable Development, 2012: 3). El uso de éstas se ha ido incrementando, adquiriendo cada día un papel preponderante en el devenir de la sociedad, no importando los ámbitos en los que se encuentren, sean éstos de carácter social, cultural, económico, político

o educativo entre otros. Bien sea en una sociedad o una comunidad con menor o mayor progreso tecnológico en alguno de esos ámbitos. Tan es así que en comunidades que están alejadas de las grandes ciudades, y que parecieran más desfavorecidos o con poco progreso, el uso de las TICs, así sea por un miembro de la comunidad que pueda acceder a contenidos de información y conocimiento, usando cualquier tipo de dispositivo, marca una diferencia en el contexto en el que se está inmerso.

En el ámbito de la educación superior, contar con habilidades digitales por lo menos al ingreso a un centro de formación universitario es un requisito necesario. Entendiendo que el estudiante puede hacer uso de las TICs para buscar, localizar, recuperar y analizar información no importando el formato digital en el que se encuentre (sonido, imagen o texto), así como para comunicarse y colaborar con otros semejantes; esto con el fin de satisfacer necesidades particulares sean para la formación y/o recreación, atendiendo necesidades específicas o colectivas como usuarios de la red. Así mismo, el que cuenten con nociones básicas sobre el uso de los diferentes equipos (computadora, *tablet*, celular...) y los programas que permiten una mejor inmersión en el medio digital.

Quienes han analizado el tema de las competencias digitales en el contexto del covid-19 son: Díaz-Arce y Loyola-Illescas (2021); George Reyes (2021); Holguin-Alvarez, Jhon *et al.* (2021); y Sierralta Pinedo (2021). Estos autores hacen un análisis de las implicaciones que la educación a distancia tuvo en docentes, estudiantes, directivos e instituciones educativas, durante la pandemia del Covid 19. Coinciden en señalar las adaptaciones que cada actor tuvo que implementar para continuar en la educación en la modalidad virtual. En ese escenario las diferencias se vieron más marcadas entre quienes contaban con los recursos y quienes no tenían capacidad tecnológica, lo que incidía en la atención pedagógica de los estudiantes.

Los docentes juegan un papel importante en el proceso de aprendizaje de sus estudiantes, su rol inicia desde la planeación de sus cursos hasta la motivación para atrapar su atención, pasando por los procesos de evaluación y retroalimentación. Al analizar su trabajo durante la pandemia, algunos tuvieron que adaptar las estrategias para el aprendizaje; ahora debían responder al trabajo asincrónico como el aprendizaje basado en problemas o el aula invertida (Díaz-Arce y Loyola-Illescas, 2021). El docente pasa a un segundo plano y el estudiante es quien dirige su aprendizaje a través de la resolución de situaciones problemáticas.

En el caso de los estudiantes ¿cuál es la percepción de las estrategias de enseñanza-aprendizaje implementadas por los docentes? George Reyes (2021) hizo un análisis de las competencias digitales básicas para garantizar

la continuidad académica provocada por el Covid 19 en estudiantes de bachillerato. Analizando la interacción comunitaria para el aprendizaje, la selección de materiales y el uso de herramientas tecnológicas para la evaluación lo que se encontró fue una aceptación de las actividades realizadas por los docentes con herramientas digitales. Las áreas de oportunidad se dieron en cuanto a las acciones para la evaluación y el aprovechamiento eficiente de las herramientas para debatir con los compañeros.

Con respecto a las “competencias digitales”, los autores analizados señalan que no existe un consenso en su definición, ya que en la literatura sobre el tema existen otras denominaciones como “alfabetización digital” y “alfabetización informacional” lo que ocasiona confusión para identificar la especificidad de cada uno de esos conceptos. La competencia digital es multidimensional

[...] en el que se engloba un conjunto de habilidades y actitudes interrelacionadas entre sí que abarcan aspectos técnicos, informacionales, creación de contenidos, mediáticos, comunicativos, solución de problemas, así como la toma de decisiones estratégicas y éticas (Reyes, 2021: 126).

Dado el carácter multidimensional del término “competencias digitales”, se tiende a generar confusión para identificar los elementos sustantivos que lo integran. Ya que toca diferentes aspectos que tienen relación con la tecnología, el uso y análisis de la información, así como de su aplicación en el contexto educativo. fue que se consideró el término “habilidad digital” para hacer una distinción respecto al término “competencia digital”, ya que éste presenta una mayor complejidad para su uso, desde diferentes visiones epistémicas.

Al hablar de habilidades digitales en el contexto educativo es necesario poner en el centro del análisis a los actores (estudiantes y profesores). El estudiante aun y cuando se da por hecho que domina la tecnología lo cierto es que, para el uso con fines de aprendizaje, muestra pocas habilidades particularmente para la apropiación de contenidos. En lo referente a los docentes, enfrentan un triple reto en su práctica educativa, deben poseer no solamente el conocimiento de su disciplina, sino también saberes pedagógicos y tecnológicos (Díaz-Arce y Loyola-Illescas, 2021).

Desde esa perspectiva y de acuerdo con Lemus Pool “el término habilidades digitales nos permite alejarnos de la problemática de la definición de competencias” (2017: 117), ya que se buscó obtener claridad y pertinencia en el presente estudio. Esta autora declara que son “el conjunto de saberes y capacidades que se requieren para generar procesos didácticos en el entorno virtual” (2017: 118), esto en primera instancia para los docentes; pero para los estudiantes, más que “saberes y capacidades [...] para procesos didácticos”,

esos saberes y capacidades se deben considerar para procesos de aprendizaje de los contenidos ofrecidos en un curso lectivo en función de estrategias pedagógico-didácticas adaptadas al medio digital. Por su parte, los alumnos deben mostrar sus destrezas en el uso y manejo de las TICs para su aprendizaje en este contexto.

Así mismo, cabe mencionar que las habilidades digitales están integradas por diferentes dimensiones para una mejor comprensión de las mismas. Lemus Pool (2017) las sintetiza en las siguientes: instrumental, cognitiva, comunicativa, ética y didáctica, proporcionando una breve definición de cada una de éstas.

- *Instrumental*: referida al conocimiento, dominio técnico y uso de las tecnologías;
- *Cognitiva*: habilidad para desarrollar procesos de búsqueda, análisis y procesamiento de diferentes fuentes de información;
- *Comunicativa*: habilidades para generar la retroalimentación, el intercambio y gestionar la interacción didáctica en los ambientes virtuales;
- *Ética*: el respeto a las fuentes de información; y, por último,
- *Didáctica*: está asociada con las estrategias de enseñanza aprendizaje apoyadas en las TICs (2017: 124-5).

De acuerdo con los autores analizados, se da evidencia sobre el uso del término “habilidades digitales” en el contexto del presente estudio, derivado de la situación de premura que exigía dar respuesta puntual e inmediata sin un proceso de reflexión y planeación, y más bien se iba actuando conforme se presentaron los eventos. Cada actor asumió, desde su rol, la responsabilidad y el reto de dar respuesta a la circunstancia vigente en el dominio de la tecnología (instrumental) y las estrategias de enseñanza-aprendizaje (didáctica).

METODOLOGÍA

Las habilidades digitales se analizan desde dos perspectivas, una es la pedagógica, donde se hace la reflexión del modelo educativo a distancia y las estrategias de enseñanza-aprendizaje; desde el ámbito de la gestión de la información, el uso eficiente de los recursos disponibles para ofertar una educación de calidad. Esta investigación es de carácter exploratorio con el objetivo de dar cuenta sobre cómo se adaptaron los actores del proceso educativo ante una situación emergente. Partimos del supuesto de que el acontecimiento no

previsto y de apremio para los estudiantes y docentes no los detuvo, por el contrario, se adaptaron a las circunstancias conforme se fueron presentando. Este documento da cuenta de la experiencia de los actores ante esa realidad, así como la incorporación de tecnología en su proceso educativo y los cambios que el mundo virtual exige.

El periodo en el que se realizó el estudio abarcó desde el inicio del confinamiento en marzo de 2020, a junio de 2021. El universo fue la Licenciatura en Gestión de la Información de la Facultad de Ciencias de la Información de la Universidad Autónoma de San Luis Potosí. Cada semestre se imparten 52 asignaturas a cargo de aproximadamente 20 profesores. Se seleccionó de cada semestre una muestra de estudio intencionada, tratando de abarcar docentes con diferentes habilidades digitales, así como materias con contenidos en: administración, normalización, organización y servicios. La muestra estuvo integrada por 10 profesores, impartiendo 17 materias diferentes, y 130 estudiantes de ese programa educativo. La muestra se integró de la siguiente forma:

Semestre	Profesores	Materias	Alumnos
Marzo – junio 2020	3	5	35
Agosto – diciembre 2021	4	8	65
Enero – junio 2021	3	4	30
Total	10	17	130

Tabla 1. Muestra de estudio
Fuente: Elaboración propia

Se aplicaron dos instrumentos donde se indagó sobre las habilidades digitales y pedagógicas, así como de aprendizaje y de contexto durante el proceso de formación en línea. Con base en las dimensiones propuestas por Lemus Pool (2017), se elaboraron los instrumentos para la recolección de datos, así como para el análisis de los mismos. Se determinó como categorías de análisis la dimensión instrumental, referida al conocimiento, dominio técnico y uso de las tecnologías; y la dimensión didáctica, asociada con las estrategias de enseñanza-aprendizaje apoyadas en las TICs.

