

OCTUBRE/DICIEMBRE,
2018, VOL. 32, NÚM. 77,
MÉXICO, ISSN 2448-8321

Investigación Bibliotecológica

ARCHIVONOMÍA, BIBLIOTECOLOGÍA E INFORMACIÓN



Investigación Bibliotecológica

ARCHIVONOMÍA, BIBLIOTECOLOGÍA E INFORMACIÓN

Vol. 32, núm. 77, octubre/diciembre, 2018, México, ISSN: 2448-8321
DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77>



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Instituto de Investigaciones Bibliotecológicas y de la Información

Contenido

INVESTIGACIÓN BIBLIOTECOLÓGICA, VOL. 32, NÚM. 77, OCTUBRE/DICIEMBRE, 2018, MÉXICO, ISSN: 2448-8321
DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77>

ARTÍCULOS

- **Impacto de las revistas españolas de Biblioteconomía y Documentación y repercusión de las autocitas en su índice h** [Impact of the Spanish Library and Information Science journals and impact of self-citation on their H-index]
José Antonio Salvador-Oliván, Gonzalo Marco-Cuenca y Rosario Arquero-Avilés 13-30
- **Origen y desarrollo de la editorial Espasa-Calpe Mexicana (1927-1975)** [Origin and history of the Espasa-Calpe Mexicana publishing house (1927-1975)]
José Miguel Sánchez-Vigil 31-57
- **O ensino pela pesquisa: interlocuções com Paulo Freire na docência em biblioteconomia** [The teaching for the research: interlocutions with Paul Freire in teaching in library]
Jorge Santa Anna 59-72
- **La opinión en textos con un enfoque interdisciplinar: propuesta de contexto y léxico ad hoc** [A context and ad hoc lexicon proposal to an interdisciplinary approach to opinion analysis]
Silvana Grazia Temesio-Vizoso 73-99
- **Gestor de contenidos de código abierto para archivos digitales sonoros que preservan materiales de investigación** [Open source content management for digital sound archives that preserve research materials]
Perla Olivia Rodríguez-Reséndiz, Joséphine Simonnot y Dafne Citlalli Abad-Martínez 101-115
- **The impact of open access citation and social media on leading top Information Science journals** [Impacto del acceso abierto en citaciones y medios sociales de las principales revistas de Ciencia de la Información]
Paulo Roberto Cintra, Ariadne Chloe Furnival y Douglas Henrique Milanez 117-132
- **Tendencias de la innovación tecnológica en Colombia 1991-2013 a partir del análisis de patentes** [Patent analysis of Colombian technological innovation 1991-2013]
Rosa Elena Gómez-Hurtado 133-150

- **Redimensionamento do espaço e do tempo: a incerteza e os desafios da sociedade, da informação e do planejamento do marketing** [Re-dimensioning space and time: The uncertainty and challenges in society, information and marketing planning]
Roberto Vilmar-Satur, Júlio Afonso Sá de Pinho-Neto e Emeide Nóbrega-Duarte 151-173
- **Representación temática y recuperación de los noticieros de televisión en la web: propuesta y aplicación de un modelo metodológico** [Subject area representation and retrieval of television news on the web: proposal and application of a methodological model]
Catalina Naumis-Peña y Silvano Soto-Hernández 175-191
- **Mujeres investigadoras en las primeras estructuras de organización en ciencias exactas e ingenierías en México de 1900-2000: Estudio Bibliométrico** [A Bibliometric study of women researchers in early organizational structures of exact sciences and engineering in Mexico from 1900 to 2000]
María Elena Luna-Morales y Evelia Luna-Morales 193-215

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. 32, No. 77, octubre-diciembre 2018, es una publicación trimestral editada por la Universidad Nacional Autónoma de México, Ciudad Universitaria, Delegación 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, Del. 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, Del. Coyoacán, Ciudad de México, C.P. 04510, Tel. 56230337, fecha de la última modificación, Octubre 2018.

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.

DOL: <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77>

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 IIBI: <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: www.elsevier.es/unam/investigacionbibliotecologica

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. ADOLFO RODRÍGUEZ GALLARDO
Universidad Nacional Autónoma de México
Ciudad de México, México

DRA. SUELI ANGÉLICA DO AMARAL
Universidad de Brasília
Brasília, Brasil

M.SC. SARAY CÓRDOBA GONZÁLEZ
Universidad de Costa Rica
San José, Costa Rica

PhD BIRGER HJØRLAND
University of Copenhagen
Copenhagen, Dinamarca

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

DRA. GLORIA PÉREZ SALMERÓN
Presidenta electa de IFLA
Barcelona, España

DRA. FERNANDA RIBEIRO
University of Porto
Porto, Portugal

DR. ELÍAS SANZ CASADO
Universidad Carlos III
Madrid, España

DR. EMILIO SETIÑO QUESADA
Biblioteca Nacional José Martí
Habana, Cuba

COLABORADORES EN ESTE NÚMERO

INVESTIGACIÓN BIBLIOTECOLÓGICA, vol. 32, núm. 77, octubre/diciembre, 2018, México, ISSN: 2448-8321

Ariadne Chloe Furnival

Departamento de Ciência da Informação - DCI
Universidade Federal de São Carlos - campus
São Carlos Rodovia Washington Luis, km 235 -
São Carlos - SP - Brasil CEP: 13565-905
Brasil
Tel./fax: (16) 3351-8374
chloe@ufscar.br

Catalina Naumis Peña

Instituto de Investigaciones Bibliotecológicas
y de la Información, Universidad Nacional
Autónoma de México
Piso 12, Torre II de Humanidades. Ciudad
Universitaria. Avda. Universidad 3000
C P 04510, Coyoacán, Ciudad de México
Tel. 56230346
naumis@unam.mx

Dafne Citlalli Abad Martínez

Instituto de Investigaciones Bibliotecológicas
y de la Información de la Universidad Nacional
Autónoma de México
Circuito Interior s/n, Torre II de Humanidades,
pisos 11, 12 y 13, Ciudad Universitaria, C.P.
04510, Del. Coyoacán, México, D.F.
Tel. 52 55 5623300
dafne@iibi.unam.mx

Douglas Henrique Milanez

NIT-Materiais, Universidade Federal de
São Carlos - campus São Carlos Rodovia
Washington Luis, km 235 - São Carlos - SP -
Brasil CEP: 13565-905
Brasil
Tel./fax: (16) 3351-8551.
dougmilanez@gmail.com

Emeide Nóbrega Duarte

Universidade Federal da Paraíba (UFPB)
Programa de Pós-Graduação em Ciência
da Informação (PPGCI) do CCSA da UFPB,
Campus I, João Pessoa/PB, CEP: 58051-900
Brasil
Tel. +55 83 3216 7200
emeide@hotmail.com

Evelia Luna-Morales

Cinvestav-IPN, Coordinación General de
Servicios Bibliográficos, Ciudad de México
Av. Instituto Politécnico Nacional 2508, Col.
San Pedro Zacatenco, Delegación Gustavo A.
Madero, México DF, Código Postal 07360,
Apartado Postal: 14-740, 07000, México DF
México
Tel. 5747-3800 (ext. 1778)
eluna@cinvestav.mx

Gonzalo Marco-Cuenca

Departamento de Ciencias de la Documenta-
ción e Historia de la Ciencia, Universidad de
Zaragoza
C/ Pedro Cerbuna, 12. 50009, Zaragoza
España
Tel. 876553559
gmarco@unizar.es

Jorge Santa Anna

Universidade Federal de Minas Gerais (UFMG)
Rua José Alves da Silva, n. 595, Bairro
Caiçaras, Belo Horizonte, Brasil,
CEP: 30775-390
Brasil
Tel. 55 (27)99267-1527
professorjorgeufes@gmail.com

Joséphine Simonnot

Centre de Recherche en Ethnomusicologie,
Paris, Francia
MAE - 21, allée de l'Université
Bat. Max Weber
92023 Nanterre cedex
Tel. 01 46 69 26 69
josimonnot@gmail.com

José Antonio Salvador-Oliván

Departamento de Ciencias de la
Documentación e Historia de la Ciencia,
Universidad de Zaragoza
C/ Pedro Cerbuna, 12. 50009, Zaragoza
España
Tel. 876553562
jaso@unizar.es

Juan Miguel Sánchez Vigil

Facultad de Ciencias de la Documentación
C/ Santísima Trinidad, 37, 28010 Madrid
España
Tel. 0034 /913946674
jmvigil@ucm.es

Júlio Afonso Sá De Pinho Neto

Universidade Federal da Paraíba (UFPB)
Programa de Pós-Graduação em Ciência
da Informação (PPGCI) do CCSA da UFPB,
Campus I, João Pessoa/PB, CEP: 58051-900
Brasil
Tel. +55 83 3216 7200
sadepinho@gmail.com

María Elena Luna-Morales

Cinvestav-IPN, Coordinación General de
Servicios Bibliográficos, Ciudad de México
Av. Instituto Politécnico Nacional 2508, Col.
San Pedro Zacatenco, Delegación Gustavo A.
Madero, México DF, Código Postal 07360,
Apartado Postal: 14-740, 07000, México DF
México
Tel. 5747-3800 (ext. 1778)
meluna@cinvestav.mx

Paulo Roberto Cintra

Universidade Estadual de Campinas –
UNICAMP – Brasil
Cidade Universitária Zeferino Vaz - Distrito
de Barão de Geraldo Caixa Postal 6152 CEP:
13083-970 Campinas. S.P. Brasil
Tel./fax: (19) 3521 4555.
paulocntr@yahoo.com

Perla Olivia Rodríguez-Reséndiz

Instituto de Investigaciones Bibliotecológicas
y de la Información, Universidad Nacional
Autónoma de México
Circuito Interior s/n, Torre II de Humanidades,
pisos 11, 12 y 13, Ciudad Universitaria, C.P.
04510, Del. Coyoacán, México,
Ciudad de México
Tel. 52 55 56230357
perlaolivia@gmail.com

Roberto Vilmar Satur

Universidade Federal da Paraíba (UFPB)
Departamento de Mediações Interculturais
(DMI) do CCHLA da UFPB, Campus I, João
Pessoa/PB, CEP: 58051-900
Brasil
Tel. +55 83 3216 7200
robertosatur@yahoo.com.br

Rosa Elena Gómez Hurtado

Facultad de Ingeniería y Nuevas Tecnologías,
Fundación Universitaria UNINPAHU
Calle 83ª No. 116ª-85 Casa 98
Colombia
Tel. 3208470840
rhgomez02@uninpahu.edu.co

Rosario Arquero-Avilés

Departamento de Biblioteconomía y
Documentación, Universidad Complutense de
Madrid
C/ Santísima Trinidad, 37. 28010, Madrid
España
Tel. 876553559
carquero@ucm.es

Silvana Grazia Temesio Vizoso

Universidad de la República, Facultad de
Información y Comunicación
San Salvador 1944 Despacho D3 , 4to. Piso.
CP 11200
Montevideo, Uruguay
Tel. +598 99 056262
silvana.temesio@fic.edu.uy

Silvano Soto Hernández

IDMS Consulting SC
Zacatecas 201. Col. Valle Ceylán. Tlalnepantla,
Estado de México. CP. 54150.
Tel. 53917914
silvano_soto@idms.com.mx

Impacto de las revistas españolas de Biblioteconomía y Documentación y repercusión de las autocitas en su índice h

José Antonio Salvador-Oliván*

Gonzalo Marco-Cuenca*

Rosario Arquero-Avilés**

Artículo recibido:

17 de octubre de 2017

Artículo aceptado:

10 de agosto de 2018

Artículo de investigación

RESUMEN

El número de citas que recibe un artículo es un indicador de su impacto en la comunidad científica. El objetivo de este trabajo es analizar el impacto de los artículos más citados que han sido publicados en las dos revistas españolas de Biblioteconomía y Documentación incluidas en el *Journal Citation Reports (JCR)*, así como determinar la influencia de las autocitas en su índice h. Los resultados muestran que los artículos son citados principalmente por autores y revistas iberoamericanas, con un impacto internacional muy bajo debido, prácticamente en su totalidad, a la presencia

* Departamento de Ciencias de la Documentación e Historia de la Ciencia
Universidad de Zaragoza, España. jaso@unizar.es
gmarco@unizar.es

** Departamento de Biblioteconomía y Documentación
Universidad Complutense de Madrid, España. carquero@ucm.es

de artículos escritos en inglés. Las revistas más citantes son las mismas para ambos casos, la revista mexicana *Investigación Bibliotecológica*, las revistas españolas *El Profesional de la Información* (EPI) y *Revista Española de Documentación Científica* (REDOC), y una revista internacional (*Scientometrics*). La tasa de autocitación conjunta de revistas y autores es muy alta (40 %); la eliminación de las autocitas de revista afecta al índice h y proporciona una medida más objetiva y real del impacto.

Palabras clave: Revistas Españolas; Biblioteca y Documentación; Web of Science; Impacto; Visibilidad; Autocitas; Índice h

Impact of the Spanish Library and Information Science journals and impact of self-citation on their H-index

José Antonio Salvador-Oliván, Gonzalo Marco-Cuenca and Rosario Arquero-Avilés

ABSTRACT

The number of citations article paper receives serves as an indicator of its impact in the scientific community. This paper performs an analysis of the impact of the most frequently cited papers published in two Spanish Library and Information Science journals included in *Journal Citation Reports* (JCR) and determines the influence of self-citations on their H index. Results show that the papers are cited largely by Latin-American authors and journals with little international impact, whose scant impact, in fact, is almost entirely due to papers written in English included in the journal. The most frequently citing journals for both papers are the Mexican journal *Investigación Bibliotecológica*, the Spanish journals, *El Profesional de la Información* (EPI), and *Revista Española de Documentación Científica* (REDOC), and the international journal *Scientometrics*. The self-citation rate of these journals and authors is very high (40%); and the elimination of self-citations impacts the H-index while yielding a more objective, realistic measures of impact.

Keywords: Spanish Journals; Library and Information Science; Web of Science; Impact; Visibility; Self-citation; LIS Spanish Journals; H-index

INTRODUCCIÓN Y OBJETIVOS

En la cultura científica actual predomina el reconocimiento profesional que supone la publicación en revistas de prestigio (Schekman, 2013). Este prestigio se basa en el impacto, medido a través del número de citas que reciben las publicaciones y/o revistas. La principal fuente de datos de la que se obtienen las citas y los indicadores derivados de ellas es la base de datos Web of Science (WoS), que contiene las revistas más importantes (*mainstream*) de las distintas áreas científicas; la publicación de artículos en estas revistas es un requisito en el procedimiento de evaluación de la investigación de profesores universitarios e investigadores. Sin embargo, se han señalado algunas limitaciones como su sesgo en la cobertura idiomática (principalmente revistas en lengua inglesa) y otras que afectan a los indicadores de impacto.

La indización de una revista en WoS se considera un símbolo de prestigio y su clasificación en cuartiles, según el número de citas recibidas, revela su importancia e impacto en la comunidad científica internacional. Habitualmente, se utilizan como indicadores de impacto el factor de impacto (FI) y el FI de 5 años para las revistas del *Journal Citation Reports* (JCR), y el *Scimago Journal Rank* (SJR) para las revistas indizadas en Scopus. Aunque son muchas las razones y los factores que contribuyen a que un artículo sea citado, el número de citas demuestra que un artículo ha sido empleado en la literatura y, por tanto, estos indicadores se han utilizado como una aproximación del nivel de influencia e internacionalización de las revistas.

Entre las muchas críticas que han recibido los indicadores de impacto (Bordons, Fernández y Gómez, 2002; Vanclay, 2012) se encuentra la influencia de las autocitas. El término autocita se ha utilizado para indicar que el artículo citante y el citado tienen uno o más autores en común (autocita de autoría) o cuando los artículos citantes están publicados en la misma revista que los artículos citados (autoría de revista) (Egghe y Rousseau, 1990).

Las autocitas forman parte natural de la comunicación científica y son, en sí mismas, útiles, relevantes y funcionales en el proceso de publicación y transferencia del conocimiento (Ophhof, 2013), además de configurarse como un fenómeno que es necesario medir (Gorbea-Portal y Suárez-Balseiro, 2007). No obstante, también pueden ser inapropiadas y utilizarse como estrategia de promoción, tanto de los propios autores para mejorar su visibilidad e impacto (Hyland, 2003) como de las revistas cuando los editores sugieren a los autores que citen artículos publicados en sus revistas –*coercive journal self citation*– (Ioannidis, 2015; Wilhite y Fong, 2012; Foley y Della Sala, 2010), consiguiendo con estas malas prácticas sobredimensionar su impacto percibido, un fenómeno cada vez más frecuente en todos los ámbitos

científicos (Chorus y Waltman, 2016). Por estas razones, son varios los autores que se decantan por su eliminación, ya que no reflejan el impacto real de una publicación (Aksnes, 2003; Schreiber, 2007) y suponen una perversión de los propios índices (Espinet Rubio, 2006); incluso algunos autores proponen la penalización de las autocitas de revista (De Granda Orive, Alonso Arroyo y Aleixandre Benavent, 2014).

Son varios los estudios que han analizado los efectos de las autocitas en el factor de impacto de las revistas y con resultados contradictorios. En algunos se ha hallado que las autocitas aumentan el factor de impacto de las revistas (Krauss, 2007; Frandsen, 2007), mientras que otros sugieren que su efecto es insignificante (Finardi, 2013; Campanario, 2010), incluido un estudio propio del antiguo Thomson Reuters en el que se concluye que el rango de las revistas en el *JCR* se ve poco alterado, sobre todo en aquellas revistas con mayor factor de impacto (McVeigh, 2002).

A lo largo de los años se han ido introduciendo modificaciones en los indicadores con el propósito de eliminar o mitigar el efecto de las autocitas y, de esta manera, reflejar con mayor precisión su impacto en la comunidad científica. Así, el factor de impacto, el *Eigenfactor Score* y el *Article Influence Score* excluyen las autocitas, mientras que el *SJR* limita el número de autocitas en su fórmula.

También se han desarrollado otros indicadores como el índice *h*, que combina tanto la producción científica como el impacto. Propuesto inicialmente por Hirsch para cuantificar la producción científica de los investigadores (Hirsch, 2005), comenzó a utilizarse también para evaluar las revistas (Braun, Glänzel y Schubert, 2006), siendo un indicador que aporta nueva información y que, por tanto, puede complementar a los indicadores de impacto (Salvador Oliván y Agustín Lacruz, 2015).

Se ha cuestionado la robustez del índice *h* ya que, al igual que cualquier indicador que se base en el número de citas, es susceptible de ser alterado por la autocitación; a pesar de ello, es un índice que se publica en las distintas bases de datos sin ser corregido por las autocitas. La mayoría de las publicaciones que han estudiado esta relación se centra en el índice *h* de los autores, y aunque Hirsch argumentó que el efecto de las autocitas es relativamente pequeño (Hirsch, 2005), son varios los estudios que han demostrado que puede ser importante (Bartneck y Kokkermans, 2011), sobre todo en los autores con índices *h* más bajos (Gianoli y Molina-Montenegro, 2009) y se recomienda eliminarlas de su cálculo (Schreiber, 2007). En cuanto al estudio de la relación de las autocitas con el índice *h* de las revistas, hemos encontrado sólo un artículo en el área de aplicaciones informáticas de la base de datos Scopus, concluyendo que afecta al rango de las revistas (Ferrara y Romero, 2013).

La constatación de la escasez de publicaciones sobre la visibilidad de las revistas iberoamericanas en Biblioteconomía/Bibliotecología y Documentación (ByD) a través de las citas recibidas en WoS (anteriormente ISI)¹ y cómo afectan las autocitas al impacto de las publicaciones y al índice h de las revistas han sido las razones que nos han llevado a realizar este trabajo.

En línea con lo expuesto, el objeto general de este artículo es el estudio de la repercusión de los artículos de mayor impacto en WoS publicados en las dos revistas españolas de ByD listadas en el *Journal Citation Reports* (JCR), a saber: *El Profesional de la Información* (EPI) y *Revista Española de Documentación Científica* (REDOC).

En relación con dicho objeto general, los objetivos específicos de nuestro trabajo son:

- Conocer el impacto, medido en citas recibidas, según la procedencia de autores y revistas: iberoamericanas e internacionales.
- Contrastar si hay diferencia en el impacto de las dos revistas españolas listadas en el *Journal Citation Reports* (JCR): *El Profesional de la Información* (EPI) y *Revista Española de Documentación Científica* (REDOC).
- Determinar las revistas de las que procede el mayor número de citas.
- Conocer las tasas de autocitación, tanto de las revistas como de los autores.
- Determinar el impacto de las autocitas de las revistas en su índice h corregido y en el *ranking* de los artículos.

METODOLOGÍA

Fuentes de datos. Se seleccionaron las revistas españolas de ByD listadas en el JCR: EPI y REDOC. El 1 de junio de 2017 se buscaron en la base de datos WoS los artículos publicados en dichas revistas.