Para el instrumento aplicado a los profesores se tomaron en cuenta únicamente las dimensiones: instrumental y didáctica y se agregó, siguiendo George Reyes, la habilidad de “selección y diseño de materiales digitales” (2021) para apoyar la dimensión didáctica (de enseñanza). En el instrumento aplicado a los estudiantes se cambió la dimensión didáctica por la de aprendizaje; de acuerdo con Parra Rodríguez, se entiende ésta como el cambio que

manifiesta el alumno a través de sus acciones con relación a un estado anterior personal y que de alguna manera se pueda detectar que es permanente (2005). Además, se incluyó el contexto socioeconómico y la accesibilidad a internet. El instrumento se integró por tres partes: contexto de los actores, proceso de comunicación y las estrategias didácticas para el aprendizaje en docentes y estudiantes. Las preguntas fueron abiertas con la finalidad de dar la oportunidad para exponer la experiencia en el proceso de enseñanza-aprendizaje en el que se encontraron inmersos.

La aplicación del instrumento se hizo de manera virtual. En el primer semestre se elaboró en Word y se envió por correo electrónico a los actores para ser respondido, lo que implicó que se enviaran recordatorios para tener el 100% de la muestra. A partir del siguiente semestre, el instrumento se envió a través de la plataforma Teams, la cual permite enviar recordatorios a quienes no habían contestado, obteniendo nuevamente el 100% de respuestas. En lo que concierne al análisis de los datos se transcribieron cada una de las respuestas y se organizaron conforme a las categorías antes mencionadas.

LA CONTINGENCIA SANITARIA SARS-COV 2 Y EL CAMBIO EN EL PROCESO DE ENSEÑANZA-APRENDIZAJE

Ante la contingencia sanitaria, uno de los primeros escenarios a enfrentar por parte de los estudiantes y profesores fue el de guardar confinamiento en sus casas por un tiempo indeterminado. En un principio, en un escenario muy favorable se preveía que ese confinamiento llevaría unas cuatro semanas; pero una vez que se tuvo más información al respecto y que se continuaría hasta concluir ese semestre, implicó un cambio de modalidad que poco a poco se fue perfeccionando.

Por lo anterior se hizo necesario y hasta urgente asumir otra dinámica para dar continuidad con ese proceso de educación, así como para cumplir con los programas propuestos en los diferentes semestres, para lo cual los docentes deberían generar nuevas estrategias en su práctica didáctica (Holguin-Alvarez et al., 2021). Ante el confinamiento la única alternativa fue la de optar por el uso de las TICs como herramienta y sus habilidades digitales aplicadas en el ámbito de la educación para transitar de un modelo presencial a uno a distancia.

Proceso para iniciar la educación en situación emergente

Como ya se señaló, la situación emergente inicia prácticamente a mitad del semestre, de cierta forma eso representó una ventaja; el docente ya tenía conocimiento de sus estudiantes, es decir, ya se había establecido el proceso de comunicación, quienes conocían el programa de la asignatura, se había avanzado un 50% del curso. Concluir el ciclo escolar a distancia generaba incertidumbre de cómo se iba a llevar a cabo; en un primer momento fue de angustia pues no se estaba preparado para dar continuidad al semestre de manera virtual.

En este punto es necesario señalar que, en cada inicio de curso, los docentes presentan el plan de trabajo, los objetivos a lograr, el sistema de evaluación, así como la atención de tipo académico que recibirán a lo largo del periodo de enseñanza-aprendizaje. Uno de los puntos que se señalan al inicio, es la comunicación que se dará durante ese periodo, en el cual el docente, además de atender a sus alumnos durante el tiempo de clase, se cuenta con asesorías y/o sesiones de tutoría donde se le proporciona apoyo adicional al estudiante. Frente a ese panorama el primer reto fue la comunicación con éstos vía remota, estrategia necesaria para mantener el contacto entre los principales actores (Holguin-Alvarez et al., 2021). Este tipo de comunicación, asincrónica fuera de tiempo presencial, a través de correo electrónico, ya que de inicio en los cursos se les envía información sobre las diferentes asignaturas como lo es el programa, avisos y material de apoyo para la marcha de programa. No todos los estudiantes usaban la cuenta de correo institucional, por lo que representó un reto lograr la comunicación con todos a través de ese medio.

También se propuso la comunicación (sincrónica), la cual supliría la presencialidad en el aula física. Esto es a través de sistemas de video-conferencia, de los cuales se propusieron los siguientes recursos Zoom, Google Meet y Skype, así como los espacios que ofrece Facebook como “aulas virtuales”. En las asignaturas analizadas, en principio se optó por Zoom y Facebook con el conocimiento que tenían 70% de los docentes y el 80% de los estudiantes, además de su disponibilidad y acceso para todos. Los estudiantes mencionaron que se consumían muchos bytes o datos¹; para los alumnos que viven en zonas aledañas a la ciudad implicaba menor conexión a las sesiones sincrónicas porque la red fallaba y se les iba la sesión en tratar de conectarse nuevamente en cuanto los sacaba de la red.

1 Coloquialmente se utiliza el término datos para decir que tienen crédito o aún tienen saldo en sus cuentas de teléfono celular (dispositivo móvil) y pueden continuar con la conexión a la red.

Un aspecto que influyó en este proceso fue el económico, ya que algunas aplicaciones consumían muchos datos (Zoom), aspecto que se complicaba debido a la situación económica de las familias, quienes al dedicarse al sector servicios se vieron afectados en sus ventas durante la pandemia. Es así como algunos estudiantes tuvieron la necesidad de combinar el estudio con el trabajo para apoyar la economía familiar y contribuir con el gasto. En otros casos, la situación económica de la familia no tenía la capacidad para aumentar el consumo de internet o invertir en equipos tecnológicos para continuar con las clases a distancia.

Por último, ante ese escenario en línea, las estrategias de intervención didáctica debían cambiar radicalmente; no es lo mismo estar en el aula de manera presencial interactuando con los participantes del curso a estar como observador del monitor escuchando la exposición del docente. Fue necesario cambiar la estrategia de enseñanza para captar la atención de los estudiantes, cambiando la figura del docente como guía en el aprendizaje y centrando el proceso de enseñanza en el estudiante (Díaz-Arce y Loyola-Illescas, 2021; George Reyes, 2021) quienes se convirtieron en actores de su propio aprendizaje.

En ese mismo tenor, la evaluación tuvo que adaptarse a un entorno en línea (George Reyes, 2021). Para lo cual fue necesario considerar las siguientes dimensiones: la antropológica, la funcional, los medios y recursos utilizados, y por último los contenidos de las asignaturas impartidas.

En la dimensión antropológica, particularmente en su relación como personas, es decir, dimensión donde se pone atención al estudiante no solamente en su desempeño intelectual y académico sino de manera paralela se presta atención a su parte emocional y afectiva en relación con su desempeño personal. En lo que respecta a la dimensión funcional tiene que ver con las diferentes actividades didácticas que se diseñaron para el aprendizaje de los contenidos; con esto se relacionó con los productos de desempeño (exámenes en línea, ensayos, reportes de lectura, observación y análisis de videos, entre otros productos) señalados como elementos para dar cuenta de ese aprendizaje. En la dimensión de medios para la interacción educativa se tuvo acceso a diferentes plataformas electrónicas, una institucional, otra por contrato, y las libres mencionadas anteriormente: DidacTic institucional, y Microsoft Teams, con apoyo de los recursos electrónicos que ofrecen éstas; los recursos más utilizados fueron los programas de Office de la segunda plataforma. Por último, en lo referente al contenido de las asignaturas del estudio, están alineadas a las líneas curriculares que integran el plan de estudios, cuyos contenidos son: administración, normalización, organización y servicios.

De los profesores

En el instrumento aplicado a los maestros emergieron tres aspectos: instrumental, didáctica y de aprendizaje. En el primero se refiere al uso de la tecnología y las habilidades digitales para incorporarlo a la práctica docente; en el didáctico hace alusión a la incorporación de la enseñanza a distancia en sus cursos, por ejemplo, grabar videos con la explicación de algunos temas y subirlos a la plataforma para que los estudiantes los consulten en cualquier momento, uso de guías instruccionales para orientar las acciones que deben emprender sus estudiantes en su proceso formativo; en el aprendizaje, se refiere a aprender a aprender, el docente junto con los alumnos consideran el estar atentos a los aspectos que pueden mejorar para enriquecer su práctica docente (Lemus Pool, 2017; Sierralta Pinedo, 2021).

En los resultados, se identifican tres grupos de maestros (véase *Figura 1*), si se visualiza en una barra en un extremo están aquellos docentes que se han quedado en el límite del acceso a la tecnología, representado por un 20% de la muestra, utilizan el correo o algún sistema de comunicación como el WhatsApp. El uso de otras herramientas tecnológicas y el acceso a recursos digitales como pueden ser blogs, videos en línea o páginas web que sean pertinentes para su curso, entre otros similares, representa un esfuerzo para entenderlos e integrarlos en la planeación de sus cursos. Esos maestros se conectan diariamente a las clases y las siguen desarrollando en manera magistral, apoyados en el mejor de los casos con alguna presentación, pero todo el tiempo de clase es exposición. Esta forma de comunicación representa todo un reto para captar la atención de sus estudiantes. Las estrategias de evaluación para mediar los aprendizajes siguen siendo los exámenes en línea, pero se aplicaron exámenes orales, con el fin de evitar la copia, lo cual resulta tedioso para los estudiantes y poco adecuado para la modalidad en línea.

En un segundo grupo de docentes, el 50%, se encuentran los que han desarrollado las habilidades digitales y poco a poco las han incorporado a su práctica docente, saben usarlas eficientemente para el cumplimiento de los objetivos del curso, han cambiado la enseñanza tradicional en el salón de clase para adaptarse al contexto digital; estos docentes se encuentran en un proceso de aprendizaje y están incrementando el uso pedagógico de las TICs, y han tomado cursos para mejorar sus habilidades digitales y pedagógicas. Se puede afirmar que este grupo realizó su mejor esfuerzo para integrarse a la educación en línea.

Finalmente, el último grupo, el 30% de los docentes que dominan las habilidades tecnológicas (instrumentales) y las incorporan a su planeación docente, sea porque ya los impartían en línea o porque se han adaptado fácilmente a esta

modalidad de enseñanza, reconocen que la educación a distancia es muy diferente a la educación presencial. Para mantener el contacto con los estudiantes hacen presentaciones o videos donde dan la bienvenida al curso, explican algunos temas clave, además, diseñan diversas actividades donde favorecen la comunicación entre los mismos alumnos a través de chats y otras aplicaciones para compartir sus dudas e inquietudes, hacen uso de recursos que integran otros medios además del simple texto, o la conexión a través de una cámara para la cátedra.

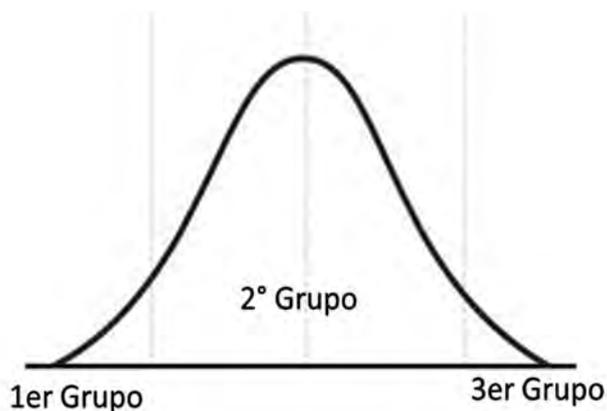


Figura 1. Los docentes y su proceso de adaptación a la enseñanza en línea
Fuente: Elaboración propia

Como se ha señalado, los docentes se tuvieron que adaptar a las exigencias de la educación en línea, incrementando sus habilidades en el uso de la tecnología, sobre todo en cómo usar estas herramientas de información y comunicación para la intervención pedagógico-didáctica que incida en el aprendizaje de sus estudiantes, modificando su práctica docente con el fin de responder a las expectativas de sus estudiantes en el entorno digital.

De los estudiantes

En la categoría del contexto los estudiantes también tuvieron que adaptarse, (véase *Figura 2*). Primero en la situación económica que se vieron afectados por el confinamiento, como se describió en líneas previas, la actividad económica de sus padres se ubica en los servicios y comercio, la venta de productos es su actividad primordial. En otros casos, al reducirles la jornada de trabajo, por ende se les redujo el salario, lo que implicó un cambio en la economía familiar. Los alumnos de la FCI, entre el 50 y 60% trabaja para solventar sus

gastos; durante la pandemia también vieron reducido su sueldo, en los casos más drásticos se quedaron sin empleo.

Otra situación que enfrentaron los estudiantes fue la disponibilidad de equipo tecnológico para conectarse a las clases virtuales, el 70% manifestó que en su familia disponían de un equipo de cómputo portátil o de escritorio; el 50% lo compartía con otros miembros de la familia, por lo que se debían organizar en el uso conforme a horarios para conectarse en clase. El 98% cuenta con teléfono celular, el cual usa para conectarse a clase. Aunado a ello, el 35% de los alumnos son de comunidades rurales, por lo que la conectividad es más lenta.

Frente a ese panorama, la falta de equipo de cómputo actualizado y problemas con la conectividad representó un problema a la hora de conectarse a las sesiones sincrónicas de la clase, lo que impidió participar activamente. Si el docente grababa la sesión, facilitaba el acceso a los estudiantes, una vez resuelta la conectividad, pero si no, se quedaron grabadas las sesiones, lo que ocurrió en casi el 50% de las clases, el estudiante no pudo ponerse al corriente.



Figura 2. Aspectos que influyen en los docentes y estudiantes
Fuente: Elaboración propia

El 60% de estudiantes manifestaron no tener tanto problema con las habilidades digitales, dominan el acceso a las redes sociales, a las plataformas y las habilidades en la búsqueda de información necesaria. Cabe señalar que fue necesario un proceso previo de aprendizaje, dependiendo del sistema o programa que se propuso para la interacción formativa, lo cual impactó en el tiempo de seguimiento del curso. Se detectaron discrepancias en el uso de diferentes dispositivos electrónicos, así como en el uso de las plataformas

propuestas por sus docentes; esto conflictuó el desempeño de algunos estudiantes, sobre todo cuando en los celulares no se podía visualizar igual que en una computadora o laptop.

Otro aspecto fue la necesidad de demostrar habilidades que se supone ya deberían poseer y que no son propiamente ubicadas como digitales, como son las de lectura y escritura, así como habilidades de manejo del tiempo y el poder cumplir con las metas propuestas para la entrega en tiempo de las actividades asignadas. Otra habilidad que se detectó fue la de autogestión del aprendizaje autónomo. Los alumnos por momentos quedaban a la expectativa de la actividad que el docente enviara al repositorio para su consulta. Y donde se detectó pensamiento particionado, es decir, que no se notaba en la construcción de su aprendizaje una continuidad con respecto a los conocimientos previos y conectarlos con el siguiente asunto o tema a revisar durante el curso. Esto representó todo un reto para éstos, ya que están acostumbrados a ser más receptivos sin reflexionar sobre el ejercicio o actividad propuestos, limitándose a cumplir con lo que se les pide; con este cambio de modalidad en línea, implica un papel mucho más activo para conocer el tema previo a la sesión sincrónica y estar en posibilidad de plantear todas sus dudas en el momento que el docente estaba conectado para la asesoría, aun y cuando se usaron diferentes canales de comunicación (George Reyes, 2021), las sesiones asíncronas no permitían aclarar las dudas que se presentaban en el momento de generar los productos necesarios para su evaluación.

CONCLUSIONES

Con el análisis del proceso se detectó la necesidad de contar con otros conocimientos y habilidades; son solamente en relación con las digitales ya que de alguna manera contribuyen a fortalecer estas últimas. A continuación, se presentan unas conclusiones de ese análisis.

En los docentes, se notó la necesidad de trabajar de manera colaborativa con los pares académicos. Entendiendo que existe un conocimiento previo para poder hacer uso de esta actividad. Ya que una buena parte de los estudiantes coincidieron en expresar que se sintieron abrumados en relación con la carga de trabajo, así como de las diferentes actividades a realizar para el aprendizaje de los contenidos de las asignaturas que tenían inscritas. Poco a poco se fueron implementando distintas estrategias para la impartición de cursos, en cuanto a una modalidad a distancia, de acuerdo con los objetivos del curso y las características propias del grupo, así como el uso de diferentes recursos tecnológicos lo cual generó certidumbre entre los actores

para continuar con el proceso de enseñanza-aprendizaje (Díaz-Arce y Loyola-Illescas, 2021; Sierralta Pinedo, 2021),

Una habilidad muy marcada en los estudiantes fue la de aprender a manejar el tiempo, esto se notó cuando éstos se abrumaban con la entrega de evidencias de desempeño (trabajos, actividades y exámenes en línea, entre otros), pero conforme trascurría el proceso del curso, poco a poco fueron organizando sus tiempos, lo cual permitió, al final del semestre el que llevaran un control de sus actividades con relación al tiempo, el que los estudiantes agendaran sus actividades y la dedicación de las mismas para concluir las conforme a lo planeado. Otro aspecto en el manejo del tiempo es que no se respetaron los horarios de clase previamente determinados para la asistencia a clase presencial, lo que ocasionó que ellos tuvieran dos clases en el mismo horario; al solventar ese punto, se pudo continuar con la realización de actividades respetando los tiempos asignados desde el inicio del semestre.