Selección de los artículos. De cada revista se seleccionaron únicamente los artículos más citados: aquellos con un número de citas recibidas igual o superior al índice h. Para ello, se ordenaron los registros recuperados por el número de citas recibidas. El índice h de la revista EPI era de 14 y se seleccionaron 14 artículos. El índice h de REDOC era 11 y se seleccionaron 13 artículos que habían sido citados al menos 11 veces.

Extracción de datos. Se descargaron los datos bibliográficos de los artículos seleccionados y de los artículos citantes. De estos últimos, se examinaron

1 Respondiendo a dicho enfoque sólo hemos hallado un artículo (Herrero Solana y Liberatore, 2008), el cual no incluía la revista *El Profesional de la Información*, ya que en ese momento no estaba recogida en la base de datos citada.

los autores y los títulos de las fuentes de publicación para calcular los siguientes indicadores:

- Autocitas de revista: número de citas recibidas por artículos publicados en la misma revista.
- Citas de revistas iberoamericanas: artículos publicados en español o portugués.
- Citas de revistas internacionales: artículos publicados en idioma distinto al español o portugués.
- Autocitas de autores: al menos un autor del artículo citante es autor del artículo citado, independientemente de su posición en el orden de firma en ambos artículos.

Para los propósitos de este trabajo, se ha utilizado el término “Iberoamérica” para referirnos a las revistas o autores de países de habla española o portuguesa de América, Portugal y España, y el término “Internacional” para revistas o autores cuyo idioma nativo es el inglés o cualquier otro que no sea español o portugués.

Además, se clasificaron las fuentes citantes en tres categorías: revistas incluidas en el JCR, revistas pertenecientes al *Emerging Sources Citation Index (ESCI)* y publicaciones procedentes del *Conference Proceedings Citation Index*.

Análisis de datos. El proceso y análisis estadístico de los datos se realizó con el programa SPSS versión 22. Para contrastar las diferencias entre las dos revistas, se aplicó el test de chi cuadrado o la t de Student según el nivel de medida de las variables. Se estudió también la correlación entre el índice h y el corregido sin autocitas.

RESULTADOS Y DISCUSIÓN

Características generales de todos los artículos de las revistas

El número de artículos publicados por EPI e indizados en WoS es casi tres veces más que en el caso de la revista REDOC (955 *vs* 346), proporción similar al número total de citas recibidas en ambas revistas (1 603 *vs* 613). A pesar de estas diferencias en las cifras absolutas, el porcentaje de artículos que no reciben ninguna cita es muy similar en ambas revistas (48.4 % *vs* 49.4 %).

En cuanto a las citas recibidas de la propia revista, el porcentaje de autocitas en EPI es mayor que en REDOC (37.4 % *vs* 32.1 %), en consonancia

con un menor porcentaje de artículos citantes procedentes de otras revistas (70.8 % *vs* 75.8 %).

| | <i>EPI</i> | <i>REDOC</i> | p^a |
|--------------------------------------------------|-------------------|-----------------|-------|
| Nº de artículos publicados | 955 | 346 | |
| Nº artículos no citados | 462 (48.4 %) | 171 (49.4 %) | 0.739 |
| Nº de artículos en que se cita | 1 105 | 468 | |
| Nº de artículos en que se cita sin citas propias | 783 (70.8 %) | 355 (75.8 %) | 0.043 |
| Total de citas | 1 603 | 613 | |
| Autocitas de revista | 600 (37.4 %) | 197 (32.1 %) | 0.020 |
| Total sin autocitas revista | 1 003 (62.6 %) | 416 (67.8 %) | |
| Índice h | 14 | 11 | |
| Nº de artículos que cumplen con el índice h | 14 | 13 | |

Tabla 1. Características generales de los artículos de *EPI* y *REDOC* en WoS

^a Comparación de proporciones: chi cuadrado

Características de las citas de los artículos más citados

El número total de citas recibidas por los artículos de *EPI* es muy superior a las recibidas por *REDOC*, con un promedio de 18.5 citas recibidas por documento en *EPI* frente a las 13.1 citas por documento en *REDOC* ($p=0.002$).

Al analizar la procedencia de las citas, aunque no se observan diferencias estadísticamente significativas, destaca lo siguiente:

- El porcentaje de autocitas de revista es algo superior al 25 % en ambas revistas, siendo un poco más alto en *EPI* que en *REDOC*, si bien esta diferencia es menor (1.6 %) que cuando se analizan todos los artículos publicados (5.3 %).
- El porcentaje de citas procedentes de revistas iberoamericanas es prácticamente el mismo en las dos revistas, el 42 %. Sin embargo, al distinguir entre nacionales (entendidas, dado el alcance del trabajo, como españolas) o no nacionales, *EPI* recibe un 2.5 % más de citas nacionales mientras que *REDOC* recibe más citas de revistas de otros países iberoamericanos.
- En lo referente a las citas de revistas internacionales, ambas revistas obtienen porcentajes similares, casi un punto más en *REDOC* debido a las

citas de los propios autores o de autores iberoamericanos, mientras que *EPI* recibe más citas de autores no iberoamericanos.

- El impacto de los artículos publicados en *EPI* y *REDOC* es principalmente en revistas en español o portugués, ya que dos de cada tres citas proceden de revistas iberoamericanas o son autocitas de revista. La repercusión en revistas en inglés es muy baja, sólo una de cada 10 es llevada a cabo por autores de países no iberoamericanos, y la mayoría de estas citas son a artículos publicados en inglés; así, en *EPI*, hay tres artículos publicados en inglés y escritos por autores no iberoamericanos que reciben 25 de las 30 citas, y en *REDOC*, 13 de las 17 citas las recibe un artículo publicado en inglés.

| | <i>EPI</i> | <i>REDOC</i> | <i>P^b</i> | | |
|------------------------------------------|------------|--------------|----------------------|---------|-------|
| | Nº | % | Nº | % | |
| <i>Citas de la propia revista</i> | 71 | 27.3 % | 44 | 25.7 % | 0.717 |
| <i>Citas de revistas iberoamericanas</i> | 110 | 42.3 % | 73 | 42.7 % | 0.937 |
| Revistas nacionales | 69 | 26.5 % | 41 | 24.0 % | 0.551 |
| Revistas no nacionales | 41 | 15.8 % | 32 | 18.7 % | 0.425 |
| <i>Citas de revistas internacionales</i> | 79 | 30.4 % | 54 | 31.6 % | 0.793 |
| Por propios autores | 22 | 8.5 % | 16 | 9.4 % | 0.748 |
| Por autores iberoamericanos | 27 | 10.4 % | 21 | 12.3 % | 0.540 |
| Por autores no iberoamericanos | 30 | 11.5 % | 17 | 9.9 % | 0.603 |
| <i>Total de citas</i> | 260 | 100.0 % | 171 | 100.0 % | |

Tabla 2. Características de citación de los artículos más citados de *EPI* y *REDOC* en WoS

^b Comparación de proporciones: chi cuadrado

Revistas citantes

Excluyendo las autocitas de las revistas, casi las tres cuartas partes de las citas proceden de revistas incluidas en el *JCR* y muy pocas son de actas de congresos. Existe un porcentaje importante de citas, el 22.2 %, que proceden de revistas que están en el *Emerging Sources Citation Index* (base de datos lanzada en 2015 para ampliar la cobertura geográfica a revistas de otros países de lengua no inglesa que están en proceso de evaluación antes de ser incluidas en el *JCR*), por lo que no tienen factores de impacto, pero sí se utilizan sus artículos en el conteo de citas; contribuyen, de esta manera, a engrosar los factores de impacto de las revistas *JCR* de aquellos países con menos impacto debido a la barrera del idioma en que publican.

Al examinar las diferencias entre las dos revistas, en *REDOC* hay un porcentaje más alto de citas de revistas *JCR*, mientras que en *EPI* es mayor el de citas procedentes de revistas del *Emerging Sources Citation Index*. Así se desprende también de la información presentada en las *Tablas 4 y 5*, en *EPI* hay cuatro revistas de fuentes emergentes en los primeros lugares, mientras que en *REDOC* sólo hay una revista proveniente de *ESCI* y con dos citas.

| | <i>EPI</i> | <i>REDOC</i> | Total | | |
|----------------------------------------------|------------|--------------|-------|---------|-----|
| | Nº | % | Nº | % | Nº |
| Revistas <i>JCR</i> sin autocitas de revista | 133 | 70.4 % | 98 | 77.2 % | 231 |
| <i>Emerging Sources Citation Index</i> | 48 | 25.4 % | 22 | 17.3 % | 70 |
| <i>Conference Proceedings Citation Index</i> | 8 | 4.2 % | 7 | 5.5 % | 15 |
| Total sin autocitas de revista | 189 | 100.0 % | 127 | 100.0 % | 316 |

Tabla 3. Distribución de citas según la base de datos de revistas

El número de revistas que citan los artículos de *EPI* es muy superior al de *REDOC* (82 *vs* 59), algo lógico ya que el número de citas recibidas también es más alto. Según el idioma de publicación de la revista, en *EPI* es mayor el porcentaje de revistas internacionales (52.4 %), mientras que en *REDOC* es mayor en las revistas iberoamericanas (61 %).

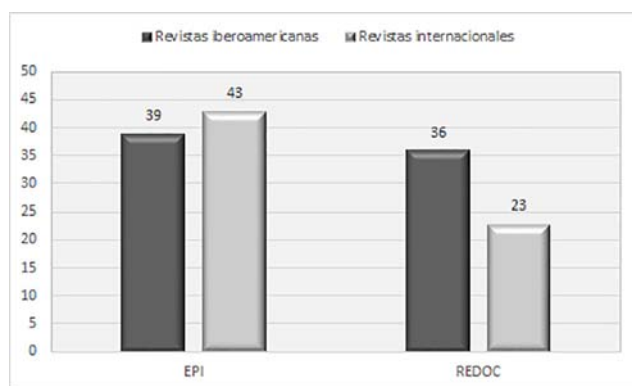


Figura 1. Distribución del número de revistas citantes según el idioma/país

Las revistas que más citan se presentan en las *Tablas 4 y 5*. En ambos casos coinciden las tres revistas que más citas producen: la revista mexicana (*Investigación Bibliotecológica*), una española, que coincide con *REDOC* o *EPI*

(según sea el caso) y una revista en inglés, *Scientometrics*. Cabe señalar que aunque el número de citas que suman estas tres revistas es más alto para *EPI* (56 citas) que para *REDOC* (33 citas), proceden del mismo número de artículos, 21 en ambos casos. También destaca que entre las revistas que más citan a *EPI* se encuentran más revistas procedentes del *Emerging Sources Citation Index* que en el caso de *REDOC*, y que las revistas internacionales más citantes están especializadas en temas de bibliometría y cienciometría.

| | | Nº citas | Nº artículos |
|------|----------------------------------------------------------------------------|----------|--------------|
| JCR | Revista Española de Documentación Científica | 21 | 8 |
| JCR | Investigación Bibliotecológica | 19 | 9 |
| JCR | Scientometrics | 16 | 4 |
| ESCI | Revista Latina de Comunicación Social | 7 | 4 |
| JCR | Journal of Informetrics | 7 | 3 |
| JCR | Comunicar | 6 | 4 |
| ESCI | Historia y Comunicación Social | 4 | 4 |
| ESCI | Revista Icono 14- Revista Científica de Comunicación y Tecnologías | 5 | 3 |
| JCR | Transinformação | 3 | 3 |
| ESCI | Observatorio | 3 | 3 |
| JCR | Journal of the Association for Information Science and Technology | 3 | 3 |
| CPCI | International Society of Scientometrics and Informetrics Conference (ISSI) | 3 | 3 |
| JCR | Information Research | 3 | 3 |
| | Total sin autocitas | 189 | 133 |

Tabla 4. Revistas que más citan a *EPI*

* Las celdas sombreadas son revistas internacionales y las no sombreadas son revistas latinoamericanas

| | | Nº citas | Nº artículos |
|------|----------------------------------------------------------------------------|----------|--------------|
| JCR | Investigación Bibliotecológica | 14 | 9 |
| JCR | El Profesional de la Información | 10 | 5 |
| JCR | Scientometrics | 9 | 7 |
| CPCI | International Society of Scientometrics and Informetrics Conference (ISSI) | 7 | 3 |
| JCR | Journal of the Association for Information Science and Technology | 7 | 2 |
| JCR | Comunicar | 6 | 3 |

| | | | |
|---------------------|----------------------------------------------------------------|-----|----|
| JCR | <i>Journal of Informetrics</i> | 5 | 3 |
| JCR | <i>PLOS One</i> | 3 | 3 |
| JCR | <i>Transinformação</i> | 2 | 2 |
| JCR | <i>Revista Mexicana de Psicología</i> | 2 | 2 |
| JCR | <i>Revista de Psicología del Deporte</i> | 2 | 2 |
| JCR | <i>Online Information Review</i> | 2 | 2 |
| JCR | <i>Library Collections, Acquisition and Technical Services</i> | 2 | 2 |
| JCR | <i>Information Research</i> | 2 | 2 |
| ESCI | <i>Aula Abierta</i> | 2 | 2 |
| JCR | <i>Adicciones</i> | 2 | 2 |
| Total sin autocitas | | 127 | 94 |

Tabla 5. Revistas que más citan a REDOC

* Las celdas sombreadas son revistas internacionales y las no sombreadas son revistas latinoamericanas

Autocitación e índice h

Si se tiene en cuenta de manera conjunta las autocitas de la revista y de los autores, se obtienen unas tasas de autocitación elevadas y muy parecidas en las dos revistas, 41.9 % en *EPI* y 43.8 % en *REDOC*. En ambos casos, la autocitación de revista es más alta que la que procede de los autores (*Tabla 6*).

Al analizar la distribución en las dos revistas, el porcentaje de autocitas de revista en *EPI* es ligeramente mayor que en *REDOC*, mientras que en ésta es mayor el porcentaje de autocitas de autores. La tasa de autocitación de autores, entre el 19.2 % y el 24.6 %, es bastante más baja que el 47 % obtenido por Jiménez Contreras, Delgado López Cózar y Ruiz Pérez (2006), lo que pensamos que puede ser debido a varios factores: en primer lugar, a que esta tasa se calculó con todos los trabajos publicados por autores españoles en WoS; en segundo lugar a que, tal y como observaron los propios autores, aquellos con más citas son los que presentan menor tasa de autocitación; en último lugar, al hecho de que el foco de nuestro estudio son los artículos más citados.

Al comparar la tasa de autocitación de la revista para todos los artículos incluidos en WoS (*Tabla 1*), se observa que en los artículos más citados es menor en ambos casos, 10.1 puntos menos en *EPI* y 6.4 puntos menos en *REDOC*, lo que significa que los artículos menos citados en estas revistas aún tienen menos impacto real al presentar más citas que proceden de la propia revista.

| | <i>EPI</i> | | <i>REDOC</i> | |
|---------------------------------|------------|---------|--------------|---------|
| | Nº | % | Nº | % |
| Total de autocitas | 109 | 41.9 % | 75 | 43.8 % |
| Autocitas de autores | 50 | 19.2 % | 42 | 24.6 % |
| Autocitas de revista | 71 | 27.3 % | 44 | 25.7 % |
| Autocitas autores en la revista | 12 | 4.6 % | 11 | 6.4 % |
| Total de citas | 260 | 100.0 % | 171 | 100.0 % |

Tabla 6. Tasa de autocitación en los artículos más citados de *EPI* y *REDOC*

En las *Tablas 7 y 8* se presentan los artículos analizados con el número de citas totales y excluidas las autocitas, así como el rango obtenido una vez eliminadas las autocitas de revista, que son las únicas autocitas que se tienen en cuenta al calcular y corregir los indicadores de impacto de las revistas *JCR*. Estamos de acuerdo con Jiménez Contreras, Delgado López-Cózar y Ruiz Pérez (2006) cuando afirman que al establecer un escalafón deben excluirse las autocitas y, aunque en el contexto de su estudio se refieren a las autocitas de autor, creemos que debe extenderse a todo tipo de autocitas, de ahí que en las tablas presentemos la exclusión de las autocitas de autor y de revista.

Al excluir las autocitas de revista, en la revista *EPI* el índice *h* bajaría de 14 a 9, y en *REDOC* de 11 a 9, con lo que se igualarían en la base de datos. El descenso en el índice *h* depende no sólo de las autocitas sino también de que haya más artículos con un número de citas próximo al índice *h*; en este caso, se debe al mayor número de autocitas en *EPI*.

En cuanto al rango de los artículos, existen cambios importantes en algunos de ellos, aunque lógicamente, aquellos con mayor número de citas se ven menos afectados por la autocitación de la revista.

Sí se ha hallado en *EPI* una relación entre el número de autores firmantes y el número de autocitas de autores, de manera que a mayor número de autores mayor es la tasa de autocitación de autores ($\rho=0,606$, $p=0,022$); sin embargo, en *REDOC* no se ha encontrado ningún tipo de relación ($\rho=-0,216$; $p=0,479$). En la literatura también existen resultados contradictorios; así, Costas, Van Leeuwen y Bordons (2010) encontraron que el número de autocitas suele aumentar conforme mayor es el número de autores, pero Shah, Gul y Gaur (2015) no encontraron ningún tipo de relación en el campo de la ByD.

Aunque el número de artículos analizados en nuestra investigación no permite afirmar si existe algún tipo de patrón que pudiera explicar las tasas de autocitación, cabe destacar que los tres artículos publicados en *EPI* por

autores internacionales (R1=3, R1=6 y R1=8) tienen como máximo una autocita de la revista, si bien hay otros artículos de autores españoles que tampoco tienen autocitas de revista (R1=5) o muy pocas (R1=10, R1=14), coincidiendo en dos de ellos algunos autores. En *REDOC* no se observa ningún patrón, salvo la coincidencia de que dos artículos firmados por los mismos autores tienen tasas altas de autocitación (R1=8, R1=10), coincidiendo también en *EPI* con el artículo con mayor tasa de autocitación de autores (R1=1).

| | A | B | C | D | R1 | R2 |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|----|----|----|----|----|
| Aleixandre-Benavent, R.; Valderrama-Zurian, J.C.; Gonzalez-Alcaide, G. "Scientific journals impact factor: limitations and alternative indicators". 2007. | 28 | 23 | 15 | 11 | 1 | 1 |
| Casero-Ripolles, A. "Newspapers on the internet: new business models in the convergence era". 2010. | 27 | 16 | 24 | 14 | 2 | 5 |
| Leydesdorff, L.; Wagner, C.S.; Park, H-W; Adams, J. "International collaboration in science: the global map and the network". 2013. | 20 ^a | 20 | 16 | 16 | 3 | 2 |
| Rodriguez-Martinez, R.; Codina, L.; Pedraza-Jimenez, R. "Online journalism and web 2.0: A model for analysis and report on results". 2010. | 21 | 12 | 19 | 10 | 4 | 9 |
| Torres-Salinas, D.; Jimenez-Contreras, E. "Introduction and comparative study of the new scientific journals citation indicators in Journal Citation Reports and Scopus". 2010. | 20 | 20 | 20 | 20 | 5 | 3 |
| Jeremic, V.; Jovanovic-Milenkovic, M.; Radojicic, Z.; Martic, M. "Excellence with leadership: the Crown indicator of Scimago institutions rankings Iber report". 2013. | 18 | 17 | 8 | 8 | 6 | 4 |
| Guallar, J.; Abadal, E. "Evaluation of digital press archives: indicators and examples of good practice". 2009. | 18 | 9 | 13 | 6 | 7 | 10 |
| Leydesdorff, L. "World shares of publications of the USA, EU-27, and China compared and predicted using the new Web of Science interface versus Scopus". 2012. | 16 | 15 | 15 | 14 | 8 | 6 |
| Diaz-Noci, J. "Trends in internet media". 2010. | 16 | 6 | 16 | 6 | 9 | 14 |
| Torres-Salinas, D.; Bordons, M.; Gimenez-Tolledo, E.; Delgado-Lopez-Cozar, E.; Jimenez-Contreras, E.; Sanz-Casado, E. "Integrated scientific journals classification (CIRC): A proposal for categorization of journals in social sciences and humanities". 2010. | 16 | 14 | 11 | 10 | 10 | 7 |

| | | | | | | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|----|----|----|----|----|
| Peset, F.; Ferrer-Sapena, A.; Subirats-Coll, I. "Linked open data and Open data. Its impact in the field of libraries and information science". 2011. | 15 | 7 | 14 | 7 | 11 | 12 |
| Guallar, J.; Rovira, C.; Ruiz, S. "Multimedia in digital media. Multimedia elements and retrieval systems in the leading Spanish online newspapers". 2010. | 15 | 7 | 12 | 7 | 12 | 13 |
| Cabezas-Clavijo, A.; Torres-Salinas, D.; Delgado-Lopez-Cozar, E. "Science 2.0: Tools catalogue and consequences for scientific activity". 2009. | 15 | 9 | 12 | 8 | 13 | 11 |
| Costas, R.; Bordons, M. "A critical view of the h-index: observations based on a practical application". 2007. | 15 | 14 | 15 | 14 | 14 | 8 |
| <p>A: N° de citas en la base de datos; B: N° de citas sin autocitas de la revista; C: N° de citas sin autocitas de los autores; D: N° de citas sin autocitas; R1: Rango del artículo según número de citas; R2: rango del artículo sin autocitas de revista ^a Corresponde al número real de citas que se pueden descargar del WoS, aunque en el artículo aparece con 1 cita más</p> | | | | | | |

Tabla 7. Rango y número de citas sin autocitas de los artículos de EPI

| | A | B | C | D | R1 | R2 |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|----|----|----|----|----|
| Torres-Salinas, D.; Robinson-Garcia, N.; Jimenez-Contreras, E.; Delgado Lopez-Cozar, E. "Towards a Book Publishers Citation Reports. First approach using the Book Citation Index". 2012. | 25 | 24 | 18 | 18 | 1 | 1 |
| Bordons, M.; Sancho, R.; Morillo, F.; Gómez, I. "Scientific activity profile of Spanish universities in four thematic areas: a multifactor analysis". 2010. | 17 | 9 | 15 | 9 | 2 | 6 |
| Rodriguez-Martinez, R.; Codina, L.; Pedraza-Jimenez, R. "Indicators to evaluate the quality of online journalism web sites: an analysis of Web 2.0 interaction and adoption". 2012. | 13 | 10 | 11 | 8 | 3 | 3 |
| De Filippo, D.; Sanz-Casado, E.; Urbano Salido, C.; Ardanuy, J.; Gomez-Caridad, I. "The role of institutional databases in the analysis of universities' scientific activity". 2011. | 13 | 9 | 12 | 9 | 4 | 7 |
| Dorta-Gonzalez, P.; Dorta-Gonzalez, M.I. "Bibliometric indicator based on the h-index". 2010. | 13 | 12 | 6 | 6 | 5 | 2 |

| | | | | | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|----|----|---|----|----|
| Fernández Quijada, F. "The profile of Spanish journals in the field of communication (2007-2008)". 2010. | 12 | 10 | 5 | 4 | 6 | 4 |
| Ortega, J.L.; Aguillo, I. "Structural analysis of the Iberoamerican academic web". 2009. | 12 | 10 | 11 | 9 | 7 | 5 |
| Gonzalez Alcaide, G.; Castellano Gomez, M.; Valderrama Zurian, J.C.; Aleixandre Benavent, R. "Scientific literature by Spanish authors on the analysis of citations and impact factor in Biomedicine (1981-2005)". 2008. | 11 ^a | 5 | 6 | 3 | 8 | 13 |
| De Filippo, D.; Morillo, F.; Teresa Fernandez, M. "Indicators of scientific collaboration between CSIC and Latin America through international databases". 2008. | 12 | 8 | 8 | 6 | 9 | 10 |
| Gonzalez Alcaide, G.; Valderrama Zurian, J.C.; Aleixandre Benavent, R. "Analysis of the internationalization process of Spanish research in science and technology (1980-2007)". 2012. | 10 ^a | 6 | 6 | 4 | 10 | 11 |
| Jimenez-Contreras, E.; Robinson-Garcia, N.; Cabezas-Clavijo, A. "Productivity and impact of Spanish researchers: reference thresholds within scientific areas". 2011. | 11 | 6 | 10 | 5 | 11 | 12 |
| Torres-Salinas, D.; Muñoz-Muñoz, A.M.; Jimenez-Contreras, E. "Bibliometric analysis of the situation of female researchers in the fields of social sciences and law in Spain". 2011. | 11 | 9 | 11 | 9 | 12 | 8 |
| Carlos Sierra, J.; Buella-Casal, G.; Bermudez Sanchez, M.P.; Santos-Iglesias, P. "Opinion of Associate Professors and Full Professors regarding the criteria and standards for the accreditation of Spanish university teachers". 2009 | 11 | 9 | 10 | 8 | 13 | 9 |
| <p>A: Nº de citas en la base de datos, B: Nº de citas sin autocitas de la revista; C: Nº de citas sin autocitas de los autores; D: Nº de citas sin autocitas; R1: Rango del artículo según número de citas; R2: rango del artículo sin autocitas de revista</p> <p>^a Corresponde al número real de citas que se pueden descargar del WoS, aunque en el artículo aparece con 1 cita más</p> | | | | | | |

Tabla 8. Rango y número de citas sin autocitas de los artículos de REDOC

CONCLUSIONES Y RECOMENDACIONES

De los resultados obtenidos, y atendiendo a los objetivos propuestos, se derivan las siguientes conclusiones y recomendaciones:

La repercusión de los artículos más citados publicados en las revistas españolas *EPI* y *REDOC* se produce principalmente en la comunidad científica iberoamericana, lo que significa que su impacto internacional es muy bajo, ya que sólo una de cada 10 citas proceden de autores no iberoamericanos en revistas internacionales, y la mayoría de ellas las reciben artículos publicados en inglés. Por lo tanto, los artículos analizados publicados en las revistas españolas seleccionadas tienen muy bajo impacto internacional y son citados por autores que publican en revistas que utilizan este idioma como vehículo de comunicación.

El impacto de los artículos de *EPI* y *REDOC* en revistas iberoamericanas e internacionales es muy similar. En *EPI* la revista con mayor número de citas es *REDOC*, seguida de la revista mexicana *Investigación Bibliotecológica* y de una revista internacional publicada en inglés, *Scientometrics*. En *REDOC*, las tres revistas más citantes son las mismas, salvo que el orden es primero para *Investigación Bibliotecológica* y *EPI* en segunda posición.

Existe una tasa muy alta de autocitación (cerca del 40 %), por lo que la repercusión externa de los artículos es baja y su impacto (número de citas) depende en cierta medida de los propios autores y de la revista.

La eliminación de las autocitas de revista afecta al índice *h* y al orden de los artículos. Al igual que se calculan en el *JCR* los indicadores de impacto corregidos por las autocitas, pensamos que sería recomendable que se publicara también el índice *h* sin autocitas en todas las bases de datos, con la finalidad de dar una aproximación más real y objetiva de su impacto en la comunidad científica.

Hay ocasiones en que la autocitación está justificada, pero otras veces se detecta una tasa excesivamente alta de autocitas o están injustificadas, al no tener relación con trabajos previos. Las autocitas de autores proporcionan información valiosa sobre el proceso de comunicación científica, pero si se desea evaluar el impacto de un trabajo y/o autor, las citas externas son más relevantes, ya que reflejan la repercusión en la comunidad científica y, por consiguiente, pensamos que deberían ser eliminadas las autocitas.

Sería deseable averiguar si la autocitación y el impacto de los estudios publicados en revistas iberoamericanas en otras áreas de conocimiento siguen los mismos patrones que los obtenidos en el campo de la ByD y si su repercusión se centra casi exclusivamente entre la comunidad científica de la misma región.

REFERENCIAS

- Aksnes, Dag W. 2003. "A macro-study of self-citation". *Scientometrics* 56 (2): 235-246. doi: 10.1023/A:1021919228368
- Bartneck, C. y S. Kokkermans. 2011. "Detecting h-index manipulation through self-citation analysis". *Scientometrics* 87 (1): 85-98. doi: 10.1007/s11192-010-0306-5
- Bordons, María, M. T. Fernández e Isabel Gómez. 2002. "Advantages and limitations in the use of impact factor measures for the assessment of research performance in a peripheral country". *Scientometrics* 53 (2): 195-206. doi: 10.1023/A:1014800407876
- Braun, Tibor, Wolfgang Glänzel y András Schubert. 2006. "A Hirsch-type index for journals". *Scientometrics* 69 (1): 169-173. doi: 10.1007/s11192-006-0147-4
- Campanario, Juan Miguel. 2010. "Self-citations that contribute to the journal impact factor: An investment-benefit-yield analysis". *Journal of the American Society for Information Science and Technology* 61 (12): 2575-2580. doi: 10.1002/asi.21439
- Costas, Rodrigo, Thed N. van Leeuwen y María Bordons. 2010. "Self-citation at the meso and individual levels: effects of different calculation methods". *Scientometrics* 82 (3): 517-537. doi: 10.1007/s11192-010-0187-7
- Chorus, Caspar y Ludo Waltman. 2016. "A large-scale analysis of impact factor biased journal self-citations". *PLOS One* 11 (8): e0161021. doi:10.1371/journal.pone.0161021
- De Granda Orive, José Ignacio, Adolfo Alonso Arroyo y Rafael Aleixandre Benavent. 2014. "Autocitación: ¿debemos penalizarla?". *Archivos de Bronconeumología* 50 (10): 458. doi: 10.1016/j.arbres.2013.10.008
- Egghe, Leo y Ronald Rousseau. 1990. *Introduction to Informetrics: quantitative methods in library, documentation and information science*. Amsterdam: Elsevier Science Publishers.
- Espinete Rubio, Pablo. 2006. "Perversiones, paradojas y traumas en la evaluación científica". *Anales de Química* 102 (4): 60-63.
- Ferrara, Emilio y Alfonso E. Romero. 2013. "Scientific impact evaluation and the effect of self-citations: mitigating the bias by discounting h-index". *Journal of the American Society for Information Science and Technology* 64 (11): 2332-2339. doi: 10.1002/asi.22976
- Finardi, Ugo. 2013. "Correlation between journal impact factor and citation performance: An experimental study". *Journal of Informetrics* 7 (2): 357-370. doi: 10.1016/j.joi.2012.12.004
- Foley, Jennifer A. y Sergio Della Sala. 2010. "The impact of self-citation". *Cortex* 46 (6): 802-810. doi: 10.1016/j.cortex.2010.01.004
- Frandsen, Tove Faber. 2007. "Journal self-citations – analyzing the JIF mechanism". *Journal of Informetrics* 1 (1): 47-58. doi: 10.1016/j.joi.2006.09.002
- Gianoli, Ernesto y Marco A. Molina-Montenegro. 2009. "Insights into the relationship between the h-index and self-citation". *Journal of the American Society for Information Science and Technology* 60 (6): 1283-1285. doi: 10.1002/asi.21042
- Gorbea-Portal, Salvador y Carlos A. Suárez-Balseiro. 2007. "Análisis de la influencia y el impacto entre revistas periféricas no incluidas en el Science Citation Index". *Revista Interamericana de Bibliotecología* 30 (2): 47-70.
- Herrero Solana, Víctor y Gustavo Liberatore. 2008. "Visibilidad internacional de las revistas iberoamericanas de Bibliotecología y Documentación". *Revista Española de Documentación Científica* 31 (2): 230-239.

- Hirsch, J. 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. doi:10.1073/pnas.0507655102
- Hyland, Ken. 2003. "Self-citation and self-reference: Credibility and promotion in academic publication". *Journal of the American Society for Information Science and Technology* 54 (3): 251-259. doi: 10.1002/asi.10204
- Ioannidis, John P.A. 2015. "A generalized view of self-citation: Direct, co-author, co-laborative, and coercive induced self-citation". *Journal of the Psychosomatic Research* 78 (1): 7-11. doi: 10.1016/j.jpsychores.2014.11.008
- Jiménez Contreras, Evaristo, Emilio Delgado López-Cózar y Rafael Ruiz Pérez. 2006. "Producción española en biblioteconomía y documentación con visibilidad internacional a través del Web of Science (1995-2004)". *El Profesional de la Información* 15 (5): 373-383.
- Krauss, Jochen. 2007. "Journal self-citation rates in ecological sciences". *Scientometrics* 73 (1): 79-89. doi: 10.1007/s11192-007-1727-7
- McVeigh, M. E. 2002. Journal self-citation in the *Journal Citation Reports – Science Edition* (2002). Fecha de consulta: 13 de septiembre de 2017. <http://wokinfo.com/essays/journal-self-citation-jcr/>
- Opthof, T. 2013. "Inflation of impact factors by journal self-citation in cardiovascular science". *Netherlands Heart Journal* 21 (4): 163-165. doi: 10.1007/s12471-013-0384-0
- Salvador Oliván, José Antonio y Carmen Agustín Lacruz. 2015. "Correlación entre indicadores bibliométricos en revistas de Web of Science y Scopus". *Revista General de Información y Documentación* 25 (2): 341-359. http://dx.doi.org/10.5209/rev_RGID.2015.v25.n2.51241
- Schekman, Randy. 2013. "How journals like Nature, Cell and Science are damaging science". *The Guardian*. Fecha de consulta: 12 de septiembre de 2017. <https://www.theguardian.com/commentisfree/2013/dec/09/how-journals-nature-science-cell-damage-science>
- Schreiber, M. 2007. "Self-citation corrections for the Hirsch index". *Europhysics Letters* 78 (3): 30002. doi: 10.1209/0295-5075/78/30002
- Shah, Tariq Ahmad, Sumeer Gul y Ramesh C. Gaur. 2015. "Authors self-citation behavior in the field of Library and Information Science". *Aslib Journal of Information Management* 67(4): 48-468. doi: 10.1108/AJIM-10-2014-0134
- Vanclay, Jerome K. 2012. "Impact factor: outdated artefact or stepping-stone to journal certification?". *Scientometrics* 92 (2): 211-238. doi: 10.1007/s11192-011-0561-0
- Wilhite, Allen W. y Eric A. Fong. 2012. "Coercitive citation in academic publishing". *Science* 335 (6068): 542-543. doi: 10.1126/science.1212540

Para citar este texto:

Salvador-Oliván, José Antonio, Gonzalo Marco-Cuenca y Rosario Arquerro-Avilés. 2018. "Impacto de las revistas españolas de Biblioteconomía y Documentación y repercusión de las autocitas en su índice h". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 13-30.

<http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57852>

Origen y desarrollo de la editorial Espasa-Calpe Mexicana (1927-1975)

Juan Miguel Sánchez Vigil*

Artículo recibido:
10 de diciembre de 2017

Artículo aceptado:
8 de agosto de 2018

Artículo de revisión

RESUMEN

La comercialización del libro en América fue una de las actividades económicas que más contribuyó al desarrollo industrial español a partir de la década de los veinte del siglo pasado, especialmente en México, donde se desarrolló una ardua tarea en favor de la cultura. Una de las primeras editoriales en establecerse allí fue Espasa-Calpe, que llevó a cabo una labor excepcional en la publicación y venta de obras de autores españoles y mexicanos. Es objeto de este trabajo presentar la actividad de Espasa-Calpe Mexicana desde su origen en 1927 hasta el final de la dictadura del general Franco en 1975, analizando las distintas fases de su desarrollo con el objetivo específico de aportar nuevos datos para la historia de la edición en ambos países.

* Facultad de Ciencias de la Documentación
Universidad Complutense de Madrid, España. jmvigil@ucm.es

Palabras clave: Edición en México; Edición en España; Espasa-Calpe Mexicana; Industria Editorial

Origin and history of the Espasa-Calpe Mexicana publishing house (1927-1975)

Juan Miguel Sánchez-Vigil

ABSTRACT

Marketing of books in America, and especially in Mexico, was an important economic activity contributing to Spain's industrial development from the 1920s forward, as it undertook the difficult task of enriching the cultural landscape. Publishing works by both Spanish and Mexican authors, Espasa-Calpe was among the first publishing houses to open in Mexico. This study provides an overview of the development of Espasa-Calpe Mexicana from its founding in 1927 until the end of the Franco dictatorship in 1975. By examining the distinct stages of its growth, this paper provides new insights into the history of publishing in both Spain and Mexico.

Keywords: Publishing in Mexico; Publishing in Spain; Espasa-Calpe Mexicana; Publishing Industry

INTRODUCCIÓN

La industria del libro vivió un importante cambio en España a partir de 1920 cuando las grandes editoriales decidieron acometer políticas de expansión en América. La fusión de Espasa y Calpe en diciembre de 1925 convirtió a la nueva empresa en la mayor de España en fondo y volumen de negocio, con especial interés por el mercado americano, donde Calpe disponía de una delegación en Buenos Aires desde 1922.

En 1927, al tiempo que Leopoldo Calvo Sotelo publicaba *Memoria sobre el libro español en América*, el Consejo de Administración decidió abrir un depósito comercial en México, origen de Espasa-Calpe Mexicana. Con carácter internacional, la industria editorial tuvo numerosos altibajos si tenemos en

cuenta que se sufrieron las dos guerras mundiales y que hasta bien entrados los sesenta no se llevaron a cabo políticas comunes que confluyeron en lo que Robert Escarpit vino en denominar en 1965 “la revolución del libro”. A ello cabe añadir los problemas internos de España y de México, en ocasiones tan graves y penosos como la Guerra civil española (1936-1939).

Es objeto del artículo analizar la creación, actividad y desarrollo de la editorial Espasa-Calpe Mexicana entre 1927 y 1975, desde su creación hasta el final de la dictadura franquista, con el objetivo general de aportar nuevos datos para la historia de la edición española y mexicana en lo que se refiere a industria, comercio y situación del sector. Como objetivo específico se pretende dar a conocer a quienes hicieron posible el proyecto, poniendo en valor su aportación a la industria y cultura de ambos países.

La metodología ha sido cualitativa y cuantitativa, analizando 561 actas del Consejo de Administración de la editorial Espasa-Calpe de Madrid fechadas entre el 28 de enero de 1926 y el 28 de octubre de 1975, de las que se ha recuperado toda la información relacionada con México. Para ello se elaboró una base de datos con cuatro campos: fecha, personas citadas, temas específicos y resumen del contenido de cada documento. La lectura se ha llevado a cabo en orden cronológico, estableciendo así una estructura en tres etapas coincidentes con los momentos clave de la historia de España y de la empresa editorial: la primera desde la fundación de Espasa-Calpe (enero de 1926) hasta la Guerra civil española de 1936-1939, la segunda desde el comienzo de la posguerra hasta la constitución de la Sociedad Espasa-Calpe Mexicana (1940-1948), y la tercera durante el desarrollo de esta empresa hasta el final de la dictadura y el comienzo de la transición democrática en España (1948-1975).

Asimismo, se ha realizado una revisión bibliográfica sobre la política del libro y la edición en España y México, y se han analizado varios catálogos de la editorial en el periodo estudiado, de especial utilidad para conocer la publicación y difusión de las obras. Con el fin de contextualizar la información, se han tenido en cuenta los acontecimientos socioculturales y políticos de ambos países, y al objeto de completar el texto se han elaborado tablas explicativas sobre ventas, disponibilidades económicas y países importadores de libros. También se han incorporado reproducciones fotográficas de catálogos y libros significativos editados en España y México, procedentes del archivo de la propia editorial y de la colección del autor.

En cuanto a las fuentes, se han empleado catálogos generales de Espasa-Calpe editados entre 1937 y 1975, más varios de Espasa-Calpe Argentina y Mexicana. De ésta última han sido de gran interés los impresos en 1950 y 1954. Sobre la edición y el libro en México se han consultado las monografías

Autores, editoriales, instituciones y libros: estudios de historia intelectual de García Diego (2015), *Entre el oficio y el beneficio: el papel del editor: práctica social, normatividad y producción editorial* de Kloss Fernández del Castillo (2007), y *Breve historia del libro en México* de Torre Villar (1999). En lo que se refiere a España las principales fuentes han sido *Historia de la edición en España*, coordinada por Martínez Martín (2009 y 2015), *La edición en España* de Sánchez Vigil (2009) y el *Bosquejo de una política del libro* de Gili Roig (1944). Sobre la relaciones entre ambos países se han consultado el *Informe con que tropieza en varios países de la América Española la exportación de nuestros libros*, editado por el Instituto Nacional del Libro Español (INLE, 1962), *Los signos de la noche. De la guerra al exilio. Historia peregrina del libro republicano entre España y México* de Santonja (2003) y *Las relaciones entre España y México, 1810-2010* de Sánchez Andrés y Pérez Herrero (2010). La información se ha contextualizado con datos históricos tomados de *Historia de México* de Fernández (2008).

ANTECEDENTES

Desde comienzos del siglo XX los editores españoles tuvieron como objetivo desarrollar el mercado americano, entonces en manos de la industria francesa que editaba en lengua española (Martínez Rus, 2000: 32). Los años veinte comenzaron a ser fructíferos en políticas culturales, empezando con la creación en 1921 de la Oficina de Relaciones Culturales Españolas (denominada Junta de Relaciones Culturales a partir de 1926), cuyo objetivo era difundir la cultura y facilitar intercambios. En 1923 se celebró el Primer Congreso Nacional del Comercio en Ultramar donde se trató el tema del libro prioritariamente y la apertura del mercado a otros productos (Martínez Rus, 2003: 293), y en 1927 se organizó en Madrid la Conferencia Nacional del Libro para analizar los problemas editoriales.

Una de las primeras editoriales en América fue Calpe, propiedad de La Papelera Española y dirigida por Nicolás Urgoiti, quien en 1922 viajó a Buenos Aires para crear una delegación y establecer relaciones con Uruguay y Chile. Calpe formaba parte del grupo Prensa Gráfica, junto con los diarios *El Sol* y *La Voz*, las revistas *Mundo Gráfico*, *Nuevo Mundo* y *La Esfera*, y la librería Casa del Libro inaugurada en 1923. En diciembre de 1925 Calpe se fusionó con Espasa, editora de la *Enciclopedia Universal Ilustrada*, y activó el comercio en México contratando como representante a Edmundo Mas.