En ese mismo tenor, se detectó la falta de rutinas de trabajo autónomo en la cotidianidad del día a día de los diferentes actores del proceso. Una vez organizados, permitió seguridad y certidumbre para la buena marcha del proceso. Una habilidad que no se pudo mejorar fue la lectura y la escritura en el ámbito digital. Por tanto, es necesario reforzar esas habilidades en particular. Exige otro tipo de atención, básicamente para el análisis de los documentos (imagen, sonido o texto) para realizar adecuadas síntesis con respecto a las metas que se proponen en los planes de clase.

Para ambos actores, estudiantes y docentes, es necesaria la claridad en el sistema de evaluación (George Reyes, 2021) La generación de rúbricas que permitan la orientación para la valoración de los aprendizajes obtenidos por alumnos clarifica desde el inicio del proceso los aspectos que se deben tener claros y éstos puedan identificar aquello en lo que necesiten más orientación. El reconocimiento de los contextos espacio-temporales de cada uno de los actores del proceso facilita la comunicación entre ambos, así como la planeación del tiempo para la realización de las actividades. Finalmente, el aprendizaje en lo relacionado con la parte psico-emocional ante situaciones de virtualidad o en línea, es un aspecto que, si se obvia, se puede caer en una situación de olvidar una parte importante de lo humano en la labor del docente y del estudiante.

REFERENCIAS

- Asociación Nacional de Universidades e Instituciones de Educación Superior. 2020. *Anuario estadístico de la educación superior. Ciclo escolar 2019-2020*. <http://www.anui.es.mx/informacion-y-servicios/informacion-estadistica-de-educacion-superior/anuario-estadistico-de-educacion-superior>

- Díaz-Arce, Dariel, y Efraín Loyola-Illescas. 2021. "Competencias digitales en el contexto COVID 19: una mirada desde la educación". *Revista Innova Educación* 3 (1). <https://www.revistainnovaeducacion.com/index.php/rie/article/view/181>
- George Reyes, Carlos Enrique. 2021. "Competencias digitales básicas para garantizar la continuidad académica provocada por el Covid-19". *Apertura* 13 (1): 36-51. <http://www.udgvirtual.udg.mx/apertura/index.php/apertura/article/view/1942>
- Gisbert, Mercè, Juan González, y Francesc Marc Esteve. 2016. "Competencia digital y competencia digital docente: una panorámica sobre el estado de la cuestión". *Revista Interuniversitaria de Investigación en Tecnología Educativa*, no. 0 (junio): 74-83. <https://revistas.um.es/riite/article/view/257631/195811>
- Holguín-Alvarez, Jhon, Juan Ápaza-Quispe, Jenny María Ruiz Salazar, y Juan Antonio Picoy Gonzales. 2021. "Competencias digitales en directivos y profesores en el contexto de educación remota del año 2020". *Revista Venezolana de Gerencia* 26 (94): 623-643. <https://www.produccioncientificaluz.org/index.php/rvg/article/view/35761/38066>
- Lemus Pool, María Consuelo. 2017. *El desarrollo de habilidades digitales en profesores universitarios, el caso mexicano de H@bitat Puma en la UNAM*. España: Editorial Académica Española.
- Leymonié, Julia. 2010. "Nativos e inmigrantes digitales: ¿cómo aprendemos y enseñamos?" *Revista de Comunicación Dixit* 12. <https://revistas.ucu.edu.uy/index.php/revistadixit/article/view/292>
- Parra Rodríguez, Jaime. 2005. *Tendencias de estudio en cognición, creatividad y aprendizaje*. Colombia: Pontificia Universidad Javeriana.
- Prensky, M. "Nativos e inmigrantes digitales". [https://marcprensky.com/writing/Prensky-NATIVOS%20E%20INMIGRANTES%20DIGITALES%20\(SEK\).pdf](https://marcprensky.com/writing/Prensky-NATIVOS%20E%20INMIGRANTES%20DIGITALES%20(SEK).pdf)
- Sierralta Pinedo, Sheila. 2021. "Digital skills in times of COVID-19, challenge for the teachers of the Educational Institution CECAT 'Marcial Acharán'". *Mendive: Revista de Educación* 19, 3 (July-September). <https://mendive.upr.edu.cu/index.php/MendiveUPR/article/view/2569>
- Universidad Autónoma de San Luis Potosí, Secretaría Académica. 2016. *Modelo Educativo*. <http://www.uaslp.mx/Secretaria-Academica/Paginas/Modelo-Educativo.aspx>
- World Business Council for Sustainable Development. 2012. *Information and communication technology, an enabler for inclusive business solutions*. Switzerland: WBCSD. Org.

Para citar este texto:

- Oliva-Cruz, Eduardo, y Adriana Mata-Puente. 2022. "Uso de las habilidades digitales en el proceso de enseñanza-aprendizaje en ciencias de la información en un entorno virtual durante la pandemia por Covid 19". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 177-193. <http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58627>

La producción científica de la Universidad Autónoma de Querétaro indizada en WoS, 2010-2020

Ramsés Jabín Oviedo Pérez*
Ángel Bravo Vinaja**

Artículo recibido:
15 de junio de 2022

Artículo aceptado:
29 de septiembre de 2022

Artículo de investigación

RESUMEN

El presente artículo caracteriza la producción científica de los investigadores de la Universidad Autónoma de Querétaro (UAQ) indizada en las bases de datos Science Citation Index Expanded, Emerging Sources Citation Index, Social Science Citation Index y Arts & Humanities Citation Index de la Web of Science, entre los años 2010 y 2020. El análisis tuvo el objetivo de medir la producción de artículos a partir de indicadores bibliométricos unidimensionales y multidimensionales. El trabajo recuperó un total de 2 343 publicaciones distribuidas en 199 revistas. Los resultados muestran la evolución de la producción, áreas de investigación,

* Facultad de Ciencias de Información, Universidad Autónoma de San Luis Potosí, S.L.P., México a332982@alumnos.uaslp.mx

** Programa Innovación en el Manejo de Recursos Naturales, Colegio de Postgraduados. Campus San Luis Potosí, S.L.P., México abravo@colpos.mx

revistas, categorías, idioma, autores, colaboración nacional e internacional, y ordenamiento de trabajos más citados. Se concluye que las áreas más productoras de documentos científicos de la UAQ son ingeniería, tecnología, química y física, firmados en redes de coautoría nacional e internacional.

Palabras clave: Bibliometría; Producción científica; Evaluación científica; Universidad Autónoma de Querétaro

The scientific production of the Autonomous University of Querétaro indexed in WOS, 2010-2020

Ramsés Jabín Oviedo Pérez and Ángel Bravo Vinaja

ABSTRACT

This article has as objective, to characterize the scientific production of researchers from the Autonomous University of Querétaro (UAQ) indexed in the databases as Science Citation Index Expanded, Emerging Sources Citation Index, Social Science Citation Index and Arts & Humanities Citation Index of the Web of Science, among the years 2010 and 2020. The analysis is to measure the production of articles based on unidimensional and multidimensional bibliometric indicators. The work retrieved a total of 2,343 publications distributed in 199 journals. The results show the evolution of production, research areas, journals, categories, language, authors, national and international collaboration, and ranking of the most cited works. It is concluded that the areas that produce the most scientific documents at the UAQ are engineering, technology, chemistry and physics, signed in coauthorship in national and international networking.

Keywords: Bibliometrics; Scientific production; Scientific evaluation; Autonomous University of Querétaro (UAQ)

INTRODUCCIÓN

El papel relevante que tienen las universidades en la búsqueda de una sociedad más justa y preparada surge del vínculo entre el conocimiento técnico y científico, el histórico-político, y los recursos epistemológicos, teóricos y críticos de una educación humanista (González Casanova, 2014). La docencia, la extensión y la investigación son actividades esenciales en el funcionamiento orgánico de las instituciones de educación superior (IES). Una parte medular en la consecución de los objetivos de la labor investigativa requiere de la comunicación académica y científica, lo que ha llevado a valorar las publicaciones emitidas por expertos como medios idóneos para el desarrollo de la ciencia. Las revistas científicas forman parte del engranaje general de la universidad. Sin libros, sin revistas, una universidad no es tal; las bibliotecas y los proyectos editoriales son la base de su desarrollo intelectual, científico y tecnológico.

En este contexto, la investigación universitaria tiene la responsabilidad de crear nuevos conocimientos. Sin embargo, aun cuando la actividad científica es un fenómeno multidimensional, hay una dimensión basada en la política científica nacional, que ha hecho a las universidades priorizar la difusión de resultados de investigación en revistas de alto impacto (Lloyd et al., 2012), y esto quiere decir revistas indexadas en bases de datos internacionales evaluadoras de la producción científica (Scopus, Web of Science). Por ello, en las últimas décadas la determinación de recursos públicos instauró un sistema de evaluación de la actividad científica asentado en indicadores de producción e impacto científico (Conacyt, 2021).