Hasta abril de 1927 el Consejo de Administración de Espasa-Calpe no planteó la apertura de un depósito en México, al tiempo que designó

representantes en Cuba y Colombia para operaciones comerciales. Para desarrollar el proyecto encargaron un estudio al citado Mas, quien informó positivamente: “La situación económica por la que atravesó México entre 1924 y 1928 fue delicada. Las deudas externa e interna eran grandes y el sistema bancario inoperante” (Fernández, 2008: 59).

La editorial encargó al viajante Jacinto López Casares que realizara un periplo durante ocho meses hasta llegar a México, donde junto a Edmundo Mas debería realizar un informe con criterios de actuación. Se le adjudicó la cantidad de 3 200 pesetas mensuales, más el 7 % sobre el importe líquido de ventas del fondo de Calpe y de ediciones cuyos derechos tuviera cedidos la empresa. En carta enviada por López Casares desde El Salvador al Consejo de Administración el 29 de abril de 1928 (Acta n° 29, 30 de mayo de 1928: 148), se indica que el viaje se retrasó y hasta abril de 1928 no llegó a El Salvador, donde las ventas fueron reducidas pues sólo encontró un librero interesado, M. J. Méndez. Desde allí pasó a Guatemala, y en junio se trasladó en barco a Veracruz vía La Habana.

CREACIÓN Y DESARROLLO DE LA DELEGACIÓN DE MÉXICO (1928-1936)

En México López Casares consiguió pedidos de dos clientes de relevancia: Andrés Botas e Hijos y Sucesores de Herrero Hermanos, e informó de la imposibilidad de visitar Guadalajara, Monterrey y Puebla debido a la Guerra cristera (1927-1929) durante la presidencia de Plutarco Elías Calles, que enfrentó a la Iglesia con el Estado. Los ingresos por venta de fondos ajenos en distribución superaron hasta casi tres veces los de los libros editados por la empresa (*Tabla 1*).

| Cientes | Fondo Espasa-Calpe | Fondo ajeno | Enciclopedias | Total pesetas |
|------------------|--------------------|-------------|---------------|---------------|
| Herrero Hermanos | 3 347.50 | 5 446.15 | 1 090 | 9 883.65 |
| Andrés Botas | 1 565.05 | 6 783.25 | | 8 348.30 |
| TOTAL | 4 912.55 | 12 229.40 | 1 090 | 18 231.95 |

Tabla 1. Ventas de Jacinto López Casares en México en 1928

En abril de 1929 el Consejo de Administración decidió transformar el depósito de México en delegación, con funciones similares a Buenos Aires, poniendo al frente a López Casares temporalmente hasta que llegara desde España el primer apoderado: José Jiménez Nicolau. Los sueldos anuales fueron respectivamente de 22 000 y 10 000 pesetas, más un porcentaje por ventas.

En España continuaban las iniciativas para apoyar la expansión del libro en América. En la primavera de 1929 se celebró el II Congreso Nacional del Comercio Español en Ultramar, donde se planteó que el Estado devolviera a los editores los derechos arancelarios por exportación, y que se anulara el recargo del 100 % sobre la contribución que se pagaba. El 30 de octubre se constituyó el Consorcio Nacional de Editores Exportadores en Barcelona, cuyo objetivo fue difundir el libro español en América. Ambos eventos, junto con la creación el 14 de marzo de 1930 del Sindicato Español del Libro Español (SELE), influirían en la política comercial de Espasa-Calpe en México.

En diciembre de 1929 se procedió al alquiler de un local en el número 6 de la Avenida Isabel la Católica por un importe mensual de 175 pesos, con una inversión de 2 000 pesos en acondicionamiento. La actividad dio paso a las tareas editoriales y en junio de 1930 se presentó la primera obra: *Plasmogenia*, de Alfonso L. Herrera, que ya había sido editada en 1924 por Herrero Hermanos y que no fue aceptada por considerarla especializada y con poca salida.

En junio de 1930 llegó a México Francisco Rubio para relevar a López Casares y dirigir la sección comercial, mientras que Jiménez Nicolau se dedicó a tareas administrativas. Las gestiones dieron excelentes resultados: se importó una remesa de libros por 40 000 pesetas y aumentaron las ventas, que en 1930 alcanzaron una media mensual de 10 000 pesetas. Con esas perspectivas económicas, el Consejo de Administración autorizó un viaje por el norte del país en busca de clientes. Otra de las referencias a la labor editorial fue la propuesta de publicación en noviembre de 1930 de la obra *Mala yerba*, de Mariano Azuela, de quien Espasa ya había editado *Los de abajo*.

En marzo de 1931 consiguieron las mejores ventas del año, sumando 30 000 pesos a pesar de la situación del país, y en junio se realizó un estudio para determinar si el servicio de libros a Estados Unidos podría realizarse desde Cuba o México con el fin de abaratar el transporte. Finalmente se optó por seguir haciendo los envíos desde Madrid, salvo en casos solicitados expresamente por los clientes. Dato de interés es el sueldo mensual de los apoderados, fijado en 650 pesos.

Como se ha indicado, en los dos primeros semestres de 1930 y 1931 las ventas subieron considerablemente, pasando de 60 877 pesetas a 412 053, con un crecimiento del 576.85 %. El aumento del saldo líquido por ventas de Espasa-Calpe en los primeros semestres de ese periodo también fue considerable (*Tabla 2*).

| Año | Espasa-Calpe Madrid | | Casa del Libro | Espasa-Calpe Barcelona | | Total |
|------|---------------------|-------------|----------------|------------------------|-------------|--------------|
| | Fondo propio | Fondo ajeno | Fondos varios | Fondo propio | Fondo ajeno | |
| 1930 | 456 138.63 | 414 750.80 | 392 400.93 | 58 953.12 | 41 175.75 | 1 363 419.23 |
| 1931 | 925 449.18 | 336 836.33 | 397 930.10 | 58 436.50 | 35 798.90 | 1 754 451.01 |

Tabla 2. Ventas en Espasa-Calpe (España) en los primeros semestres de 1930-1931

Consecuencia de los buenos resultados fue el aumento de sueldo a los apoderados, que en octubre de 1931 pasó a 838.50 pesos mensuales. En ello influyó el cambio de precio en la moneda efectuado por las autoridades mexicanas con arreglo al nuevo patrón de plata. En agosto de 1932 el Consejo aprobó que los apoderados se encargaran de gestionar la delegación en Cuba, elaborando informes económicos y la viabilidad para determinar acciones futuras.

En 1934 se fundó en Madrid la Agrupación de Editores Españoles con el objetivo prioritario de defender los mercados americanos. Rafael Giménez Siles, secretario de la organización, propuso depósitos comunes en varios países, iniciativa que Espasa-Calpe y otras editoriales como Aguilar llevaban realizando ya por tres lustros (Giménez Siles, 1978: 18). Durante aquel año la prioridad en México fue comercial, y se recibieron propuestas de publicación como la biografía de Francisco Morazán por Hernán Robleto en julio de 1934 para la colección “Vidas Españolas e Hispanoamericanas”, que fue desestimada, o la del general López de Santa Anna de Rafael F. Muñoz, que fue informada por el historiador Melchor Fernández Almagro por su interés histórico y literario, y se publicó con una tirada de 3 000 ejemplares en 1936 en la colección citada con el título *Santa Anna, el que todo lo ganó y todo lo perdió*.

En abril de 1935 se creó en España el Instituto del Libro Español para impulsar el comercio en América mediante nuevos depósitos. De abril a junio se recogieron en las Actas del Consejo de Administración los datos económicos, indicando las disponibilidades, entre ellas las cantidades recibidas desde las delegaciones de Buenos Aires y México. De la primera se recibieron 292 204.34 pesetas y de la segunda 84 245.70 pesetas, 60 000 francos y 200 libras (Tabla 3). A finales de 1935 se produjo un descenso en los envíos de divisas a España (35 272.70 pesetas) que preocupó al Consejo. La justificación de los apoderados fue la situación social y la falta de suministro de obras para su venta a plazos, es decir, de libros en gran formato de alto precio, lo que hizo que durante el segundo semestre se perdieran unas 100 000 pesetas de ingresos (Acta n° 119, 28 de noviembre de 1935).

En el primer trimestre de 1936 Espasa-Calpe Mexicana envió a la casa matriz 501 264.51 pesetas, y presentó un informe de actuaciones. En abril, el apoderado Jiménez Nicolau solicitó autorización para trasladarse a España en verano debido a una afección reumática, pero la respuesta fue negativa porque según su contrato no le correspondían vacaciones. Este detalle sería anecdótico si no se hubiera producido el golpe de Estado que dio lugar a la Guerra civil en julio de 1936 y que no le permitió regresar hasta el otoño de 1939.



Figura 1. Publicidad de Espasa-Calpe con las direcciones de Barcelona, Buenos Aires y México (Avenida Isabel la Católica, 6). (Catálogo de la Cuarta Feria del Libro de Madrid, 1936)

| Entidad | Abril | Mayo | Junio | Julio |
|-----------------------------|--------------|------------|------------|------------|
| Banco de Bilbao | 1 650 748.94 | 823 902.07 | 691 553.27 | 761 798.58 |
| Banco de España | 3 579.86 | 3 381.93 | | |
| Banco Popular | 393.55 | 494.95 | | |
| Banco Hispano Americano | 11 740.00 | 13 887.93 | | |
| Banco de Vizcaya | 1 293.11 | 2 077.11 | | |
| Banco Anglo South Americano | 3 983.25 | 5 862.75 | | |
| International Banking Corp. | 59 954.67 | 163 910.97 | 169 015.67 | 111 418.80 |

| | | | | |
|----------------------------------------------------|---------------------|------------|-------------------|--------------------------------------------------|
| Banco Mercantil e Industrial | 520.45 | 520.45 | | |
| Otros Bancos | | | 18 094.69 | 14 473.90 |
| Caja de Madrid | 52 235.45 | 57 267.55 | 60 756.34 | 68 193.90 |
| Caja de Barcelona | 7 755.82 | 3 495.53 | 2 185.59 | 4 045.01 |
| Caja de Casa del Libro | 300.00 | 300.00 | 300.00 | 300.00 |
| TOTAL | 1 792 505.10 | | 941 905.56 | 960 230.19 |
| Banco de Bilbao París (Francos) | 10 999.46 | 200 575.26 | 23 333.46 | 81 569.36 |
| Banco de Bilbao Londres (Libras) | 25-2-6 | 16-3-3 | 162-13-2 | 349-7-2 |
| The Anglo South American Trust N. York (Dólares) | 100.24 | 220.16 | 194.88 | 406.75 |
| Deutsche Bank Leipzig (Marcos) | 221.91 | 4 805.16 | 1 590.63 | 72.08 |
| Cartera | 19 850.05 | 19 575.64 | 15 575.00 | 9 307.14 |
| Recibido de Buenos Aires | 485.67 | 147 936.55 | 51 388.12 | 62 500.00 |
| Recibido de México | 40 200.17 | 39 730.6 | 2 422.86 | Ptas 1 892.07 Fran 60 000.00 Libras 200.00 |
| Valores al cobro | 174 700.79 | | 136 531.05 | 164 953.36 |
| Importe del Dividendo Activo | | 960 000.0 | | |
| Pagado hasta hoy 29.717 cupones | | 822 566.56 | | |
| Cantidades bloqueadas por envío directo en pesetas | | | 100 311.03 | 96 897.97 |
| Cantidades bloqueadas. Monedas varias: | | | | |
| - Chile (Pesos argentinos) | | | 673.45 | 673.45 |
| - Chile (Pesos chilenos) | | | 2 450.98 | 2 450.98 |
| - Uruguay (Pesos oro uruguayo) | | | 31 995.61 | 31 995.61 |
| -Uruguay (Bonos oro Francos) | | | 85 500.00 | 85 500.00 |

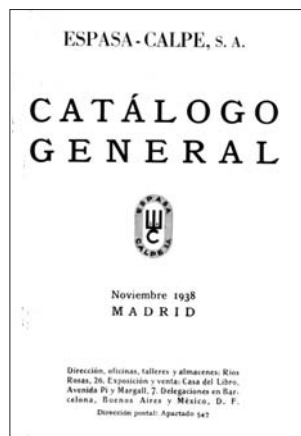
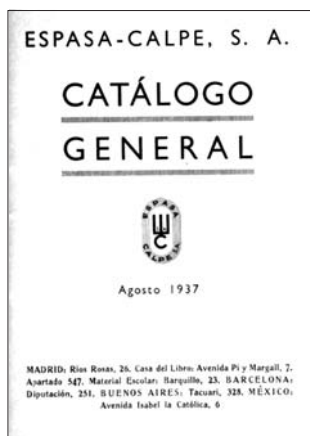
Tabla 3. Disponibilidades de Espasa-Calpe entre abril y julio de 1935
(Fuente: Actas del Consejo de Administración. Elaboración propia)

DE LA GUERRA CIVIL ESPAÑOLA A LA CONSTITUCIÓN DE LA SOCIEDAD ESPASA-CALPE MEXICANA (1936-1948)

Durante la Guerra civil (1936-1939) los locales de Espasa-Calpe en la calle Ríos Rosas de Madrid fueron incautados para la impresión de publicaciones oficiales. El Consejo de Administración se trasladó a San Sebastián, controlada por los golpistas, y desde allí se organizó el negocio, encargando a Manuel Olarra Garmendia, cuñado de José Jiménez Nicolau, que viajara a Buenos Aires y transformara aquella delegación en sociedad anónima debido a la gravedad de los acontecimientos en España.

El 29 de marzo de 1937 se constituyó la sociedad Espasa-Calpe Argentina y la delegación de México pasó a depender de Buenos Aires el 26 de julio mediante una operación comercial que costó a la empresa argentina 50 000 pesos al contado. Fueron nombrados tres directores: Manuel Olarra, Francisco Rubio, y Gayoso, acción que se justificó para “evitar una intervención extraña en la casa mexicana, o que las disponibilidades en la misma pudieran ser objeto de un despojo” (Olarra Jiménez, 2003: 29).

Durante la contienda las relaciones con América quedaron suspendidas. A comienzos de 1937 Gonzalo Losada y Guillermo de Torre crearon en Buenos Aires la colección Austral, diseñada por Attilio Rossi, donde se publicarían las obras de autores mexicanos, entre ellos Amado Nervo, cuya *Amada inmóvil* llevó el número 32 y salió el 15 de enero de 1938, y Juan Ruiz de Alarcón, del que se lanzaron *La verdad sospechosa* y *Los pechos privilegiados* en el número 68, editado el 15 de mayo de 1939.



Figuras 2 y 3. Catálogos generales editados en Argentina y México durante la Guerra civil (1937-1938)

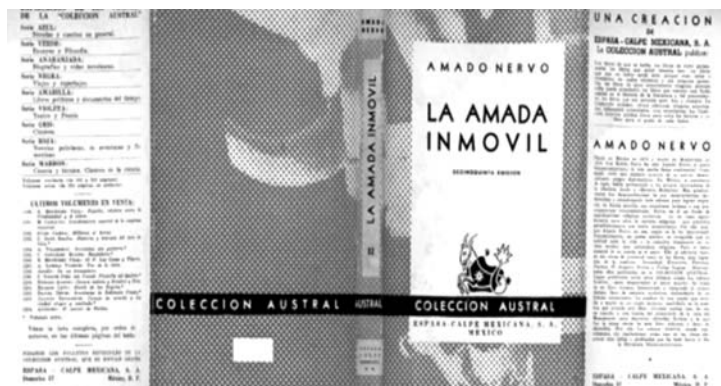


Figura 4. Amado Nervo. *La amada inmóvil*, Colección Austral, nº 32. Sobrecubierta de la decimoquinta edición, 1961

Como es sabido, el gobierno del presidente Cárdenas no reconoció a los golpistas españoles y acogió a los exiliados, por lo que se interrumpieron las relaciones diplomáticas. La primera Acta del Consejo de Administración después de la guerra está fechada el 31 de enero de 1940, y la primera referencia a México no aparece sino hasta el 30 de mayo, indicando que se reanudaba el servicio postal. Las ventas se controlaron desde Buenos Aires, si bien los pedidos y correspondencia se hicieron directamente con Madrid.

A partir de 1941 se atendieron los temas internos, recursos humanos y sueldos, que subieron a 210 pesos mensuales para los trabajadores. La actividad se reactivó con una gran remesa de libros que habría de llegar directamente al puerto de Veracruz una vez que se autorizara la importación. El proceso no fue fácil porque el Tesoro de EEUU prohibió la exportación y libre tráfico de efectos a los países que apoyaron a Alemania e Italia durante la Guerra Mundial: Suiza, Suecia, España, Portugal y Marruecos. En la década de los cuarenta la edición en México se impulsó con la apertura de numerosas empresas fundadas por los exiliados. En el catálogo de la exposición “Obra impresa del exilio español en México 1939/1979”, celebrada entre septiembre y noviembre de 1979, se relacionan más de cuarenta empresas (Ateneo Español de México, 1979), entre ellas Edición y Distribución Iberoamericana de Publicaciones S. A. (Ediapsa), Unión Tipográfica Editorial Hispano Americana (Uteha), Grijalbo, Era, Joaquín Mortiz, Leyenda, Atlántida, Séneca o Nuestro Pueblo (Armendáriz Sánchez y Ordóñez Alonso, 1999; Fernández Moya, 2008, 2009; Ixba Alejos, 2013).

Entre los intelectuales destacó Rafael Giménez Siles, que organizó el 12 de junio de 1944, junto a Martín Luis Guzmán, la Asociación de Libreros

y Editores Mexicanos, entidad que consiguió del presidente Manuel Ávila Camacho la exención de impuestos a la edición. En agradecimiento se creó un premio literario y científico con su nombre. Giménez Siles activó en diciembre de aquel año la primera feria del libro con el apoyo de la Sección de Editores de Libros de la Cámara Nacional de la Industria de la Transformación, inspirándose en la feria española inaugurada en 1934; concurren 44 empresas (Giménez Siles, 1978: 32). La segunda feria se celebraría 23 años después, el 22 de mayo de 1970, también con 44 participantes. Estos eventos coincidieron con la publicación de la conferencia de Gustavo Gili *Bosquejo de una política del libro*, que pronunció en Barcelona para solicitar al gobierno el apoyo al mercado en América.

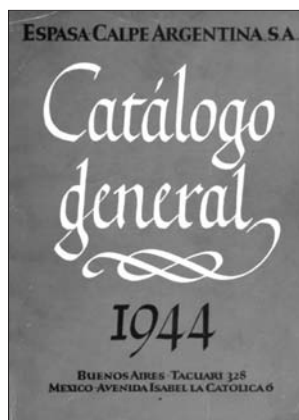


Figura 5. Segunda edición de las *Poesías Completas* de Pérez de Ayala. Impresa en Buenos Aires en 1942 y comercializada en Argentina, México y España

Un problema fue la piratería, denunciada en octubre de 1943 en un informe que México mandó a Madrid y Buenos Aires. Muchos libros protegidos por la Ley de Propiedad Intelectual española durante 80 años en México sólo lo estaban por 30 de acuerdo a la norma. Gran parte de los libros piratas que circulaban en México se imprimían en Argentina.

En el verano de 1944 Espasa-Calpe Mexicana se instaló en Donceles 57, edificio de amplia fachada, tres pisos y cuatro plantas, alquilado por 2 300 pesos mensuales. El envío a España de ingresos se canalizó desde Buenos Aires y el informe del Consejo de Administración de 21 de diciembre de 1944

indica que se había enviado medio millón de pesetas y que la recaudación prevista para 1945 era de otro medio millón.



Figuras 6 y 7. Catálogos de 1944 y 1945 donde se observa el cambio de dirección de Espasa-Calpe Mexicana



Figura 8. Edificio de Espasa-Calpe en el número 57 de la calle de Donceles, inaugurado el verano de 1944

Al término de la Segunda Guerra Mundial, España fue expulsada de Naciones Unidas por su apoyo al fascismo, y las exportaciones de libros a México continuaron desde Buenos Aires. Sólo aquellas empresas que tenían delegaciones, entre ellas Espasa-Calpe, pudieron continuar con negocios, y en julio de 1945 Olarra viajó a México para fiscalizar la delegación y recaudar efectivo. Espasa-Calpe Mexicana comercializó las ediciones que recibía de Buenos Aires y Madrid en el mercado de Centroamérica, EEUU, Ecuador, Colombia y Venezuela. Su tarea editora fue mínima, con obras de Rafael Muñoz (*Se llevaron el cañón para Bachimba*, *Vámonos con Pancho Villa*, *El águila y la serpiente*), José Vasconcelos (*Breve historia de México*) y Salvador Novo (*Nueva Grandeza Mexicana*).