Al respecto, los indicadores bibliométricos son una adecuada vía para caracterizar la actividad de investigación de las instituciones de educación superior en términos cuantitativos. La bibliometría es un área estratégica para la medición y evaluación de la producción que genera un campo científico (Van Rann, 1996). Si bien la bibliometría no muestra el estado o la evolución de un fenómeno en su totalidad, sí refleja algunos de sus aspectos y señala regularidades en su comportamiento (Von Ungern-Sternberg, 2000; Guadarrama y Manzano, 2016), por lo que resulta útil para estudiar las tendencias de la investigación universitaria.

Sin embargo, hasta el momento en que se planteó la realización de la presente investigación, se pudo constatar que no se han dado estudios bibliométricos suficientes sobre la investigación científica y tecnológica de la Universidad Autónoma de Querétaro (UAQ). Un primer estudio de Campillo Sanabria (1993) determinó la dinámica y estructura de la investigación científica a partir de la utilización de herramientas de tipo cuantitativo. El segundo estudio

pertenece a Andrade Cázares, Ibarra Rivas y González Ramírez (2012), quienes realizaron un estudio bibliométrico centrado en las tesis de Maestría en Ciencias de la Educación de la UAQ producidas en el período 2002-2012, tomando como marco teórico categorías de índole sociológica.

Derivado de lo anterior, el problema que se plantea es el de analizar en qué grado o medida ha evolucionado la producción científica representada en los artículos científicos publicados por investigadores de la UAQ en las bases de datos Science Citation Index Expanded (SCIE), Emerging Sources Citation Index (ESCI), Social Science Citation Index (SSCI) y Arts & Humanities Citation Index (A&HCI) de Web of Science (WoS) en el período 2010-2020.

Así, el objetivo general de la investigación fue medir el desarrollo y evolución de la producción e impacto de la literatura científica publicada por parte de la comunidad investigadora de la Universidad Autónoma de Querétaro de 2010 a diciembre de 2020, registrada en las revistas indizadas en las bases de datos referidas de WoS de Clarivate Analytics, de acuerdo con dos tipos de indicadores bibliométricos (Sanz Casado y Martín Moreno, 1997): unidimensionales, dedicados a analizar o medir una sola característica de los documentos publicados, como los de tipología documental, temática de investigación, año de publicación, autor, idioma, dependencias, colaboración científica, ordenamiento y citas a nivel de revistas; y multidimensionales, basados en técnicas de análisis multivariante, dedicados a medir o analizar de forma simultánea diferentes variables o múltiples interrelaciones en los documentos, como los de palabras conjuntas, citas conjuntas (Small, 1973) y redes de coautoría.

METODOLOGÍA

Con la finalidad de cumplir con el objetivo establecido, se desarrollaron una serie de acciones en la WoS, colección principal, sobre la producción científica indexada en las bases de datos SCIE, ESCI, SSCI y A&HCI. La estrategia de búsqueda utilizó los filtros “Universidad Autónoma de Querétaro” en el campo afiliación y “2010-2020” en el campo “Año de publicación”. Posteriormente, se refinó la búsqueda a “Artículos” y se obtuvieron 2 346 publicaciones científicas.

Para el análisis bibliométrico, los resultados de la búsqueda se marcaron y se exportaron en formato Excel. Una vez importados los datos, se normalizaron los campos considerados en el estudio: autores, publicaciones, revistas, año, categoría WoS, instituciones y artículos más citados. La metodología descrita se aplicó en un momento, el 31 de enero de 2022, obteniéndose después de la limpieza de datos un resultado de 2 343 registros.

Para los indicadores multidimensionales, se empleó el *software* VOS-viewer (Van Eck y Waltman, 2010). A partir de la recogida de documentos recuperados de WoS, éstos se exportaron en formato de texto plano a la herramienta bibliométrica, donde se obtuvieron los mapas bibliométricos de palabras conjuntas, citas conjuntas y redes de coautoría considerados con el *corpus* analizado. El programa utiliza el Visualization of Similarities (VOS), un método propuesto por Van Eck y Waltman (2007) como alternativa al escalado multidimensional para visualizar similitudes entre objetos. La técnica VOS permite ejecutar diferentes algoritmos de *clustering* para generar mapas equiparables a grupos temáticos.

ANÁLISIS Y PRESENTACIÓN DE RESULTADOS

Evolución de la actividad científica

La evolución anual de la producción científica de la UAQ en todas las áreas de conocimiento, indizada en el WoS en el período 2010-2020, se muestra en la *Figura 1*. La producción se incrementó en un 209%, pasando de 120 artículos en el 2010 a 371 en el 2020.

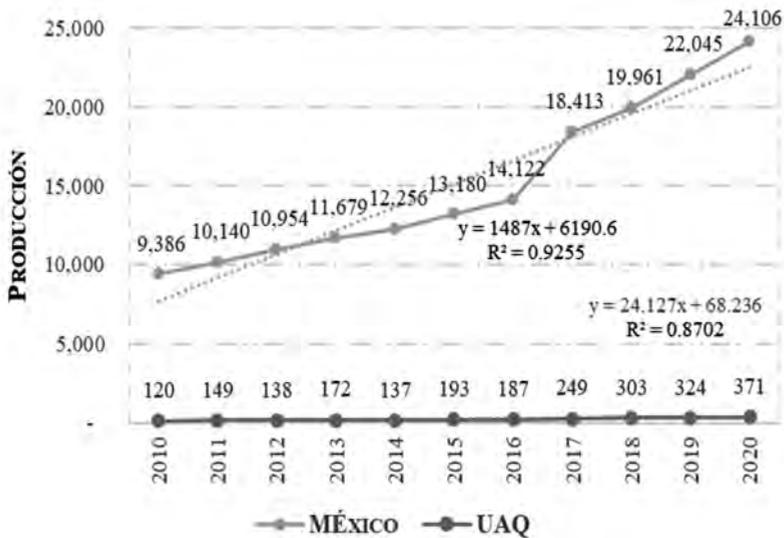


Figura 1. Comparación de la dinámica de producción científica de la UAQ y la ciencia mexicana en WoS: 2010-2020

Fuente: elaboración propia a partir de WoS

En general la producción científica publicada por la ciencia mexicana tuvo un crecimiento exponencial, siendo el año de 2013 en el que aumentó el número de publicaciones. Así, el 1.40% de la producción de la literatura nacional en el período 2010-2020 corresponde a la UAQ. En promedio, la ciencia en México publicó una media de 15 113 trabajos por año, mientras que la UAQ obtuvo una media anual de 213 artículos indexados. Durante los 10 años estudiados, las tendencias de crecimiento lineal son para México de $R^2=0.9255$ y de $R^2=0.8702$ para la UAQ; lo que implica que en esa década la universidad queretana tuvo una tendencia más baja que la producción a nivel nacional.

Cabe mencionar, por otro lado, que la UAQ tuvo en 2010 y en 2014 los años de menor producción: apenas alcanza el 10.97% del período analizado (con 257 publicaciones en dos años). El año que muestra una tendencia hacia mayor producción se da a partir de 2017, como se ve reflejado en los siguientes números: 2017 (249 trabajos), 2018 (303 trabajos), 2019 (324 trabajos) y 2020 (371 trabajos).

Evolución de la actividad científica por materias

La distribución anual de los trabajos de acuerdo con los índices SCIE, ESCI, SSCI y A&HCI de WoS se puede visualizar en la *Figura 2*. El análisis advirtió que hay crecimiento y decrecimiento de las publicaciones de la UAQ en los índices estudiados del WoS en el período 2010-2020. Sin embargo, es en el SCIE donde está indizada la mayor parte de su producción científica, mientras que la menor cantidad de publicaciones capturadas están en el A&HCI. En general, hay un patrón de publicación descrito por estudios previos (Lloyd et al., 2012): tanto Scopus como WoS se centran principalmente en las revistas especializadas en ciencias duras. Así, la presencia de las áreas de ciencias sociales, del comportamiento y de humanidades tienen escasa visibilidad en revistas de corriente principal (Tarango y Ruiz-Domínguez, 2012).

El aumento en la producción científica puede tener como causa principal el hecho de que, desde la década del 2010, la universidad queretana ha fomentado políticas de investigación, con especial atención a los indicadores de productos académicos, maximizando el uso de recursos en la investigación. Así, Baltazar Vargas (2021) considera que el elemento de la medición de la producción científica en la UAQ se desarrolla como parte de la coyuntura evaluativa centrada en la planeación y control. A su vez, García Gasca (2019: 33) sostiene que, a partir de datos obtenidos de Scopus en el período 2012-2018, el aumento en el número de publicaciones “muestra un avance en la capacidad y productividad académica de los profesores de la UAQ”.

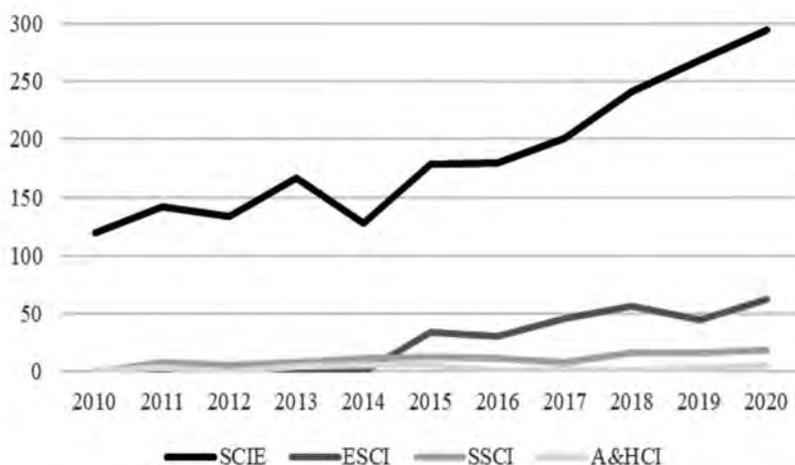


Figura 2. Evolución de la producción científica de la UAQ indizada en las bases de datos SCIE, ESCI, SSCI y A&HCI: 2010-2020
Fuente: WoS

Categorías temáticas

Con relación a las categorías temáticas y áreas de investigación de la producción científica, las 2 343 publicaciones científicas se publicaron en 199 revistas indizadas en la WoS. La *Tabla 1* muestra las 10 principales categorías temáticas en las que publica la UAQ.

Orden	Categoría	No. de trabajos	Porcentaje (%)
1	Food Science Technology	342	14.59
2	Materials Science Multidisciplinary	200	8.53
3	Engineering Electrical Electronic	174	7.42
4	Biochemistry Molecular Biology	150	6.40
5	Nutrition Dietetics	136	5.80
6	Physics Applied	114	4.86
7	Chemistry Applied	105	4.48
8	Engineering Multidisciplinary	105	4.48
9	Plant Sciences	105	4.48
10	Chemistry Multidisciplinary	104	4.43
Total	NA	1 535	65.51

Tabla 1. Principales categorías temáticas del WoS por el número de trabajos publicados: 2010-2020
Fuente: WoS. NA: No aplica

Temas de investigación

Considerando que el análisis de co-palabras es uno de los métodos más eficaces para descubrir las tendencias de investigación en un campo científico (Gálvez, 2018), la *Figura 3* muestra un mapa con las co-ocurrencias de 272 palabras de mayor frecuencia (aparecidas al menos 20 veces) tomadas del título y el resumen de los artículos de la universidad queretana. El mapa bibliométrico, al 60% de términos más relevantes, revela que de 2010 a 2020 la UAQ ha configurado cuatro grandes grupos temáticos. Mediante una escala cromática de grises se exhiben las principales regiones (clústeres) de investigación.

El clúster 1 “Ingeniería y tecnología” (derecha arriba) agrupó un total de 93 ítems, las palabras clave con mayor peso fueron: *technique; paper; approach; methodology; performance; information; signal; problem; detection*. El clúster 2 “Ciencia y tecnología de los alimentos” (izquierda abajo) incluyó un total de 68 ítems, entre las palabras clave de mayor peso se situaron: *activity; concentration; protein; composition; product; acid; content; compound; property*. El clúster 3, “Ciencias agrícolas y medicina” (derecha abajo), agrupó un total de 57 ítems, las palabras clave con mayor peso: *Mexico; species; year; population; disease; Querétaro; diversity; risk; age*. El clúster 4, “Ciencias físicas y químicas” (izquierda arriba), incluyó un total de 54 ítems; las palabras clave mayor peso fueron: *temperature; cell; degrees C; characterization; solution; phase; formation; surface; x ray diffraction*.

El mapa permite identificar hacia dónde están orientados los temas de investigación. Es el caso de la mecatrónica, una rama de la ingeniería que une conocimientos de mecánica, electrónica, informática y sistemas de control; es un área que involucra el desarrollo de dispositivos robóticos para las industrias automotriz, de la salud y de manufactura. También en donde hay más investigación es en los temas de biotecnología y biosistemas, que se sustentan en conocimientos de biología, química y otras disciplinas afines, y que utilizan sistemas biológicos y organismos vivos o sus derivados para crear o modificar productos o procesos para beneficios humanos.

Idiomas de publicación

Con respecto al idioma preferente de publicación, el 86.72% fue el inglés, el 13.23% fue el español e, igualmente, sólo 0.04% el alemán (un trabajo) y 0.04% el francés (un trabajo). Este hecho confirma el inglés como el idioma preferencial para la difusión del conocimiento en bases de datos internacionales. Cabe destacar que existe un incremento significativo de artículos publicados en castellano en la WoS, puesto que en 2010 sólo se registraron 9

Título de la publicación	Acceso abierto	Base de datos e impacto (2020)	Categoría JCR, Posición y Cuartil	No. de artículos	%
Sensors	Sí	SCIE, FI 3.031	Instruments & Instrumentation, 15/61 (C1)	35	1.48
Food Research International	Sí	SCIE, FI 6.475	Food Science & Technology, 9/143 (C1)	28	1.18
Food Chemistry	No	SCIE, FI 7.514	Chemistry, Applied, 7/74 (C1)	27	1.14
Tecnología y Ciencias del Agua	Sí	SCIE, FI 0.367	Engineering, Civil, 135/137 (C4)	24	1.01
Applied Sciences-Basel	Sí	SCIE, FI 2.679	Chemistry, Multidisciplinary, 101/178 (C3)	22	0.93
International Journal of Hydrogen Energy	No	SCIE, FI 5.816	Chemistry, Physical, 48/162 (C2)	22	0.93
IEEE Access	Sí	SCIE, FI 3.367	Engineering, Electrical & Electronic, 94/273 (C2)	20	0.84
Plos One	Sí	SCIE, FI 3.24	Multidisciplinary Sciences, 26/72, (C2)	20	0.84
Total	NA	NA	NA	198	8.34

Tabla 2. Revistas con 20 o más artículos de la UAQ publicados en WoS: 2010-2020
Fuente: WoS. NA: No aplica.

Co-citación de las revistas

Debido a que la red de relaciones entre revistas son también una estructura institucional de la actividad científica, las citas entre revistas pueden utilizarse como indicador de la organización disciplinaria de las ciencias (Ding, Chowdhury y Foo, 2000) y, por lo tanto, de las áreas de investigación científica de la UAQ. En este sentido, la *Figura 4* visualiza la conformación de redes académicas entre revistas citadas en la producción científica de la UAQ.

Así, el mapa permite leer la distribución de 746 revistas con al menos 20 citas. Las publicaciones más altamente citadas (clústeres más grandes) son las revistas que se citan en los artículos que publica la UAQ, es decir, las fuentes que

Por otra parte, la *Tabla 3* también muestra el índice H de los autores. Se trata un índice bibliométrico que sirve para medir la relación entre la cantidad de publicaciones y las citas recibidas (Hirsch, 2005). Se agrega como sub indicador ya que es útil para identificar a los autores más destacados en cada disciplina. Tiene la conveniencia de combinar productividad con impacto. De esa manera, se puede observar que los artículos firmados por Flavia Loarca y Rosado Loria, procedentes de disciplinas científicas congéneres, han obtenido más citas en WOS durante los años estudiados.

Orden	Investigador	Facultad de adscripción	Área de especialización	No. de trabajos	Índice H*
1	Roque Alfredo Osornio Ríos	Facultad de Ingeniería, Campus San Juan de Río	Mecatrónica	110	27
2	Janet Ledesma García	Facultad de Ingeniería, Campus Cerro de las Campanas	Materiales nanoestructurados	90	27
3	Rosalía Reynoso Camacho	Facultad de Química, Campus Cerro de las Campanas	Nutrición	63	25
4	René de Jesús Romero Troncoso	Facultad de Ingeniería, Campus San Juan de Río	Mecatrónica	61	27
5	Guadalupe Flavia Loarca Piña	Facultad de Química, Campus Cerro de las Campanas	Compuestos bioactivos	60	32
6	Ramón Gerardo Guevara González	Facultad de Ingeniería, Campus Amazcala	Biosistemas	59	27
7	Minerva Guerra Balcázar	Facultad de Ingeniería, Campus Cerro de las Campanas	Nanomateriales	57	20
8	Juvenal Rodríguez Reséndiz	Facultad de Ingeniería, Campus Cerro de las Campanas	Mecatrónica	54	13

9	Irineo Torres Pacheco	Facultad de Ingeniería, Campus Cerro de las Campanas	Biología	44	23
9	Jorge Luis Rosado Loria	Facultad de Ciencias Naturales, Campus Juriquilla	Nutrición	44	31
10	Francisco De Moure Flores	Facultad de Química, Campus Cerro de las Campanas	Caracterización de materiales	39	13

Tabla 3. Investigadores de la UAQ con más publicaciones en WoS: 2010-2020

Fuente: Elaboración propia con base en datos de WoS

*Índice H obtenido de WoS

Redes de coautoría

Por el análisis de la producción científica, se pudo apreciar que existen diferentes redes de colaboración y coautoría interna en la UAQ. Así, en la *Figura 5* se visualizan distintos clústeres de investigadores de la IES queretana con al menos 15 documentos indexados en WoS. Se pudo constatar que las adscripciones académicas de los autores con más publicaciones conforman las facultades dedicadas a las ciencias duras y naturales, cuyo resultado está relacionado con las tendencias de investigación que muestra esta universidad.