En 1946 se aprobó en España la Ley de Protección del Libro, con el objetivo de afrontar la carestía y conseguir papel de calidad, pero los resultados fueron malos. Ese año se produjeron tensiones políticas entre México y Argentina, reflejadas en la edición. En el artículo sin firma titulado “El auge del libro” de la revista *La Propiedad* (14 de septiembre) se informaba sobre el cambio positivo del momento:

La lucha editorial entablada por México en contra de Argentina y España apenas se inicia y no podemos vaticinar un éxito tan rápido y definitivo como el que se apuntó nuestro cine en su competencia contra el cine argentino y español. Solo en 1942 Argentina editó 15 millones de libros, 10 de los cuales exportó y 5 destinó a consumo interno. Entre los importadores México figuró en primer lugar adquiriendo dos millones de ejemplares impresos en Argentina, la mitad de los cuales están representados por ediciones de Espasa-Calpe [...]. México está ganando mercados exteriores al tiempo que ensancha su campo nacional. Cuba acaba de comprar el remanente de libros mexicanos de la pasada feria en 25.000 dólares y sabemos que los libreros argentinos han contratado un buen porcentaje de la producción de las más importantes casas mexicanas [...] Todo, pues, concurre a beneficiar el auge del libro mexicano: protección gubernamental, fenómenos económicos de esta postguerra, falta de transportes en el sur de América y una creciente sed de lectura en el pueblo.

En enero de 1947 se ordenó desde Madrid una valoración sobre autores de más venta, libros de éxito, precios y relación con el costo en España, papeles de impresión, posibilidades del mercado papeler y materias primas necesarias. En junio, el consejero comercial de España en Washington, Luis García Guijarro, trató de que el gobierno mexicano reconociera a Franco pero no lo logró, aunque se reabrieron varios puertos y se restablecieron relaciones económicas (Sánchez Andrés y Pérez Herrero, 2010: 37). Olarra regresó a México en noviembre para estudiar la independencia de la delegación frente a Buenos Aires, tomando como ejemplo la editorial Labor que había realizado una transformación similar.



Figura 9. *El reloj del señor Humphrey*, de “Carlos” Dickens.
Colección Austral n° 658, 2ª edición, 5 de junio de 1947

A comienzos de 1948 los apoderados se quejaron de las remesas clandestinas a bajo precio importadas a los gremios de edición y librería. Se vendieron a 15 centavos por peseta cuando el cambio oficial era de 46, situación que estaba vinculada a otras actividades: “Las importaciones se hacen movilizandando pesetas en España para efectuar así una exportación clandestina de nuestra moneda a base, principalmente, de los fondos obtenidos por la exhibición de películas mexicanas en España, y también por una importante remesa de medias nylon que, al parecer, han sido introducidas en España” (Acta n° 231, 29 de abril de 1948). Además de desvirtuar el mercado, la consecuencia fue un arancel al libro español, ya que las importaciones se consideraban *dumping*, práctica consistente en la venta de un producto por debajo de su precio, incluso de su coste de producción, con el fin de eliminar competencia y apoderarse del mercado.

En abril de 1948 se reestructuró la delegación para constituir una nueva sociedad. Manuel Olarra, director en Argentina, redujo el personal un 27 %, actualizando los pagos pendientes a librerías y auditando compras y ventas. A finales de mayo de 1948 se firmó la escritura de constitución de Espasa-Calpe Mexicana y comenzó una nueva etapa.

SOCIEDAD ESPASA-CALPE MEXICANA

El 27 de mayo de 1948 se firmó la escritura de constitución de la sociedad y el 31 se redactó el acta del primer Consejo, nombrando presidente a José Izu, vicepresidente y secretario a Manuel Olarra, vocales a Antonio de Sabatés, Gervasio Collar y Manuel Quintana, y comisario a José Manuel Escobedo. Los apoderados, elegidos temporalmente mientras eran nombrados los nuevos desde España, fueron Olarra, que también ejerció de director, Francisco Rubio, Manuel G. Escobedo, y el citado Gayoso.

Primera etapa, 1948-1960

Las tensas relaciones entre México y EEUU agudizaron la crisis. El informe económico del segundo trimestre fue negativo porque el peso bajó su cotización de 6 a 4.85 con respecto al dólar. El comercio operaba con un alza del 30 % en artículos importados y un 15 % en nacionales, por lo que los editores tuvieron que elevar precios. En el caso de Aguilar y Espasa-Calpe se aplicó entre un 20 y un 30 % de subida, que supuso una importante pérdida al convertir a dólares los créditos por venta a plazos. En septiembre el Consejo de Administración nombró a los dos nuevos apoderados: José María Mendieta, procedente de Espasa de Barcelona, y Lorenzo Salgado de Madrid, ambos con un sueldo de 1 000 pesos mensuales más beneficios.

Las noticias que llegaron de México en enero de 1949 fueron las trabas del gobierno para la importación, especialmente de libros procedentes de España y Argentina, y la prohibición de entrar en el país los apoderados. Esto obligó a solicitar del director de Comercio y Política Arancelaria español, José Ibáñez Martín, que solventara las dificultades en la exportación a países americanos y especialmente a México.

Los dirigentes de Espasa en Buenos Aires y México pidieron a Madrid que la comercialización se realizara exclusivamente desde aquellas sociedades, pero se consideró inviable por las normativas de los Institutos del Libro y de la Moneda, y sobre todo porque “Los intereses de Espasa-Calpe son más y el comercio de exportación de la editorial debe supeditarse en cada momento a la conveniencia comercial que se deriva de la actual situación del comercio mundial” (Acta nº 242, 31 de marzo de 1949).

Olarra comunicó en abril el descenso de ventas por la crisis política en México y por el impuesto del 40 % sobre las importaciones. En cuanto a las publicaciones notificó la firma de un contrato para la edición de un tomo de poesías de Torres Bodet, ministro de Relaciones Exteriores y director general de Unesco, y de los libros *Técnica Quirúrgica Operatoria* y *Dermatología*, de los doctores Segovia y Bejarano.

En mayo de 1949 se acordó pagar las cantidades adeudadas a los editores mexicanos por sus exportaciones a España a cambio de que se retirara el impuesto del 40 % a los libros españoles. La intervención de Francisco Rubio ante el ministro español Ibáñez Martín fue decisiva para el desbloqueo. Los pagos a México no se realizaron con inmediatez, ya que el Instituto de la Moneda carecía de divisas en dólares, y tampoco se regularizó la importación debido al mismo problema, lo que hizo que numerosos envíos se paralizaran en la aduana de Madrid a la espera del dinero para retirarlos, situación que se normalizó en diciembre de 1949.

En enero de 1950 la empresa contaba con 34 empleados. La producción resultaba más barata en España que en México por el valor de la moneda, de ahí que se subieran los precios de la colección Austral en 50 céntimos, pasando a 3.50 pesos el volumen sencillo y a 5 el extra. El 11 de marzo de 1950 se celebró una junta general en México para el tema económico, fondos y remesas. Durante el primer cuatrimestre de ese año se observó falta de libros españoles en el mercado y la persistencia de la entrada de ejemplares piratas a mitad de precio, que influyeron concretamente en dos obras de Espasa: *Diccionario Manual* de la Real Academia de la Lengua y *Diagnóstico Etiológico* de Gregorio Marañón.

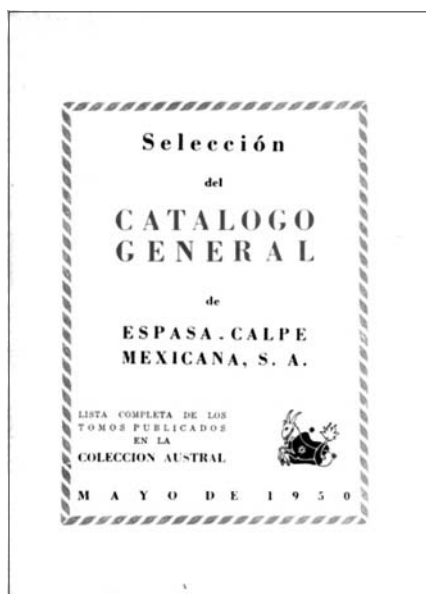


Figura 10. Catálogo general de Espasa-Calpe Mexicana publicado en mayo de 1950

Otro problema fue la obligación de pagar en el momento de la retirada de mercancías de la aduana por imperativo del Banco Central de México, lo que suponía solicitar anticipos con intereses. Debido a esto desde abril quedaron retenidas 130 cajas de libros y 1 970 paquetes, retirados a medida de las necesidades mediante pagos fraccionados.

Las tensiones políticas entre México y Argentina habían paralizado la actividad durante dos años, que se reanudó en julio de 1950 con reimpressiones de obras agotadas, fundamentalmente de la colección Austral, de cuyas tiradas la mitad de ejemplares eran remitidos a Buenos Aires. En el catálogo impreso en mayo de ese año se incluyó la colección “Libros de Historia Mexicana” con 11 títulos.

El gobierno mexicano dificultó la importación de libros españoles y por temor a otro impuesto se procedió a enviar fondos desde Madrid, sobre todo tomos de la *Enciclopedia Espasa* y de grandes obras, porque “del resto del catálogo se encuentran surtidos, ya que en esas remesas que acaban de retirar se forzó mucho el envío de todos los títulos de que carecían o que estaban a punto de que se les agotaran” (Acta 258, 22 de julio de 1950). Los problemas se agudizaron y un cargamento preparado en el puerto de Bilbao fue retenido al conocer que el gobierno mexicano exigía licencia de importación. Finalmente el barco salió en diciembre hacia Veracruz, y al mismo tiempo se realizó una exportación por correo, si bien con el temor de que fuera retenida en la aduana al no aceptar el cambio libre español para las importaciones y exportaciones de libros fijado por el Instituto de la Moneda.

En marzo de 1951 Olarra realizó una revisión de la delegación y convocó una junta general. La situación en México era buena pero no en Buenos Aires debido a la escasez de papel, hasta el punto de que varios libros maquetados quedaron sin imprimir. En junio se activó el convenio de importación entre México y España, y en marzo de 1952 el apoderado Francisco Rubio enfermó y fue trasladado a una clínica de Nueva York donde falleció dos meses después. Mendieta y Salgado continuaron su labor.

Entre mayo de 1953 y julio de 1954 se generaron deudas en Espasa Argentina por valor de 2 500 000 de pesos. La reacción del Consejo de Administración de Madrid para corregir la gestión fue intensificar las ventas y restringir publicaciones. El Consejo aprobó que las filiales mexicana y argentina obraran como delegaciones de la sociedad de Madrid, procurando que se editara en España salvo excepciones (Acta nº 307, 30 de agosto de 1954). En agosto de 1954 el apoderado Lorenzo Salgado, de vacaciones en España, expuso que la situación de la sociedad era excelente y perfectamente saneada.

Sin embargo, la situación empeoró en Buenos Aires y en octubre de 1954 la deuda con Espasa México aumentó en 45 115 pesos, correspondiente a libros editados en México y enviados a Argentina. El Consejo decidió entonces ceder protagonismo a la editorial mexicana:

En nuestro deseo delimitar la actuación de Espasa Calpe Argentina, se han pasado reediciones de Austral y de obras varias a México, pero si como consecuencia de ello Argentina cursa sus pedidos por las obras cuya reimpresión ha llevado a cabo México, y no satisface su importe, no se consigue nuestro propósito, antes al contrario, se coloca a Argentina en una situación privilegiada, puesto que no tiene que desembolsar cantidades por tales reediciones. En su consecuencia el Consejo acordó se escriba a Espasa Calpe Mexicana y Argentina, en el sentido de que, en adelante, se abstendría México de remitir libros a Argentina, a no ser que ésta haya satisfecho con anterioridad una cantidad igual a los libros que solicite, a fin de que no aumente la deuda que ya tiene contraída. (Acta n° 333, 20 de diciembre de 1954)

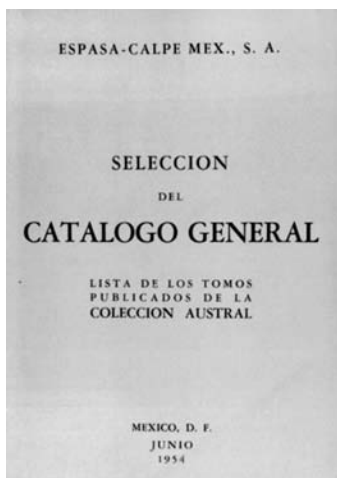


Figura 11. Catálogo de junio de 1954

La comercialización en Buenos Aires mejoró en el último trimestre de 1954. En cuanto a México, el balance se consideró satisfactorio y se autorizó el viaje de Mendieta a España para que informara. A su regreso consiguió un aumento de sueldo del 10 % para los empleados, pasando la nómina mensual de la sociedad de 11 975 a 13 337.50 pesos. En relación a los apoderados, la devaluación del peso mexicano de 5.20 a 3.10 pesetas les había supuesto una disminución en sus ingresos de 825 pesetas por cada 100 000 pesos mexicanos de venta líquida. El Consejo acordó en febrero de 1955 acceder a una mejora de 2 000 pesos, con una comisión por ventas de entre el 2.25 y el 2.50 %.

En 1956, año en que España reingresó en la ONU por mediación de EEUU, Rafael Giménez Siles fundó la Agrupación de Editores Mexicanos con la participación de 13 entidades, presidida por él mismo y por Luis Novaro (Giménez Siles, 1978: 33). La economía se reactivó y en abril los pedidos a Madrid superaron los 2 000, entre ellos 250 colecciones del *Diccionario Enciclopédico Abreviado* de Espasa en 7 tomos. Sin embargo, el Consejo había restringido al máximo las ediciones en las dos filiales durante 1955 y 1956 con una sola novedad, más 42 títulos de Austral (25 en 1955 y 17 en 1956).

En mayo de 1957 la nómina mensual total pasó de 12 985 pesos a 14 930, es decir, un 15 % más, y a los apoderados se les concedió una retribución fija de 500 pesos mensuales. México saldó con Argentina una parte de la deuda contraída, abonando 7 757 pesos. Las ventas experimentaron una baja debido a la convocatoria de elecciones presidenciales y a la carencia de nuevas obras para venta a plazos. En consecuencia, el Consejo resolvió que se editaran los libros de autores mexicanos de interés.

A comienzos de 1958 los editores españoles denunciaron el convenio bibliográfico con México y las exportaciones de libros quedaron suspendidas temporalmente. Esto afectó las ventas puesto que se acusó la carencia de títulos, como informó el apoderado Mendieta durante su estancia en Bilbao. También fue un contratiempo la decisión de la editorial Aguilar, que retiró la exclusiva de comercialización de sus libros a las dos filiales americanas. Afortunadamente, el convenio se prorrogó hasta enero de 1959 y se reanudó el envío. Con el fin de recuperar mercado se ordenó desde Madrid que se reimprimieran títulos de Austral de autores argentinos y mexicanos, y que se revisaran los libros recomendados como texto en las escuelas, así como los de autores españoles que por circunstancias excepcionales no podían editarse en España, en clara alusión a cuestiones de derechos y a temas políticos.

Las ventas de 1959 arrojaron cifras inferiores a las de 1958, pero no así los cobros. A la reducción de ingresos se sumó la subida de sueldos por disposición gubernamental, pasando el salario mínimo de 12.50 pesos diarios a 14.50, si bien la editorial decidió abonar 14.75 para que los trabajadores se acostumbraran a pagar su parte del seguro.

Segunda etapa, 1961-1975

En 1960 la incertidumbre fue remitiendo. Los ingresos subieron debido a la diferencia en el cambio de moneda, que pasó de 3.10 pesos a 4.80. A finales de ese año la comisión para los apoderados pasó del 2.5 al 3 %, sin alteración en los sueldos, mientras que el de Olarra aumentó de 11 260 pesos a 15 000. Se observa la desigualdad de trato a los dirigentes de ambas filiales, sobre

todo tras la advertencia del Consejo de Espasa-Calpe a Olarra advirtiéndole que “debía procurar por todos los medios enviar a España, aunque sólo sea en parte, los beneficios que obtenga, ya que ese fue el motivo y fundamento de establecer las delegaciones y no limitarse simplemente al pago de las facturas de los libros que le son remitidos” (Acta n° 391, 30 de agosto de 1961). El año 1961 fue de “normalidad” según manifestó Salgado al Consejo. Desde México se importaron varios títulos de Austral que fueron importante fuente de ingresos.



Figura 12. *La verdad sospechosa y Los pechos privilegiados*, de Juan Ruiz de Alarcón. Colección Austral, n° 68, 7ª edición, 4 de septiembre de 1959



Figura 13. Sellos de Espasa-Calpe de Madrid en ejemplares de la colección Austral importados de México en 1961

El informe del Instituto Nacional del Libro Español de 1962 señalaba una fuerte competencia y la exclusión de España en los proyectos mexicanos de coedición. El precio fue un problema debido a los costos de fabricación en España y a los márgenes de distribución, y en cuanto a la política arancelaria española el gravamen sobre el papel complicaba la situación. El transporte se realizaba en barco con una media de recepción de dos a tres meses, por ello era necesario hacer grandes pedidos cuyos pagos se realizaban mediante préstamos (INLE, 1962). En el primer semestre se produjo un descenso en las ventas y a finales de año el mercado era estacionario.



Figura 14. Aristóteles. *Moral a Nicómaco*. Cubierta y solapa. Colección Austral, nº 318, 5ª edición, 1962

En 1963 las actas del Consejo no reflejaron incidencias notables y el informe corrió a cargo del omnipresente Olarra. En 1964 se celebró en la capital mexicana el Primer Congreso Iberoamericano de Asociaciones y Cámaras del Libro del 18 al 23 de mayo. El representante español fue Carlos Robles Piquer, presidente del INLE, y se tomaron los siguientes acuerdos: libre exportación por vía aérea, mayor divulgación y distribución de los libros, e instauración del Día del Libro Iberoamericano el 23 de abril.

La actividad de 1965 fue regular, y a mediados de 1966 se realizó un pedido a Madrid por importe de 4 950 000 pesetas. Durante 1967 y 1968 en los informes llegados desde México se empleó el término “normalidad”. En dos años se estudió la unificación de instalaciones en el local de Avenida de la Morena, dejando las oficinas de Donceles y los almacenes y talleres de las calles Sidar y Rovirosa. El cambio se llevó a cabo en septiembre de 1969, al tiempo que la reestructuración de la empresa, fijando para los apoderados un sueldo de 44 500 pesos anuales y una comisión por ventas del 3 al 4 %.

México fue, tras Argentina, el principal importador de libros españoles entre 1968 y 1970, y el primero en 1971. Venezuela, Chile y Colombia fueron los otros grandes compradores, y en menor medida Puerto Rico, Uruguay, Estados Unidos y Brasil. Entre los países europeos, sólo Francia y Reino Unido se interesaron por la edición española (*Tabla 4*).

| ORDEN | 1968 | 1969 | 1970 | 1971 |
|-------|---------------|---------------|---------------|---------------|
| 1 | Argentina | Argentina | Argentina | México |
| 2 | México | México | México | Argentina |
| 3 | Venezuela | Venezuela | Venezuela | Venezuela |
| 4 | Chile | Chile | Colombia | Colombia |
| 5 | Colombia | Colombia | Chile | Francia |
| 6 | Perú | Francia | Francia | Chile |
| 7 | EE.UU | Cuba | Reino Unido | Reino Unido |
| 8 | Puerto Rico | EE.UU | Puerto Rico | Uruguay |
| 9 | Brasil | Perú | EE.UU | EE.UU |
| 10 | Uruguay | Puerto Rico | Perú | Perú |

Tabla 4. Países con mayor compra de libros españoles entre 1968 y 1971

Fuente: Cendán Pazos (1972: 153)

En agosto de 1970 se realizó una operación de envergadura firmada tras un viaje urgente del apoderado Salgado para estudiar con Ernesto Antón, director general de Espasa-Calpe, la venta a los almacenes Aurrerá de 15 000 ejemplares del *Diccionario Enciclopédico Abreviado*, 5 000 de ellos encuadernados en cartón y 10 000 en tela. En los dos años siguientes la delegación acumuló deudas con la casa central por el impago de las remesas. En mayo de 1972 se tomó la decisión de cortar el suministro de libros a México si el tema no se solucionaba, con gran preocupación del Consejo. En diciembre el apoderado Salgado fue relevado del cargo tras 24 años de servicio y se le autorizó a regresar a España. Fue sustituido por Francisco Cruz Rubio en enero de 1973, quien trabajaba en la Secretaría de Ediciones de Madrid, con 44 500 pesos de sueldo anuales más una comisión del 3 % sobre ventas. Sus primeras tareas fueron supervisar la auditoría de la empresa Price Waterhouse y regularizar el capital de la sociedad, que pasó a un millón de pesos después de la gestión realizada ante la Dirección General de Transacciones Exteriores del Ministerio de Comercio, aprobada en Consejo de Ministros el 23 de noviembre de 1973.

Fue este un momento de expansión derivado del decreto de 1971 por el que el Gobierno español actualizaba la Ley de Protección del Libro de 1946, con objeto de facilitar la importación de papel, bajar el gravamen fiscal por exportación y mejorar las condiciones de financiación. Con estas medidas aumentó la industria editorial española, pasando del trigésimo puesto que ocupaba en el *ranking* editorial mundial en 1949 al quinto lugar en 1974.