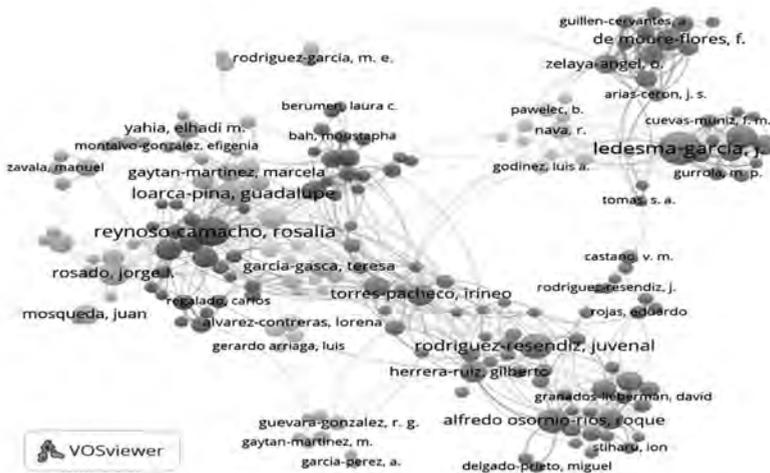


Figura 5. Redes de coautoría de la UAQ en WoS: 2010-2020

Fuente: Elaboración propia con base en datos de WoS

Colaboración entre países

En el análisis de la investigación colaborativa en las publicaciones de la UAQ, se encontró que existen ocurrencias por coautoría de múltiples investigadores a nivel nacional e internacional en el período 2010-2020. Destaca la participación de investigadores provenientes de diversos países en trabajos conjuntos con la UAQ. Así, en el análisis de las publicaciones, la universidad queretana mantiene una fuerte colaboración con instituciones de Estados Unidos de América (10.13%), seguido por España, Canadá, Inglaterra, Italia, Colombia y Francia (*Tabla 4*). Se observa que hay, en menor grado, una colaboración con países latinoamericanos. El 40.90% de la producción total se realizó en colaboración internacional. En el período estudiado, la UAQ publicó 2 343 trabajos en coautoría con 66 países del mundo. La Universidad de Illinois es la institución de EUA con más colaboración en la década estudiada. Entre otros datos no incluidos en la Tabla hay 4 países con un rango de veinte a once colaboraciones, 18 entre diez y cuatro, 2 con tres, 11 con dos y 19 países con sólo una. Este hallazgo confirma el estudio de Ríos Gómez y Herrero Solana (2005), en la medida en que los países con los que más colabora América Latina son Estados Unidos, España y Reino Unido.

País	No. de trabajos	Porcentaje %
Estados Unidos	239	10.13
España	197	8.35
Canadá	70	2.96
Reino Unido	41	1.73
Italia	35	1.48
Colombia	31	1.31
Francia	31	1.31
Argentina	25	1.05
Alemania	22	0.93
Australia	21	0.89
Chile	21	0.89
Otros 55 países	>20	68.92

Tabla 4. Colaboración internacional de la UAQ en WoS: 2010-2020
Fuente: Elaboración propia con base en datos de WoS

Colaboración con instituciones nacionales

A nivel nacional, en la *Figura 6* se puede ver la distribución de la producción científica de la UAQ con instituciones con sede en los estados de la República Mexicana, representado por una escala de grises. Se obtuvo una colaboración con 94 instituciones, la mayoría concentradas en la Ciudad de México. Un aspecto importante a considerar es que a nivel regional la UAQ colabora asiduamente con la UNAM, pero también juegan un papel relevante los centros de investigación regionales (CIR) y los centros públicos de investigación (CPI). Además de la capital de México, se alcanza a visualizar que la UAQ tiende a publicar investigaciones en colaboración con instituciones de Guanajuato, Chihuahua, Jalisco, Nuevo León, San Luis Potosí y Michoacán; y, de acuerdo con los datos reportados, los estados con los que no tiene ninguna colaboración son Campeche y Tabasco. Ambas situaciones se pueden analizar desde dos aristas. La primera, positiva, es que la UAQ mantiene redes de coautoría con la principal entidad educativa de investigación científica en México, la UNAM, lo cual redundaría en la conformación de comunidades de investigación de calidad. La segunda, negativa, es que la UAQ aún no ha establecido vínculos de colaboración científica con todos los estados de la República, a pesar de que en los últimos años en México “se ha enriquecido la estructura organizativa de las instituciones: se han incrementado los programas de investigación, los perfiles de formación, los enfoques regionales de la investigación y se han multiplicado las relaciones de colaboración entre científicos e instituciones” (Flores Vargas, 2021: 42). Estas circunstancias podrían ayudar a construir formas de organización más productivas y a incrementar la producción de conocimientos científicos, tecnológicos y humanísticos realizada en la UAQ.

Influencia de la investigación

Para poder conocer la influencia de la producción científica, se seleccionaron los artículos con mayor cantidad de citas recibidas. En la *Tabla 5* se muestran las publicaciones con más citas reportadas en la WoS en el período 2010-2020. Conforme al ordenamiento, se puede apreciar que los trabajos corresponden a las áreas de conocimiento de la ingeniería, la química y la biología; publicados entre el 2011 y el 2015 en revistas con factores de impacto altos, por redes internacionales de coautoría. Asimismo, llama la atención que sólo un artículo alcanzó más de 200 citas, recibidas en el rango temporal estudiado.



Figura 6. Distribución de la producción de la UAQ a nivel nacional en WoS: 2010-2020
Fuente: Elaboración propia con base en datos de W

Orden	Título del artículo	No. Citas
1	Vanthoor-Koopmans, M., Wijffels, R., Barbosa, M., y Eppink, M. 2013. "Biorefinery of microalgae for food and fuel". <i>Bioresource Technology</i> 135: 142-149. https://doi.org/10.1016/j.biortech.2012.10.135	248
2	Hernández-Ledesma, P., Berendsohn, W., Borsch, T., Von Mering, S., Akhani, H., Arias, S., Castaneda-Noa, I., Eggli, U., Eriksson, R., Flores-Olvera, H., Fuentes-Bazán, S., Kadereit, G., Klak, C., Korotkova, N., Nyffeler, R., Ocampo, G., Ochoterena, H., Oxelman, B., Rabeler, R., y Uotila, P. 2015. "A taxonomic backbone for the global synthesis of species diversity in the angiosperm order Caryophyllales". <i>Willdenowia</i> 45 (3): 281-383. https://doi.org/10.3372/wi.45.45301	196
3	Valtierra-Rodríguez, M., Romero-Troncoso, R., Osornio-Ríos, R., y García-Pérez, A. 2014. "Detection and Classification of Single and Combined Power Quality Disturbances Using Neural Networks". <i>IEEE Transactions on Industrial Electronics</i> 61 (5): 2473-2482. https://doi.org/10.1109/TIE.2013.2272276	186

4	Alcántara, C., Kuemmerle, T., Prishchepov, A., y Radeloff, V. 2012. "Mapping abandoned agriculture with multi-temporal MODIS satellite data". <i>Remote Sensing of Environment</i> 124: 334-347. https://doi.org/10.1016/j.rse.2012.05.019	177
5	García-Pérez, A., Romero-Troncoso, R., Cabal-Yepez, E., y Osornio-Ríos, R. 2011. "The Application of High-Resolution Spectral Analysis for Identifying Multiple Combined Faults in Induction Motors". <i>IEEE Transactions on Industrial Electronics</i> 58 (5): 2002-2010. https://doi.org/10.1109/TIE.2010.2051398	155
6	Hernández-Morales, V., Nava, R., Acosta-Silva, Y., Macías-Sánchez, S., Pérez-Bueno, J., y Pawelec, B. 2012. "Adsorption of lead (II) on SBA-15 mesoporous molecular sieve functionalized with -NH ₂ groups". <i>Microporous and Mesoporous Materials</i> 160: 133-142. https://doi.org/10.1016/j.micromeso.2012.05.004	141

Tabla 5. Publicaciones científicas de la UAQ con más citas en WoS: 2010-2020

Fuente: WoS

DISCUSIÓN

Aunque se han publicado otros estudios bibliométricos sobre la producción científica de las universidades de educación superior en México (véase Villaseñor, Arencibia-Jorge y Carrillo-Calvet, 2017), la novedad de la presente investigación residió en que se realizó desde un esquema micro con una sola institución. Para este propósito, se analizó un *corpus* de documentos de la UAQ. En la interpretación de resultados se identificó que la producción de artículos de investigadores mexicanos en revistas indexadas en WOS muestra un crecimiento sostenido, lo que coincide con investigaciones previas como la de Lloyd *et al.* (2012), en la que se aprecia la contribución de la UAQ y otras IES públicas y privadas. De acuerdo con el diagnóstico del Foro Consultivo Científico y Tecnológico, Querétaro tiene como áreas estratégicas para la competitividad, entre otras, la nanotecnología, la biotecnología y la electrónica (FCCYT, 2012), lo cual tiene relación con las temáticas de investigación más destacadas de la UAQ.