El desequilibrio en la balanza de exportaciones con América llevó al gobierno mexicano a tomar medidas, y a instancias de la Cámara Nacional de la Industria Editorial Mexicana se limitó la importación a un máximo de 1 500 ejemplares por título, de forma que las empresas extranjeras se vieran obligadas a editar en el país y creciera así el empleo en el sector de artes gráficas.

Las actas del Consejo de Administración de Espasa-Calpe de Madrid en 1974 y 1975 apenas hacen referencia a México. En julio de este último año se produjo un incremento de las ventas del 38 %, pero se interrumpió al cerrarse las fronteras a las importaciones. Las previsiones de importación desde México para 1976 se fijaron en dos millones de pesos (diez millones de pesetas), pero para obtener la autorización oficial se debía exportar con anticipación por el mismo valor. El comercio se paralizó y en octubre de 1975 se informó que el *stock* se iba agotando y las posibilidades de venta descendían. Con esta perspectiva, el 20 de noviembre murió el general Franco en Madrid y comenzó el cambio democrático. Dos años después, en 1977, la administración del presidente López Portillo reanudó las relaciones diplomáticas con España tras casi 40 años de silencio. La historia tomó otro rumbo.

CONCLUSIONES

Con carácter general, se concluye que la conservación de la documentación de las empresas editoriales es fundamental para el estudio y análisis de la historia de la edición. En el caso que nos ocupa se han localizado las actas del Consejo de Administración de Espasa-Calpe, los catálogos editoriales y varias fotografías procedentes de su archivo. En consecuencia, consideramos que las administraciones públicas deben establecer políticas, protocolos y normativas que garanticen su preservación, partiendo de la adquisición de los archivos editoriales para su posterior tratamiento documental.

La documentación editorial, y en particular las actas de Espasa-Calpe, nos permiten conocer la estructura y funcionamiento interno de la empresa en el periodo de análisis, ya que ofrece información sobre la composición y decisiones de los consejos y comités directivos, tanto en lo que se refiere a la administración como a las obras editadas.

El corpus que aquí se analiza constituye un referente para conocer las relaciones internacionales en cuanto a la industria del libro, así como las políticas para su producción y comercialización. Es por tanto una fuente de información imprescindible para conocer las actividades españolas sobre la expansión del libro en América, partiendo de tres actuaciones específicas: la celebración del II Congreso Nacional del Comercio en Ultramar (1929), la creación del Consorcio Nacional de Editores Exportadores de Barcelona (1929) y la fundación del Sindicato Español del Libro Español (1930). Por otra parte, se ponen de relieve los problemas para la comercialización de los libros en México: los aranceles por exportación, la piratería o falta de respeto a los derechos de autor, y las diferencias en cuanto a la protección de la propiedad intelectual (30 años después de la muerte del autor en México y 80 en España).

Las delegaciones de la editorial en México, Argentina y España, en especial en la década de los cincuenta, aumentaron la producción y generaron un flujo comercial mediante el intercambio de libros en función de las necesidades de cada país. Queda patente la relación cultural y comercial entre Espasa-Calpe Argentina y Mexicana, realizando actividades de compensación para acometer los problemas de una y otra empresa. Las referencias concretas nos permiten conocer los momentos críticos en lo que se refiere a las economías mexicana y argentina, así como a las políticas del libro. Si en un primer momento la empresa mexicana dependió de la delegación de Buenos Aires, su constitución en sociedad en 1948 la transformó en una entidad clave para la casa central, tanto desde el punto de vista económico como cultural al editar obras de autores mexicanos, especialmente en la colección Austral.

Por lo que respecta a la propia actividad de Espasa-Calpe Mexicana, se concluye que su papel fue decisivo para la industria del libro en España y México, como se observa en las referencias al volumen de negocio y a los movimientos de capital de la empresa. Su contribución al desarrollo de la economía española mediante el envío de recursos es un hecho, fundamentalmente en las tres décadas posteriores a la Guerra Civil española de 1936-1939. Fue también base de la industria generada por la empresa matriz durante la posguerra, en los momentos críticos de la economía española.

Por último, con esta investigación se pone en valor la aportación de la editorial Espasa-Calpe a la cultura y se aportan nuevos datos para la historia de la edición en México y España, destacando la labor de aquellos intelectuales y profesionales que dirigieron la empresa durante medio siglo entre 1927 y 1975, especialmente Manuel Olarra Garmendia, José Jiménez Nicolau y Francisco Rubio.

REFERENCIAS

- Armendáriz Sánchez, Saúl y Magdalena Ordóñez Alonso. 1999. "La aportación de los refugiados españoles a la bibliotecología mexicana: notas para su estudio". *Proyecto Clío* 8. Fecha de consulta: 24 de septiembre de 2016. <http://clio.rediris.es/pdf/exilio.pdf>
- Ateneo Español de México. 1979. *Obra impresa del exilio español en México 1939/1979*. Catálogo de la exposición. México: Inbad/Sep.
- "Auge del libro, El". 1946. *La Propiedad*, México, 14 de septiembre, en *Editor, librero e impresor. Guión autobiográfico profesional*, Rafael Giménez Siles, 125. México: Impresora Azteca, 1978.
- Calvo Sotelo, L. 1927. *El libro español en América*. Madrid: Gráfica Universal.
- Cendán Pazos, Fernando. 1972. *Edición y comercio del Libro Española, 1900-1972*. Madrid: Editora Nacional.
- Escarpit, Robert. 1965. *La revolución del libro*. Madrid: Alianza Editorial.
- Fernández, Iñigo. 2008. *Historia de México*. México: Panorama Editorial.
- Fernández Moya, María. 2009. "Editoriales españolas en América Latina. Un proceso de internacionalización secular". *Información Comercial Española. Cuadernos económicos* 848 (julio-agosto): 65-77.
- Fernández Moya, María. 2008. "Editores españoles a ambos lados del Atlántico: el sector editorial republicano y la edición en la España franquista". *Historia del presente* 12 (España): 97-110.
- García Diego, Javier. 2015. *Autores, editoriales, instituciones y libros: estudios de historia intelectual*. México: El Colegio de México.
- Gili Roig, Gaspar. 1944. *Bosquejo de una política del libro*. Barcelona: Hispano Americana.
- Giménez Siles, Rafael. 1978. *Editor, Librero e Impresor. Guión autobiográfico profesional*. México: Imprenta Azteca.
- INLE (Instituto Nacional del Libro Español). 1962. *Informe sobre las dificultades con que tropieza en varios países de la América Española la exportación de nuestros libros*, 22 de noviembre de 1962. Archivo General de la Administración, Cultura, Expediente 73-000475.
- Ixba Alejos, Elizer. 2013. "La creación del libro de texto gratuito en México (1959) y su impacto en la industria editorial de su tiempo: Autores y editoriales de ascendencia española". *Revista Mexicana de Investigación Educativa* 18 (59): 1-19.
- Kloss Fernández del Castillo, Gerardo. 2007. *Entre el oficio y el beneficio: el papel del editor: práctica social, normatividad y producción editorial*. Guadalajara: Universidad de Guadalajara y Santillana Editorial.
- Martínez Martín, Jesús Antonio, coord. 2015. *Historia de la edición en España, 1939-1975*. Madrid: Marcial Pons.
- Martínez Martín, Jesús Antonio. 2009. *Historia de la edición en España, 1836-1936*. Madrid: Marcial Pons.
- Martínez Rus, Ana. 2003. *La política del libro durante la Segunda República. Socialización de la lectura*. Gijón: Trea.
- Martínez Rus, Ana. 2000. "La proyección editorial en los mercados americanos (1901-1936)". *Pliegos de Bibliofilia* 12 (4º trimestre): 31-53.

- Olarra Jiménez, Rafael. 2003. *Espasa-Calpe. Manuel Olarra, un editor con vocación americana*. Buenos Aires: Dunken.
- Sánchez Andrés, Agustín y Pedro Pérez Herrero. 2010. *Las relaciones entre España y México, 1820-2010*. Madrid: Real Instituto Elcano.
- Sánchez Vigil, Juan Miguel. 2009. *La edición en España. Industria cultural por excelencia. Historia, proceso, gestión, documentación*. Gijón: Trea.
- Sánchez Vigil, Juan Miguel. 2005. *Calpe. Paradigma editorial*. Gijón: Trea.
- Santonja, Gonzalo. 2003. *Los signos de la noche. De la guerra al exilio. Historia peregrina del libro republicano entre España y México*. Madrid: Castalia.
- Torre Villar, Ernesto de la. 1999. *Breve historia del libro en México*. México: Universidad Nacional Autónoma de México.

Para citar este texto:

- Sánchez Vigil, Juan Miguel. 2018. "Origen y desarrollo de la editorial Espasa-Calpe Mexicana (1927-1975)". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 31-57.
<http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57896>

O ensino pela pesquisa: interlocuções com Paulo Freire na docência em biblioteconomia

Jorge Santa Anna*

*Artículo recibido:
7 de noviembre de 2017*

*Artículo aceptado:
23 de agosto de 2018*

Artículo de investigación

RESUMEN

El objetivo de este artículo es relatar y confirmar la propuesta freireana de enseñar aprendiendo al ser aplicada en el contexto de la Ciencia de la Información, una de las áreas que más exige habilidad en cuanto a los procedimientos de investigación y para compartir el conocimiento. El artículo discurre acerca de la práctica de la investigación aliada al diálogo y al intercambio de informaciones, procedimientos adoptados por un docente de Biblioteconomía en una universidad. Presenta los resultados alcanzados a partir de la aplicación de procedimientos de investigación en una disciplina respecto de la formación académica de bibliotecarios. Refuerza la importancia de esta práctica en el ejercicio de la docencia y la contribución del

* Universidad Federal de Minas Gerais, Brasil. professorjorgeufes@gmail.com

aprendizaje compartido en las aulas. La metodología adoptada caracteriza el estudio como investigación descriptiva, de abordaje cualitativo, con técnicas de observación, relato de experiencia y entrevista. Se confirmó la viabilidad en adoptar la investigación y compartir el conocimiento en la docencia en Biblioteconomía, teniendo en cuenta el dominio crítico-reflexivo proporcionado al alumnado, así como la satisfacción de esos sujetos en cuanto al uso de ese método de enseñanza en la formación académica.

Palabras clave: Enseñanza de la Investigación; Docencia en Biblioteconomía; Enseñanza- Aprendizaje; Método Freireano de Enseñanza

The teaching for the research: interlocutions with Paul Freire in teaching in library

Jorge Santa-Anna

ABSTRACT

This paper examines the Freirean proposal for teaching/learning, specifically in the field of Information Science, a field that requires both rigorous research procedures and knowledge sharing, and discusses the approach to research, adopted by a university Librarianship lecturer, that is allied with dialogue and exchange of information. Results from the application of such research procedures during the academic training of librarians are reported that serve to reinforce the importance of shared teaching/learning in classrooms. This study employs a descriptive research methodology, using a qualitative approach that includes observation techniques, experiential reporting and interviews. Through the critical-reflexive milieu provided to students, the practicality of sharing knowledge in practice of teaching in Librarianship is confirmed. Moreover, student satisfaction was enhanced through the deployment of this teaching method in academic training.

Keywords: Teaching by Research; Teaching in Librarianship; Teaching-learning; Freirean Method of Teaching

INTRODUÇÃO

A prática educativa acompanha o ser humano infinitamente, perpassando as diversas fases da vida. Desde os ensinamentos familiares, à educação básica e profissional, até à vivência diária nas relações de trabalho, o ser humano está em constante aprendizado. Essa prática manifesta-se em todas as áreas do conhecimento, intensificando-se, sobremaneira, no período correspondente à formação acadêmico-profissional, haja vista a preparação técnica e humanista do indivíduo para uma atuação efetiva no contexto social.

No âmbito da formação acadêmico-profissional, o papel desempenhado pelos docentes é de suma importância. O exercício da docência constitui um dos fazeres mais estratégicos e provocativos no que se refere à aquisição/ampliação de conhecimento. O ato de ensinar requer a preparação prévia do docente, por meio de estudo e pesquisa que o capacite ao domínio específico de determinada temática ou área do conhecimento, viabilizando o processo de produção de conhecimento individual.

Contudo, é louvável e ainda mais produtivo, aquele que conduz seus métodos de ensino por meio da filosofia do ensinar aprendendo. Além de produzir conhecimento individual e transmiti-lo ao alunado, o docente, nessa concepção, ao mesmo tempo que ensina, também recebe conhecimentos do alunado, firmando um processo educativo recíproco e interativo, com vistas a despertar uma produção coletiva de conhecimento.

A sala de aula transforma-se em um ambiente de socialização, atuando o docente como mediador que conduz as atividades mútuas de ensino-aprendizagem. Essa proposta de ensino revoluciona a forma de educação empregada de forma centralizada na figura do professor como centro das atenções, caracterizando um processo de alienação. As propostas freireanas¹ consideram a incompletude do ser humano, logo, o professor também está em constante aprendizado. Essa incompletude é justificada porque é “[...] impossível saber-se incabado e não se abrir ao mundo e aos outros à procura de explicação, de respostas a múltiplas perguntas” (Freire, 2006: 136).

Assim, o professor é o personagem que ensina e aprende ao mesmo tempo. Mantendo uma relação pacífica e conjunta com os alunos, trocando com eles informação, o professor capacita-os a adquirir competências e habilidades essenciais para resolver problemas rotineiros presentes nas relações sociais, como também, ter capacidade em comunicar-se, tomar decisões, fundamentar suas ideias, enfim, relacionar-se socialmente.

1 Refere-se aos ensinamentos propostos por um dos maiores teóricos e educadores do Brasil, Paulo Freire. O pensamento desse teórico revolucionou a pedagogia mundial, propondo novos métodos de ensino, em prol de uma educação libertadora e crítica (Guimarães *et al.*, 2017).

O objetivo de formar alunos críticos, seguros e participativos, tendo em vista concretizar seus objetivos de vida na sociedade deve ser uma das preocupações dos professores da sociedade atual. Isso porque, segundo Behrens (1999, em Rodrigues, 2010: 148), “a sociedade passa a exigir profissionais que tenham capacidade de tomar decisões, que sejam autônomos, que produzam com iniciativa própria, que saibam trabalhar em grupo, que partilhem suas conquistas e que estejam em constante formação”.

A fim de adquirir conhecimento e formar indivíduos transformadores é preciso, também, que o professor adote métodos que incentive os discentes ao hábito da pesquisa. Ensinar por meio da pesquisa desperta nos aprendizes o espírito investigativo e inovador. Assim, segundo Piaget (1974: 18), o verdadeiro professor é aquele que “[...] deixa de ser apenas um conferencista e estimula a pesquisa e esforço, em lugar de contentar-se em transmitir os problemas já solucionados”.

Mesmo cientes da importância do processo de pesquisa na educação superior, muitos desafios ainda perfazem a consolidação dessa prática, pois a pesquisa tem sido ensinada de uma maneira ineficiente, e, possivelmente com algum efeito colateral negativo. Considera-se importante introduzir o estudante no processo real de conduzir pesquisa e não apenas em ler sobre o assunto (Witter, 1987).

Sendo assim, este estudo propõe reflexões sobre a importância da prática da pesquisa no exercício da docência e a contribuição do compartilhamento de conhecimento nas salas de aula, considerando o pensamento de Paulo Freire como fundamental para o sucesso do processo educativo no decurso da formação acadêmico-profissional.

O presente texto narra a prática da pesquisa aliada ao método freireano, procedimentos esses adotados por um docente do Curso de Biblioteconomia em uma universidade. Discorre acerca dos procedimentos adotados pelo docente no intuito de viabilizar a pesquisa, o interesse e a participação conjunta entre aluno e professor, com vista a concretizar a construção coletiva de conhecimento. Além disso, o estudo apresenta, por meio de método avaliativo, a percepção dos discentes quanto às metodologias e didática adotadas pelo professor no decorrer da disciplina.

A proposta aqui formulada é conduzida por meio de uma pesquisa descritiva, de abordagem qualitativa, com técnicas de observação, relato de experiência e entrevista, considerando as atividades e vivência prática de um docente no contexto da formação acadêmica em Biblioteconomia/Ciência da Informação no Brasil.

METODOLOGIA

No intuito de relatar a aplicabilidade das propostas educativas de Freire (2006), na docência em Biblioteconomia, de modo a demonstrar a viabilidade da prática colaborativa e de pesquisa adotada no ensino dessa área de conhecimento, a pesquisa aqui apresentada caracteriza-se como descritiva, de abordagem qualitativa, com técnicas de observação, relato de experiência e entrevista.

O estudo manifesta-se com essas características metodológicas porque: descreve a percepção dos alunos acerca do método adotado na condução de uma disciplina (pesquisa descritiva); apresenta informações oriundas da literatura e das opiniões de discentes, o que não requer a quantificação/mensuração desses dados (pesquisa qualitativa); utiliza de constatações advindas de comportamentos monitorados (observação); relata atividades, fatos e ocorrências de uma realidade prática (relato de experiência); e, por fim, utiliza dados coletados em um processo de avaliação, contendo a opinião de sujeitos sobre um determinado objeto (entrevista).

Os dados coletados para aferição dos resultados, a princípio, foram extraídos da literatura até então publicada sobre métodos de ensino, com foco na obra de Paulo Freire. Esse diálogo com outras pesquisas é de fundamental importância para proporcionar consistência e comparação dos resultados alcançados com este estudo. Portanto, utilizando-se de leituras, adentrou-se a artigos científicos e livros das áreas de Educação e Biblioteconomia/Ciência da Informação, a fim de fomentar o discurso e os argumentos em comparação com os dados que foram levantados da realidade prática e devidamente interpretados.

Além disso, também se coletou dados por meio do monitoramento de uma disciplina ofertada no curso de Biblioteconomia de uma universidade. A análise dos procedimentos adotados pelo professor consolidou-se por meio da aplicação da técnica de observação e de entrevista.

O estudo de observação foi realizado junto às atividades ministradas pelo professor, incentivando os discentes a pesquisarem as temáticas propostas no programa de disciplina. Já a entrevista foi aplicada aos alunos, no final da disciplina, no intuito de analisar a percepção deles quanto a esse método de ensino utilizado pelo docente. Essas técnicas de coleta de dados, considerando a realidade prática e o acompanhamento do pesquisador, ao longo de sua experiência como docente, caracterizam, metodologicamente, o relato de uma experiência.

Em suma, no que se refere ao instrumento de coleta de dados junto aos alunos, foi elaborada uma pergunta aberta, de modo que os sujeitos descrevessem

em papel que lhes foi entregue, os pontos positivos e negativos da disciplina, considerando as metodologias de ensino adotadas, como: leituras constantes, pesquisas na literatura sobre temas específicos, rodas de conversa, seminários, trabalhos em grupo, elaboração de resumos e fichamentos, dentre outras.

Os dados coletados da realidade prática, a interlocução com os dados percorridos na literatura e os resultados possíveis alcançados a partir dessas análises são descritos na seção seguinte.

APRESENTAÇÃO E ANÁLISE DE RESULTADOS

A disciplina ministrada pelo docente de Biblioteconomia foi “Formação e desenvolvimento de coleções (FDC)”, disciplina essa ofertada no sétimo (penúltimo) período do curso. A turma continha um total de 26 alunos, os quais encontravam-se em fase final da Graduação, alguns até iniciando os trabalhos de conclusão de curso (TCCs). Todavia, faz-se necessário destacar que eles não estavam familiarizados com o método da pesquisa, uma vez que o curso e a própria universidade estimula a prática da pesquisa, com maior intensidade, no momento da elaboração dos TCCs.

O professor, no primeiro dia de aula, discorreu acerca do método freireano, assim como demonstrou as constantes atividades de pesquisa que norteariam as três unidades da disciplina. Como não haveria avaliação final escrita, os alunos seriam avaliados conforme o empenho e engajamento com as pesquisas e consequente produção de textos. Em virtude de a disciplina de FDC possui características teóricas, esse fato proporcionou a aplicação das técnicas de pesquisa, sobretudo na primeira unidade, quando se estuda a evolução histórica das coleções ao longo dos tempos.

Foi possível observar, a priori, um certo espanto e admiração dos alunos, acerca das atividades que seriam desenvolvidas, sobretudo no que se refere à aplicação de muitos trabalhos e à ausência de avaliação formal, como o público estava acostumado. Essa admiração inicial também foi constatada no estudo de Grillo *et al.* (2016), o qual defendeu haver necessidade, antes de adotar essa metodologia, de “[...] reflexão e de amadurecimento, o que implica a autoaceitação do professor como mediador no ensino, e do aluno como protagonista de sua aprendizagem” (Grillo *et al.*, 2016: 1).

A cada aula, o professor utilizava estratégias que convidavam os alunos à participação. O ato de fazer perguntas individuais, de aplicar atividades para casa, sendo discutidas na aula seguinte, de indicar livros e artigos de fácil leitura constituem alguns dos procedimentos que motivavam ou convidavam o aluno a trocar conhecimentos com o professor.

DISCUSSÃO

Por meio da análise aos diversos relatos constantes na literatura, importante citar o trabalho de Isaia (2006), para quem, a prática de pesquisa, diálogo e interação entre alunos e professor consolida a chamada aprendizagem compartilhada, que pode ser definida como um processo de construir conhecimento de forma colaborativa no âmbito do ensino nas universidades, “[...] em que docentes e alunos constroem, por meio de uma aprendizagem colaborativa, conhecimento profissional compartilhado” (Isaia, 2006: 77).

Na unidade I da disciplina foi apresentado o percurso histórico do processo de FDC. A turma foi dividida em quatro diferentes grupos, estando cada um responsável em pesquisar como as coleções bibliográficas eram monitoradas nos períodos históricos, como: na Antiguidade, Idade Média, Idade Moderna e Contemporânea.

Em cada aula, os grupos traziam as fontes que haviam pesquisados e contribuíam com o debate mediado pelo docente. As nomenclaturas mais importantes deveriam ser esclarecidas, como: bibliófilos, copistas, armazenamento de coleções, preservação/conservação, bibliotecas minerais e vegetais, códices, papiro, gestão das coleções, desbastamento, planejamento da coleção, dentre outras.

Percebe-se a fundamental importância em motivar os alunos a pesquisar em diferentes fontes, sejam livros, artigos científicos ou meramente informativos, documentos arquivísticos, dentre outras, uma vez que essa prática familiariza o pesquisador com as estruturas e especificidades de cada tipo de fonte analisada, facilitando, portanto, a localização do assunto que está sendo pesquisado.

Além de treinar o aluno com o uso das fontes de pesquisa, o ato de pesquisar permite o entendimento acerca de conceitos, terminologias, teorias, leis, paradigmas e fatos históricos, elementos de suma necessidade para enriquecer o vocabulário do pesquisador, como permitir que ele entenda as questões teóricas e epistemológicas, para, posteriormente, aplicar o conhecimento produzido. Com esse estímulo à pesquisa nas fontes e o compartilhamento das descobertas, pode-se confirmar que “saber ensinar não é transferir conhecimento, mas criar as possibilidades para a sua própria produção ou a sua construção” (Freire, 1997: 52 em Ferreira, 2017: 11).

Na unidade II, abordaram-se as etapas do processo de FDC, enfocando no Modelo Holístico de Evans, o qual propõe a delineação de seis etapas: estudo da comunidade, política, seleção, aquisição, desbaste e avaliação. Cada uma

dessas etapas foi entregue aos grupos que realizaram pesquisas e socializaram com a turma por meio de seminários e rodas de conversa².

Tendo em vista a aplicação dessas técnicas de ensino e a liberdade em pesquisar que foi conferida aos alunos, evidencia-se que esse tipo de ensino viabiliza a reconstrução contínua do conhecimento, de modo que novos pontos de vista surgem, evidenciando que não há verdade absoluta (Bolzan, 2002). Assim,

O conhecimento é gerado e co-construído coletivamente, e produzido na interatividade entre duas ou mais pessoas que dele participam, constituindo-se o núcleo da atividade. Tarefas conjuntas provocam a necessidade de confrontar pontos de vista divergentes, acerca de uma mesma atividade, o que possibilita a descentralização cognitiva e se traduz no sociocognitivo que mobiliza as estruturas intelectuais existentes e obriga os sujeitos a reestruturá-las (Bolzan, 2002: 53).

No bojo das atividades realizadas pelos alunos e a apresentação dessas atividades, importante lembrar que o processo de pesquisa não se constituía, tão somente, de leitura. O grupo também deveria, no dia da apresentação, trazer um resumo expandido da temática que estava pesquisando. Foram ensinadas técnicas de leitura, de redação e de estrutura e normalização de trabalhos técnico-científicos.

A interação entre pesquisa e produção de textos faz parte do universo do estudante universitário. Os procedimentos adotados para leitura, interpretação, elaboração e apresentação seguem um formalismo que consolida os rigores científicos, logo proporcionar habilidades ao alunado quanto a esse “modismo” é uma estratégia interessante dos docentes no decurso das atividades disciplinares. Além de ensinar esse formalismo, faz-se necessário motivar o aluno a atribuir significado àquilo que elabora (Marquesin, Benevides e Baptista, 2011).

Por fim, na unidade III, os alunos pesquisaram o que a literatura mencionava sobre processos de FDC em diferentes modalidades de biblioteca. Assim, cada grupo fez levantamentos bibliográficos, acompanhamento e visita a uma modalidade de biblioteca, entrevistando o bibliotecário e observando como os processos eram realizados, comparando com as recomendações da literatura. As pesquisas de campo foram socializadas em sala, no último dia de aula, sendo que cada contexto analisado foi comparado com os demais.

2 Essas técnicas de ensino caracterizam-se por manifestarem-se como importantes instrumentos a serem utilizados como prática metodológica de aproximação entre os sujeitos no cotidiano pedagógico. É por meio do diálogo inerente a essas técnicas que o conhecimento é compartilhado entre alunos e professores (Melo e Cruz, 2014).

Essa atividade proposta pelo professor promoveu a integração entre teoria e prática, de modo a tornar mais visível para os aprendizes do que é ensinado na sala de aula. No âmbito da Biblioteconomia/Ciência da Informação, conhecer a realidade social e as instituições bibliotecárias é um pressuposto para planejar quaisquer serviços de informação (Almeida, 2011). Além disso, ao demonstrar a realidade social das instituições e da sociedade como um todo o docente assume frente a seus alunos “[...] um compromisso com os destinos do país. Compromisso com seu povo. Com o homem concreto. Compromisso com o ser mais deste homem” (Freire, 2007: 25).

Nesse contexto, acredita-se ser viável a integração entre teoria e prática na condução de uma disciplina do ensino superior. Isso porque viabiliza o conjunto dos objetivos principais da universidade, por meio da tríade ensino, pesquisa e extensão, sendo que essas três dimensões se complementam (Emmel e Krul, 2017).

Acreditamos que a Docência no Ensino Superior está implicada em ensino, pesquisa e extensão, formando uma tríade articulada entre si, uma vez que a atividade de ensino também pode ser realizada sob uma atitude investigativa, e assim sucessivamente a extensão, que se relaciona a pesquisa, pois implica a produção de conhecimentos vinculados com a vida em sociedade (Emmel e Krul, 2017: 1, grifo nosso).

A investigação *in loco* e a devida apresentação das experiências vivenciadas por cada grupo de trabalho demonstrou o quanto os alunos já possuem algum conhecimento acerca da realidade, dos produtos, serviços e recursos presentes nas unidades de informação. Por meio das visitas, os alunos aperfeiçoaram o conhecimento, como também identificaram problemas e casos de sucesso relativos às atividades do profissional no mercado de trabalho. Esse resultado se assemelha às mesmas percepções de alunos alcançadas com o estudo de Almeida (2011), acerca das atividades propostas em disciplinas do Curso de Biblioteconomia, como: Disseminação da Informação, Fontes de Informação e Ação Cultural.

Vê-se que os alunos estiveram envolvidos com pesquisa, durante o desenvolvimento da disciplina FDC. O professor reportou-se como um verdadeiro orientador das práticas de pesquisa. Por consequência dessa metodologia, foi possível perceber que a sala de aula transformou-se em um ambiente de compartilhamento do conhecimento, um espaço para diálogo e relações harmoniosas, além de se fortalecer, nesse contexto interativo, ações como: indagação, refutação, confirmação, ética, confiabilidade, precisão,

transparência e ajuda mútua, elementos esses primordiais para garantir boas práticas no fazer científico e profissional.

A fim de analisar a percepção dos alunos a respeito do método adotado, o docente aplicou um processo avaliativo, solicitando a descrição dos pontos positivos e negativos. No dia da aplicação do questionário, estavam em sala, 20 alunos, sendo que foram devolvidos um total de 18 questionários. Em virtude das respostas serem muito parecidas, apresentam-se, por meio do quadro 1, os principais pontos descritos, sendo que a maioria registrou satisfação quanto ao método adotado na disciplina de FDC.

| Pontos positivos sobre o método de ensino | Pontos negativos sobre o método de ensino |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| "Foram aulas em que todos permaneceram atentos, se empenhando nas pesquisas" | "As aulas foram tão boas que ficávamos torcendo para que chegasse o dia da disciplina. Ficava triste quando a aula acabava. Único ponto negativo foi a intervenção do professor no final das apresentações, apresentando os erros dos alunos. Isso pode constranger alguns. Do resto foi tudo ótimo" |
| "O método de pesquisa conferiu a nós, maior autonomia e autoconfiança" | "Muito trabalho para fazer. Se fosse só uma disciplina, mas fazemos várias em um período" |
| "Evoluí muito como pessoa e como pesquisadora. Aprendi a fundamentar minhas ideias e defendê-las com embasamento" | |
| "A pesquisa dá aos alunos autoridade, clareza, independência e objetividade. Incentiva-nos a descobrir nossas potencialidades" | |
| "A melhor disciplina que já tivemos. Foi-nos dada a chance de vivenciar o sentido de uma universidade que é não apenas aprender, mas construir e ensinar" | |
| "Aulas foram dinâmicas, democráticas, levando os alunos a pensarem criticamente sobre o tema FDC" | |

Quadro 1. Descrição dos principais pontos avaliados pelos alunos
Fonte: dados da pesquisa (2015)

A partir da fala dos discentes, entende-se que o método adotado na condução da disciplina de FDC tornou-os satisfeitos. Mais que isso, pois, ao analisarmos algumas falas, encontramos expressões que demonstram o quanto produtiva e transformadora foi a disciplina na vida desses futuros bibliotecários. Pode-se inferir que, algumas características como autonomia, autoconfiança,

segurança, pensamento crítico e embasamento de ideias representam a mudança de postura do profissional e não apenas uma evolução ou desempenho na disciplina.

Importante destacar sobre a diversidade de atividades e as inovações que a disciplina proporcionou aos alunos. Se, no início da disciplina, o professor percebeu espanto por parte do alunado, no final da disciplina, manifestou-se o engajamento dos discentes com as técnicas de pesquisa, o aperfeiçoamento no ato de discursar, o encorajamento para atuação frente aos desafios do mercado de trabalho, como também o reconhecimento do que realmente é a profissão e o fazer científico.

Esse resultado pode ser constatado por meio do relato de um dos respondentes: “Essa disciplina clareou muitas coisas referentes à nossa atuação como profissional bibliotecário. Me evolui como estudante, pesquisadora e como profissional. Gostei muito da variedade de atividades realizadas”. A respeito desse resultado avaliativo, dialogamos com Pimenta e Anastasiou (2008: 103, grifo nosso):

O ensino na universidade caracteriza-se como um processo de busca e de *construção científica e crítica de conhecimentos*. As transformações da sociedade contemporânea consolidam o entendimento do ensino como *fenômeno multifacetado*, apontando a necessidade de disseminação e internalização de saberes e modos de ação (conhecimentos, conceitos, habilidades, procedimentos, crenças, atitudes).

Outro ponto que torna-se pertinente analisar com mais detalhe diz respeito à mudança de comportamento e postura do aluno quando insere-se no universo da pesquisa. Segundo um dos entrevistados: “A pesquisa dá aos alunos autoridade, clareza, independência e objetividade. Incentiva-nos a descobrir nossas potencialidades”. Assim, o aluno adquire maior segurança quando toma suas decisões ou faz seus julgamentos, algo de extrema importância ao se inserir no mercado de trabalho. Além disso, segundo Bastos *et al.* (2010), a pesquisa, assim como a iniciação científica e projetos de extensão, oportuniza a vocação no alunado a partir dos fazeres investigativos, por meio de uma postura interventiva e crítica no que está sendo pesquisado e será descoberto.

Dos 18 relatos obtidos com a aplicação da avaliação, por escrito, é possível identificar a superação de pontos positivos quanto aos pontos negativos. Identificaram-se, como pontos negativos, os seguintes: 1 - muito trabalho para realizar, e 2 - sugestões do professor, apresentadas diante de toda a turma, acerca de melhoria na forma de os alunos apresentarem os trabalhos.

Essas indicações negativas são válidas, uma vez que podem ser utilizadas como ponto de reflexão do professor, haja vista melhorar continuamente o

seu trabalho, de modo a adequar-se às necessidades do público e atingir níveis mais elevados de excelência e qualidade. No entanto, é importante considerar, outrossim, que essas questões apontadas pelos alunos são complexas, sobretudo a intervenção no final dos seminários.

As diferentes formas de avaliação e diálogo com os alunos, como também a reflexão acerca de melhorias para o aprendizado desses sujeitos deve ser uma constante no contexto da sala de aula, sobretudo no ambiente universitário. Estabelecer formas de diálogo com os alunos de modo a transparecer a necessidade de melhoria é uma ação que se justifica, pois, no processo de ensino aprendizagem isso não é desnecessário, ao contrário, é imprescindível para despertar o aprendizado.

A esse respeito, dialogamos com Brito (2012: 4), para quem a verdadeira dimensão e direção da avaliação “[...] é a aprendizagem a cada momento, a cada passo e de forma contínua e dinâmica [...]”. Assim, entendemos que essa conduta do professor não deve deixar de ser aplicada, dada a sua importância, no entanto, é preciso pensar em formas mais sensatas e discretas de realizar as correções.

CONSIDERAÇÕES FINAIS

Este estudo apresentou o método de ensino adotado na disciplina de FDC, do curso de Biblioteconomia de uma universidade. Constatou que o ensino pela pesquisa, aliado ao diálogo e ao compartilhamento de conhecimento, bases da filosofia freireana, despertou inúmeras condutas ao alunado, como: empenho, participação intensa, motivação, interesse e interação.

Essas são algumas das características fundamentais do profissional do futuro, em frente a uma sociedade mutante e desafiadora. Os resultados comprovam que o ensino pela pesquisa e a construção coletiva do conhecimento são estratégias pedagógicas viáveis, haja vista alcançar o processo de ensino-aprendizagem. No contexto da Biblioteconomia, infere-se que esse método é ainda mais necessário, em virtude de envolver uma das áreas de atuação desse profissional: o ramo da pesquisa e, atrelado a esse fazer, manifesta-se o processo de busca e recuperação de fontes de informação.

Com efeito, é possível confirmar a relevância da proposta freireana do ensinar aprendendo, especificamente no contexto da Ciência da Informação, uma das áreas que mais exige habilidade quanto aos procedimentos de pesquisa e compartilhamento de conhecimento. Confirmou-se a viabilidade em se adotar a pesquisa e a aprendizagem compartilhada, além das relações democráticas e dialógicas firmadas entre aluno e professor, no âmbito

da docência em Biblioteconomia, haja vista o domínio crítico-reflexivo proporcionado ao alunado, como também a satisfação desses sujeitos quanto ao uso desse método de ensino na formação acadêmica.

Recomendam-se estudos futuros, acerca das estratégias utilizadas em outras disciplinas (pesquisa de campo) e a devida comparação entre essas disciplinas, como também pesquisa documental, por meio da análise à ementa das disciplinas ofertadas na Graduação e do projeto político pedagógico do curso, a fim de verificar a filosofia freireana contida nesses documentos curriculares. Ademais, o estudo direciona a uma futura análise comparativa dessa filosofia entre os diversos cursos ofertados na universidade.

REFERÊNCIAS

- Almeida, Carlos Cândido. 2011. “Profissionais mediadores da informação: experiências didático-pedagógicas em dois cursos de biblioteconomia brasileiros”. *Revista EDICIC* 1 (3): 1-17. Acesso em: 28 fev. 2018. <https://repositorio.unesp.br/bitstream/handle/11449/115137/ISSN2236-5753-2011-01-03-01-17.pdf?sequence=1>
- Bastos, Flávia, Fernanda Martins, Mara Laves, Mauro Terra e Camila Souza Lemos. 2010. “A importância da iniciação científica para os alunos de graduação em biomedicina”. *Revista Eletrônica Novo Enfoque* 11 (11): 61-66. Acesso em: 28 fev. 2018. <http://www.castelobranco.br/sistema/novo enfoque/files/11/artigos/08.pdf>
- Bolzan, Dóris Pires Vargas. 2002. *Formação de professores: compartilhando e reconstruindo conhecimentos*. Porto Alegre: Mediação.
- Brito, Rosa Maria Cavalcanti. 2012. *O professor, a aprendizagem significativa e a avaliação: base para o sucesso escolar do aluno*. Acesso em: 28 mar. 2018. http://www.anpae.org.br/seminario/ANPAE2012/1comunicacao/Eixo03_38/Rosa%20Maria%20Cavalcanti%20Brito_int_GT3.pdf
- Emmel, Rúbia e Alexandre Jose Krul. 2017. “A docência no Ensino Superior: reflexões e perspectivas”. *Revista de Ensino Superior* 3 (1): 1-17. Acesso em: 15 mar. 2018. <http://dx.doi.org/10.18256/2447-3944/rebes.v7n1p42-55>
- Ferreira, Glória Izabel. 2017. A prática pedagógica do professor de Biblioteconomia. Acesso em: 28 mar. 2018. <http://www.lume.ufrgs.br/bitstream/handle/10183/10202/000294923.pdf>
- Freire, Paulo. 2006. *Pedagogia da autonomia: saberes necessários à prática educativa*. 33. ed. São Paulo: Paz e Terra.
- Freire, Paulo. 2007. *Educação e mudança*. 30ª ed. Rio de Janeiro: Paz e Terra.
- Grillo, Marlene Corroero, Dêcia Enricone, Patrícia Mattei e Josele Benedito Ferreira. 2016. “Ensino e pesquisa com pesquisa em sala de aula”. *UNirevista* 1 (2): 1-11. Acesso em: 28 mar. 2018. <http://faculda debarretos.com.br/wp-content/uploads/2015/11/pesquisa-sala-de-aula2.pdf>
- Guimarães, Alessandro, Ednalva Gutierrez Rodrigues, Fernanda M. B. Camargo, Joelma de Souza Mendes Rizzo, Margarete Sacht Goes e Shenia D’Arc Venturim Cornélio. 2017. “A palavra é... Paulo Freire”. *Pró-Discente: Caderno de Produção*

- Acadêmico-Científica* 23 (1): 64-81. Acesso em: 28 mar. 2018. www.periodicos.ufes.br/PRODISCENTE/article/download/18600/12608
- Isaia, Sílvia Maria de Aguiar. 2006. “Desafios à docência superior: pressupostos a considerar”, em *Docência na Educação Superior*, organizado por Dilvo Ristoff e Palmira Sevegnani, 63-84. Brasília: INEP.
- Marquesin, Denise Filomena Bagne, Cláudio Roberto Benevides e Denise Cristina Baptista. 2011. “Leitura e escrita no ensino superior”. *Revista de Educação* 14 (17): 9-28. Acesso em: 28 mar. 2018. <http://www.pgsskroton.com.br/seer/index.php/educ/article/viewFile/1801/1714>
- Melo, Marcia Cristina Henares de e Gilmar de Carvalho Cruz. 2014. “Roda de conversa: uma proposta metodológica para a construção de um espaço de diálogo no Ensino Médio”. *Imagens da Educação* 4 (2): 31-39. Acesso: 20 mar. 2018. <http://dx.doi.org/10.4025/imagenseduc.v4i2.22222>
- Piaget, Jean. 1974. *Educar para o futuro*. Rio de Janeiro: Fundação Getúlio Vargas.
- Pimenta, Selma Garrido e Léa das Graças Anastasiou. 2008. *Docência no Ensino Superior*. 3 ed. São Paulo: Cortez.
- Rodrigues, Mara Eliane Fonseca. 2010. “A abordagem do ensino com pesquisa: uma alternativa pedagógica para o ensino de biblioteconomia e ciência da informação”. *TransInformação* 22 (2): 147-167. Acesso: 05 abr. 2018. <http://dx.doi.org/10.1590/S0103-37862010000200005>
- Witter, Geraldina Porto. 1987. “O ensino de metodologia científica em biblioteconomia: algumas considerações”. *Ciência da Informação* 16 (2): 145-149. Acesso em: 28 mar. 2018. <http://revista.ibict.br/ciinf/article/view/257>

Para citar este texto:

Santa Anna, Jorge. 2018. “O ensino pela pesquisa: interlocuções com Paulo Freire na docência em biblioteconomia”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 57-72. <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57867>

La opinión en textos con un enfoque interdisciplinar: propuesta de contexto y léxico ad hoc

Silvana Grazia Temesio Vizoso*

Artículo recibido:
25 de octubre de 2017

Artículo aceptado:
27 de agosto de 2018

Artículo de revisión

RESUMEN

El artículo aborda la cuestión de cómo determinar la polaridad u opinión en textos, particularmente en Twitter. El análisis se hace desde múltiples perspectivas: ética de la información, comunicación, ciencia de la información e informática, enfatizando en dos análisis que se complementan: el procesamiento de lenguaje natural y la organización del conocimiento.

Se desarrolla un prototipo que propone anotar las palabras cargadas de opinión como los adjetivos y los adverbios contextualizadas por su núcleo: sustantivos y verbos respectivamente. Estos núcleos sirven también para modelar la terminología del dominio y se propone la creación de un lexicon *ad hoc*.