Por otro lado, en la interpretación de las palabras clave posicionadas en los conglomerados del mapa de co-palabras, las áreas de investigación de la UAQ se vinculan al conjunto de términos/conceptos científicos, desarrollados en los campos de conocimiento ingenieril y químico, asociados a grupos

temáticos como la mecatrónica, la biotecnología, la química y la física aplicada, como también se evidenció con la enorme producción de redes de expertos en el ámbito de las ciencias duras y exactas. El indicador de redes de coautoría mostró la imagen de un fuerte comportamiento colaborativo y su nivel de influencia entre investigadores universitarios. Típicamente, la visibilidad de las ciencias sociales y humanidades tienen una escasa representación en bases de datos de WoS.

Otra consideración es que el proceso de modernización educativa de la UAQ, adoptada desde finales del siglo XX, involucró en gran medida cambios en la dinámica de la publicación científica, según recalca el estudio de Baltasar Vargas (2021). En este sentido, a nivel geográfico esta IES ha aportado una productividad que impacta en las capacidades científicas de México (Flores Vargas, 2021). El estudio comprobó el interés de la UAQ por entrar al sistema de evaluación instaurado por la política científica nacional, y también el desinterés por registrar la información científica en acceso abierto, en tanto que, a pesar de que se sostiene con fondos públicos que deberían comprometer la difusión y el libre acceso a sus resultados de investigación, el estudio determinó que más de la mitad de los artículos producidos por la UAQ están en revistas de alto impacto, y el 61.98% se concentra en publicaciones de acceso restringido.

CONCLUSIONES Y RECOMENDACIONES

Con la aplicación de los indicadores bibliométricos se logró mostrar el crecimiento exponencial de la producción científica de la Universidad Autónoma de Querétaro indexada en WOS, durante el período 2010-2020. Las principales conclusiones que se extrajeron se exponen a continuación. Con relación a las áreas de investigación de las publicaciones, ésta mostró que la mayor producción de la Universidad se concentra en las ciencias duras y exactas. El ordenamiento de las categorías WoS comprueba una vez más que las áreas de Ciencias Sociales y Humanidades destacan por su ausencia en la distribución temática de la productividad universitaria. La preferencia idiomática sostenida por la mayoría de los trabajos indica que el inglés es el requisito más aceptado para la comunicación de resultados en bases de datos internacionales. Asimismo, se constató una alta concentración de publicaciones acaparada por investigadores de las facultades de Ingeniería, Química y Ciencias Naturales. También se mostró que las áreas más productoras de investigación de la UAQ son los estudios en ingeniería, tecnología, química y física, firmados en redes de coautoría nacional e internacional.

Finalmente, es importante señalar que incluir sólo WoS limita el análisis del perfil bibliométrico de la UAQ. Por lo que, a pesar de esta limitante, sería útil tener en cuenta el número total de sus artículos publicados en diferentes bases de datos: locales, nacionales e internacionales, para obtener una caracterización bibliométrica más completa. En consecuencia, se recomienda para futuros estudios contemplar todo el *corpus* de su producción científica, para lo cual sería importante que esta IES desarrolle y difunda una base de datos adecuada para tal faena. Atender estas recomendaciones permitirá a la UAQ dar cuenta a la sociedad de sus investigaciones científicas, tecnológicas y humanísticas.

REFERENCIAS

- Andrade Cázares, Rocío A., Luis Rodolfo Ibarra Rivas, y Sara Miriam González Ramírez. 2015. "Procesos formativos en investigación en los estudiantes de la Maestría en Ciencias de la Educación de la Universidad Autónoma de Querétaro". *Revista de Investigación Educativa de la Escuela de Graduados en Educación* 5 (10): 8-16.
- Baltasar Vargas, David. 2021. "Reforma y cambio organizacional en la Universidad Autónoma de Querétaro (1982-2015)". Tesis de doctorado, Universidad Autónoma Metropolitana-Unidad Xochimilco, División de Ciencias Sociales y Humanidades.
- Campillo Sanabria, Carlos. 1993. *Historia de la investigación en la U.A.Q. 1981-1992*. Querétaro: Universidad Autónoma de Querétaro.
- Conacyt. 2021. *Informe general del estado de la Ciencia, la Tecnología y la Innovación 2019*. Ciudad de México: Gobierno de México.
- Ding, Ying, Gobinda G. Chowdhury, y Schubert Foo. 2000. "Journal as markers of intellectual space: journal co-citation analysis of information retrieval area, 1987-1997". *Scientometrics* 47 (1): 55-73.
<https://doi.org/10.1023/a:1005665709109>
- FCCYT A. C. 2012. *Querétaro. Diagnóstico en Ciencia, Tecnología e Innovación 2004-2011*. México: Foro Consultivo Científico y Tecnológico A. C.
- Flores Vargas, Xóchitl. 2021. "Conformación geográfica de las capacidades científicas en México: Un análisis de indicadores cientiométricos: 1995-2015". Tesis de doctorado, Instituto Politécnico Nacional-Unidad Zacatenco, Centro de Investigación y de Estudios Avanzados.
- Gálvez, Carmen. 2018. "Análisis de co-palabras aplicado a los artículos muy citados en Biblioteconomía y Ciencias de la Información (2007-2017)". *Transinformação* 30 (3): 277-286.
<https://doi.org/10.1590/2318-08892018000300001>
- García Gasca, Margarita. 2019. *Primer Informe. Rectoría 2018-2021*. Querétaro: Universidad Autónoma de Querétaro.

- González Casanova, Pablo. 2014. *La universidad necesaria en el siglo XXI*. México: Era.
- Guadarrama, Víctor, y Francisco Manzano. 2016. *Indicadores de Ciencia, Tecnología e Innovación*. México: Foro Consultivo Científico y Tecnológico, A.C.
- Hirsch, Jorge E. 2005. "An index to quantify an individual's scientific research output". *Proceedings of the National Academy of Sciences of the United States of America* 102 (46): 16569-16572.
<https://doi.org/10.1073/pnas.0507655102>
- Lloyd, Marion Whitney, Alejandro Márquez Jiménez, Imanol Ordorika Sacristán, Roberto Rodríguez Gómez, Francisco Javier Lozano Espinoza, Jorge Martínez Stack, Martha Montes de Oca Cáliz, Pilar López Martínez, y Gabriela Olguín Carro. 2012. *Estudio Comparativo de Universidades Mexicanas. Tercer Reporte: Desempeño en Investigación y Docencia. Datos Básicos 2009*. México D.F.: DGEI-Universidad Nacional Autónoma de México.
<https://doi.org/10.13140/2.1.1256.2722>
- Ríos Gómez, Claudia, y Víctor Herrero Solana. 2005. "La producción científica latinoamericana y la ciencia mundial: una revisión bibliográfica (1989-2003)". *Revista Interamericana de Bibliotecología* 28 (1): 43-61.
- Sanz Casado, Elías, y Carmen Martín Moreno. (1997). "Técnicas bibliométricas aplicadas a los estudios de usuarios". *Revista General de Información y Documentación* 7 (2): 41-68.
- Small, Henry G. (1973). "Co-citation in the science literature: A new measure of the relationship between two documents". *Journal of the American Society for Information Science* 24 (4): 265-269.
- Tarango, Javier, y Arturo Ruiz-Domínguez. 2012. "Revistas periféricas de generación interna en la Universidad Autónoma de Chihuahua, México: identificación de indicadores para la calidad". *Revista General de Información y Documentación* 22: 367-394.
http://doi.org/10.5209/rev_RGID.2012.v22.39661
- Van Eck, Nes Jan, y Ludo Waltman. 2007. "VOS: A New Method for Visualizing Similarities Between Objects". En *Advances in Data Analysis*, editado por Reinhold Decker, y Hans-J. Lenz, 299-306. Berlin: Springer.
https://doi.org/10.1007/978-3-540-70981-7_34
- _____. 2010. "Software survey: VOSviewer, a computer program for bibliometric mapping". *Scientometrics* 84 (2): 523-538.
<https://doi.org/10.1007/s11192-009-0146-3>
- Van Raan, Anthony F. J. 1996. "Advanced bibliometric method as quantitative core of peer review based evaluation and foresight exercises". *Scientometrics* 36 (3): 397-420.
<https://link.springer.com/article/10.1007/BF02129602>
- Villaseñor, Elio Atenógenes, Ricardo Arencibia-Jorge, y Humberto Carrillo-Calve. 2017. "Multiparametric characterization of scientometrics performance profiles assisted by neural networks: a study of Mexican higher education institutions". *Scientometrics* 110: 77-104.
<https://doi.org/10.1007/s11192-016-2166-0>
- Von Ungern-Sternberg, Sara. 2000. *Scientific Communication and Bibliometrics*.
<http://www.abo.fi/~sungern/comm00.htm>

Para citar este texto:

Oviedo Pérez, Ramsés Jabín, y Ángel Bravo Vinaja. 2022. “ La producción científica de la Universidad Autónoma de Querétaro indizada en WoS, 2010-2020”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 36 (93): 195-215.

<http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58664>

DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2022.93.58664>