* Facultad de Información y Comunicación, Universidad de la República
Montevideo, Uruguay silvana.temesio@fic.edu.uy

Palabras clave: Twitter; Análisis de Sentimientos; PLN; Ciencia de la Información

A context and ad hoc lexicon proposal to an interdisciplinary approach to opinion analysis

Silvana Grazia-Temesio Vizoso

ABSTRACT

This paper addresses the question of how to determine polarity of opinion among texts and, more specifically, those published on Twitter. The analysis is performed from multiple perspectives, including information science, and communication and information ethics with emphasis on the analysis of natural language processing and knowledge organization.

A prototype is developed that proposes annotation of opinion conveying words as adjectives and adverbs contextualized by grammatical core, i.e., nouns and verbs, respectively. These nuclei also serve to model the terminology of the domain, and the development an ad hoc lexicon is proposed.

Keywords: Twitter; Sentiment Analysis; PLN; Information Science

INTRODUCCIÓN

El problema de la caracterización de los textos de acuerdo a los tópicos o asuntos es de importancia para lidiar con la gran cantidad de textos disponibles en la web. Una cuestión que se deriva de esto es la caracterización de esos textos en cuanto a la orientación de opinión.

La opinión es importante para medir la popularidad de un producto, una política, o cualquier otro asunto susceptible de ser analizado en forma subjetiva. Mientras que los tópicos se identifican por palabras, los sentimientos son más complejos de identificar, requieren mayor contexto, mayor comprensión del uso del lenguaje y detección de la ironía, entre otros aspectos. Según Liu (2010), la información textual puede ser categorizada en dos tipos principales: hechos y opiniones. Los hechos son expresiones objetivas acerca de entidades,

eventos y sus propiedades. Las opiniones son expresiones usualmente subjetivas que describen los sentimientos y valoraciones hacia las entidades o eventos y sus propiedades.

Una opinión se define por un conjunto de cinco elementos:

- *e*: es la entidad objetivo, que es el objeto sobre el que se emite la opinión; puede ser un producto, una persona, evento, tópico, etc.
- *a*: aspecto o característica de la entidad analizada. Existe una jerarquía de componentes o partes, subcomponentes, etc. *a* es un conjunto de atributos del objeto. El objeto puede ser representado por un árbol donde la raíz es el objeto y cada nodo un componente o subcomponente del objeto en una relación “parte de”. Cada nodo se asocia a un conjunto de atributos. Por ejemplo, si hablamos de una entidad *e*=teléfono móvil, *a* puede ser la batería, el teclado, el sistema operativo, etc., es decir, una serie de características de la entidad.
- *so*: orientación del sentimiento (valencia).
- *b*: persona quien expresa la opinión (*bolder*).
- *t*: momento en que se expresó la opinión.

Las redes sociales y en particular Twitter constituyen una fuente de información actualizada sobre los tópicos más diversos proporcionando la posibilidad de conocer la opinión acerca de éstos. El análisis de sentimientos o minería de opinión consiste en la clasificación del sentimiento de un texto respecto a su polaridad: positiva o negativa. Cuando el texto es la explicitación de un hecho, no tiene una carga de opinión, se considera neutro. Cuando hay una opinión se puede además realizar una gradación de intensidad de la polaridad: positivo, muy positivo, etc.

En este trabajo se abordan los desafíos del análisis de opinión en Twitter desde distintas ópticas: ética de la información, aspectos de comunicación, ciencia de la información e informática, con énfasis en el procesamiento de lenguaje natural (PLN) y la perspectiva de la organización del conocimiento.

La metodología empleada es la revisión y análisis del estado del arte desde una perspectiva multidimensional, reflexionando en torno a brindar mayor contexto a las palabras que expresan sentimientos. Esto se plasma en un prototipo donde se anotan las palabras cargadas de opinión como los adjetivos y los adverbios tomando en consideración sus núcleos: sustantivos y verbos respectivamente. Los núcleos modelan también la terminología del dominio por lo que son una herramienta para la creación de un lexicón *ad hoc* que capture el lenguaje usado.

CASO DE ESTUDIO: TWITTER

Twitter es una red social en la cual se puede escribir contenido con un límite de hasta 140 caracteres, aunque se ha considerado aumentar dicha extensión. Cada uno de estos textos cortos o tuits aparece en el perfil del autor con el criterio del más reciente primero. Este texto se puede reenviar (un retuit), con lo cual aparece en el usuario reenviante como un reenvío. Esta característica de reenvío permite evaluar la popularidad del tuit.

Los usuarios de Twitter se pueden suscribir al contenido que genera otro usuario, esto es una práctica muy común para seguir noticias suscribiéndose al canal de Twitter de un diario. La popularidad de los usuarios de Twitter se mide por la cantidad de “seguidores”. Los tuits, además de los 140 caracteres, tienen una serie de elementos que se denominan entidades: URL, imágenes, menciones a un usuario (@usuario) y # (*hashtag*) o etiqueta que se utiliza para agrupar tuits.

El volumen de información generado en Twitter es muy grande, por lo cual no es factible un tratamiento manual y por tanto existen diversas líneas de investigación en el tratamiento automático del sentimiento. Muchos de los estudios de sentimiento se realizaron en textos largos donde se puede expresar una idea con un contexto suficiente. Respecto al estudio en textos muy cortos, como es el caso de Twitter, establecer el contexto es un desafío. Además, el vocabulario usado en Twitter tiene características especiales en cuanto a que no se restringe a un dominio y las formas agramaticales son frecuentes.

La restricción en la extensión produce transformaciones en la modalidad expresiva y muchas veces se genera una confusión en los conceptos, se distorsiona la ortografía llegando a crearse un nuevo estilo ortográfico, la gramática se diluye y se hace un uso importante de neologismos. Por otra parte, también se propicia una capacidad de síntesis necesaria para este tipo de comunicación. Más allá de toda opinión, es una práctica comunicativa instaurada y desde el punto de vista del tratamiento del lenguaje natural subyace la cuestión de si civilizar el discurso para analizarlo o generar una nueva forma de tratamiento, porque el uso de esta comunicación constituye una tendencia ya afinada.

Algunos aspectos característicos de los aspectos sintácticos de los tuits son:

- Signos de puntuación o caracteres especiales

Los signos de puntuación suelen faltar o ser usados incorrectamente. En general se preprocesan los tuits y con diferentes criterios se homogeneizan los textos. Este tema merece un estudio por sí mismo, porque si bien hay un uso incorrecto desde el punto de vista gramatical es una situación de uso y forma parte de una comunicación hipervinculada donde el texto en Twitter

corresponde a un contexto que suele estar en otra parte y es un comentario, un aviso o un llamador que constituye un enunciado no oracional o un titular.

- Uso importante de locuciones verbales o neologismos

El lenguaje es coloquial y se usan términos que no aparecen en el diccionario y locuciones verbales que no tienen significado literal (por ejemplo, “estar en la chiquita”, “darle gas”). Estas locuciones tienen una componente de localización muy acentuada.

- Referencia a otro usuario de Twitter a través del símbolo @.

La referencia puede estar dentro de un contexto gramatical:

Habla Tania Da Rosa de @CAinfouy sobre la necesidad de empoderar a la unidad de acceso a la información pública **#UAIP2016**

o puede ser un elemento informativo vinculado fuera de la estructura gramatical:

Hay documentos que no deberían conservarse El mundo va hacia el Archivo Único Electrónico #uaip2016 @agesic @DestinoUruguay @mvdbureau

En este último caso –como información complementaria– suele ponerse al final o al principio del tuit. En cualquier caso el @ como etiqueta Part of Speech (PoS¹) constituye un nombre.

- # Hashtag que se usa para agrupar tuits

Puede ser una palabra normal que se usa con el signo # para agrupar mensajes o ser una concatenación de palabras que pueden o no tener significado. Por ejemplo:

Conoce la versión preliminar del Paquete de **#DatosAbiertos** contra la **#corrupción**

#corrupción no tiene un significado adicional, aunque también alude a la agrupación, pero desde el punto de vista sintáctico el # se puede eliminar o ignorar.

1 PoS (Part of Speech). Se trata de las categorías de palabras que tienen propiedades gramaticales similares: nombres, verbos, adjetivos, adverbios, preposiciones, entre otras.

Llega a su final el Seminario de Acceso a la Información Pública #UAIP2016 ¡Los esperamos el próximo año!

En este caso #UAIP2016 designa un evento y agrupa todos los comentarios del evento, pero no participa de la oración; desde el punto de vista sintáctico se podría sacar sin que la oración pierda sentido.

Una estrategia bastante usada es eliminar el # si está al inicio o al final del tuit presuponiendo que se quiere etiquetarlo y que no hay en este caso un aporte sintáctico o semántico.

- Palabras cambiadas para dar énfasis: jiiiiiiiiii, muuuchas gracias, Goooooooooolazo de Peñarol Junior Arias aparece el goleador
- Emoticones. Existen tablas de clasificación de emoticones con los valores positivo, negativo o neutro.
- URL. En algunos casos se utilizan para contextualizar, usando el título de la URL
- Titulares o texto difícilmente procesable:

#Fútbol #Especial #FENvsDEF Final Primer Tiempo @CAFenix_0 - 1
@DefensorSpClub #DaleDefe

FRONTERAS COMPARTIDAS

El análisis de un tuit desde el punto de vista de opinión puede hacerse desde los aspectos informáticos, desde el procesamiento de lenguaje natural (PLN), desde la implicancia comunicativa de las opiniones en los medios y desde la organización del conocimiento entre otros enfoques.

Las redes sociales y el *microblogging* constituyen una fuente de información que permite monitorear la opinión sin necesidad de estudios más complejos, como las encuestas. Por otra parte el volumen de información que produce Twitter es muy alto, así como la cantidad de usuarios en todo el mundo, que incluyen personajes políticos, celebridades, deportistas, científicos, académicos, instituciones privadas y de gobierno. Los periodistas y políticos lo utilizan como un elemento de publicidad y marca personal y como en 140 caracteres no se puede profundizar muchas veces se usa como un titular y la información se amplía en una URL o se complementa con elementos gráficos.

Desde los aspectos de comunicación, puede analizarse la red como una topología donde es posible reconocer los aspectos de alcance y difusión a

través del análisis de grafos utilizando herramientas de análisis y visualización de redes como Gephi,² que desentrañan los papeles de los nodos en la red.

Aspectos éticos del análisis de textos

Los mensajes en Twitter constituyen una situación de fragmentación de la información en la que al descontextualizarla, aislarla de la coyuntura histórica y geográfica en la que se emite, se abren múltiples perspectivas de interpretación del fragmento, algunas de las cuales pueden no corresponderse con la intención del emisor.

En otros casos de recuperación de información, como en un catálogo bibliográfico o un fondo archivístico, se recupera desde una interrogación que es en sí misma un fragmento pero que está contextualizada tanto para el buscador como para el objeto informativo con la utilización de un elemento común que es el control terminológico bajo la forma, por ejemplo, de un tesaurus. El objeto recuperado es generalmente una porción de información sin fragmentar y con un contexto embebido. Capurro (1987) comenta:

Podemos considerar al proceso de almacenamiento y recuperación de la información bajo un punto de vista hermenéutico como la articulación de la relación entre la apertura existencial al mundo del interrogador, sus distintos horizontes de pre-comprensión abiertos y compartidos socialmente y el horizonte prefijado del sistema. El proceso de búsqueda de información es básicamente un proceso de interpretación que tiene que ver con el contexto vital y el trasfondo del interrogador y el de aquellos que almacenan diferentes tipos de expresiones lingüísticas que tienen un significado dentro de contextos de comprensión fijos como son un tesaurus, palabras claves y esquemas de clasificación.

El análisis de sentimientos utiliza fragmentos (palabras de búsqueda) para recuperar otros fragmentos (mensajes de Twitter) y luego interpretar la inclinación emocional de estos últimos. Ese espacio de precomprensión que se utiliza para buscar los mensajes de Twitter a través de términos ya constituye un espacio de interpretación en sí mismo porque no funciona como ese puente preconstruido del catálogo bibliográfico.

En la segunda etapa, la interpretación del sentimiento del mensaje requiere a su vez ese contexto señalado para poder extraer la polaridad del emisor, porque en este caso no es una libre interpretación sino que se busca desentrañar el sentimiento del emisor.

2 <https://gephi.org/> es el sitio de la herramienta de software libre Gephi que realiza exploración y visualización para grafos y redes.

Para fijar este grado de incertidumbre se diseña una estrategia diferente a la de la recuperación de la información en un catálogo bibliográfico. No se establece un contexto predefinido, ese espacio de precomprensión señalado como es el caso de un tesoro, sino que se aborda el contexto dentro del propio fragmento. Lo que se predica con una polaridad se predica sobre un objeto interno al mensaje, el adjetivo se aplica a un sustantivo, el adverbio se aplica a un verbo. Ese espacio de precomprensión se construye al vuelo en el momento de analizar, porque ese espacio de precomprensión es volátil y local. Se construye en un contexto histórico y geográfico específico y nada asegura que persista.

Por otra parte se abre un ámbito de discusión respecto a la utilización de este escudriñamiento de polaridad en un espacio que es público pero abordable con una intención comercial, política o de una índole que no fue prevista por el emisor pero que puede recursivamente incidir en el propio emisor para beneficio del que analiza la polaridad. Este análisis está inmerso en un panorama más vasto que corresponde a la utilización de la información como un instrumento de manipulación, control social y político tergiversando la propia información.

Enfoque informático

Sentiwordnet³ es una lista de palabras con polaridad que establece un *ranking* de valores positivos y negativos que aplica a cada conjunto de sinónimos de Wordnet⁴ en inglés. Se puede traducir o mapear al español a través del sitio <http://multiwordnet.fbk.eu/english/home.php>. Hay estudios (Montejo-Ráez *et al.*, 2014) que utilizan este recurso para tratar la orientación de los tuits.

También cabe mencionar el lexicón de polaridad de licencia libre⁵ que fue desarrollado para TASS, que es un taller anual para análisis de sentimiento en español promovido por la Sociedad Española para el procesamiento del lenguaje natural (Urizar y Roncal, 2013).

Entre los métodos de aprendizaje computacional que usan conjuntos de entrenamiento en un aprendizaje supervisado pueden mencionarse SVM (Support Vector Machine), Naive Bayes y Máxima Entropía. Los resultados de estos métodos están en el entorno del 70 al 80 % de precisión (Pang, Lee y Vaithyanathan, 2002).

3 <http://sentiwordnet.isti.cnr.it/> presenta un recurso léxico para minería de opinión. A cada conjunto de sinónimos de WordNet le asigna uno de tres valores: positivo, negativo, objetivo.

4 <http://wordnet.princeton.edu/> alberga una base de datos léxica para el inglés que comprende nombres, verbos, adjetivos y adverbios. Se agrupan en conjuntos de sinónimos cada uno expresando un concepto diferente. Los conjuntos de sinónimos se vinculan por relaciones semántico-conceptuales y léxicas.

5 http://komunitatea.elhuyar.eus/ig/files/2013/10/ElhPolar_esV1.lex es el sitio del lexicón.

Según señalan Vilares, Alonso y Gómez Rodríguez (2015), otros estudios utilizan n-gramas, etiquetas PoS y lexicones. También se ha estudiado la orientación semántica en adjetivos, verbos y adverbios construyéndose una ecuación lineal. Se ha señalado que los métodos de aprendizaje supervisado no tienen buen rendimiento debido a que Twitter utiliza una gran cantidad de palabras fuera de los vocabularios. También se exploraron las gramáticas de dependencias y se propusieron métodos híbridos que combinan el uso de bolsas de palabras, el aprendizaje supervisado y diferentes estrategias para PoS específicos. En todos los casos se busca capturar contexto para aumentar la adecuación del análisis. Específicamente estos autores proponen un tratamiento híbrido que utiliza la gramática de dependencias utilizando triplas (Head, etiqueta arco, dependiente).

Una manera de analizar sentimientos en forma no supervisada es a través de los métodos de lexicones. Los lexicones son diccionarios que incluyen la polaridad semántica de la palabra. Se pueden crear manualmente o en forma automática. Las investigaciones se han centrado principalmente en el uso de adjetivos como indicadores de la orientación semántica. Se extraen los adjetivos de un texto y se anotan con su valor semántico usando el diccionario de sentimientos y luego se adjudica una orientación semántica al texto.

Según Taboada *et al.* (2011) los clasificadores SVM (Support Vector Machine) que entrenan en un conjunto de datos concreto utilizando unigramas o bigramas consiguen una adecuación alta en detectar la polaridad de un texto. No obstante esa adecuación se pierde en gran medida al utilizarlos en otro dominio diferente a aquél en que se entrenaron. Estos autores extraen lo que denominan palabras que involucran sentimiento e incluyen adjetivos, verbos, nombres y adverbios y utilizan estas palabras para calcular la orientación semántica incluyendo la gradación de la orientación. Sostienen que el lexicon creado (SO-CAL) es consistente a través de distintos dominios con base en la tesis de que hay una polaridad primordial que es independiente del contexto, que corresponde a la orientación semántica en la mayoría de los dominios. Si bien crear diccionarios en forma automática o semiautomática tiene algunas ventajas, a causa de la inestabilidad de los mismos estos autores deciden crearlos en forma manual anotando la polaridad de las palabras. Se crean diccionarios separados para nombres, verbos y adverbios.

La experiencia en métodos basados en lexicones o diccionarios de palabras que tienen anotada la orientación semántica para análisis de sentimientos da buenos resultados según Taboada *et al.* (2011). El resultado es estable incluso a través de distintos dominios, aunque cabe aclarar que se aplica para la clasificación de textos más largos (entradas de blogs, etc.) y se utilizan estrategias para sopesar los párrafos en la clasificación que no son aplicables en textos de las características de los tuits.

Existen empresas que brindan servicios SaaS (Software as a Service) para realizar el análisis de sentimiento de un texto, la extracción de tópicos, categorización de texto, entre otros. Meaning Cloud⁶ es uno de ellos y permite utilizar su API a través de una clave de acceso⁷ para realizar el análisis de sentimiento (Cervantes, 2016). A través de la API o de un *endpoint* se puede realizar:

- Análisis de sentimiento: se ingresa un texto y como salida se indica su polaridad, subjetividad, ironía o especificación de desacuerdo.
- Extracción de tema: identifica entidades nombradas y conceptos.
- Clasificación de texto de acuerdo a una taxonomía.

Meaning Cloud ofrece API para distintos escenarios. Estas API incluyen diccionarios, taxonomías y otras funcionalidades.

Enfoque PLN

El uso de una aproximación lingüística ofrece varios problemas, entre los cuales el uso de un lenguaje coloquial y muchas veces agramatical así como el escueto tamaño de los textos son los más notorios.

Tratamiento de la negación

Se ha estudiado que existe una tendencia a expresar una opinión negativa no directamente sino a través de la negación de una opinión positiva como una forma de suavizar la opinión negativa. En Vilares, Alonso y Gómez Rodríguez (2015) se hace una clasificación de las negaciones que pretende especificar el alcance de la negación en distintos casos, pero en el lenguaje informal de Twitter no parece aplicable.

En principio el uso del “no” elimina el carácter positivo de una afirmación, como en el caso siguiente:

El servicio es muy bueno vs. El servicio no es muy bueno

En este caso, si “bueno” se puntuara con un sentimiento positivo, al poner “no” el sentimiento de la frase no se neutraliza sino que expresa una opinión negativa (Liu, 2010). Otras partículas negativas son ni, nunca, nadie, ningún, nada, sin. También están los prefijos que tienen una orientación negativa:

6 <https://www.meaningcloud.com/es> ofrece un servicio para extraer significado de contenido no estructurado: conversaciones sociales, artículos, expedientes.

7 <https://www.meaningcloud.com/developer/> es el sitio para desarrolladores de Meaning Cloud.

- i-: legible, ilegible
- in-: transitable, intransitable
- im-: perfecto, imperfecto
- des-: favorable, desfavorable
- a-: normal, anormal

Habría que analizar a fondo esta cuestión porque puede haber casos en los que el prefijo no tenga una connotación negativa. La partícula “no” también puede tener una connotación que no sea negativa, como en este ejemplo: “El conductor *no* solamente nos acompañó sino que también nos dio indicaciones claras de cómo proceder”. Liu (2010) plantea que la negación de una opinión positiva constituye una opinión negativa. También que la negación de una opinión negativa constituye una opinión positiva; sin embargo, también se sostiene que existe la tendencia de suavizar las opiniones negativas por lo cual en muchos casos la negación de una opinión positiva constituye una opinión negativa o a lo sumo neutral como forma educada de exponer las opiniones. Por ejemplo:

El lugar no es tan malo.
La comida no resultó tan incomedible.

Se analizan algunos casos para el tratamiento de la negación:

- Hay un elemento de polaridad positiva que al tener una negación cambia su polaridad a negativa: Es muy bueno/no es muy bueno.
- Hay un no que refuerza lo negativo: No se puede ir. Es inaguantable.
- Hay un no que no produce un efecto negativo: ¿Quién dijo que no se puede comer un helado en invierno?
- La gradación de la negatividad (Vilares, Alonso y Gómez Rodríguez, 2015): negativo, muy negativo.

Como se puede ver hay una casuística variada por lo que un tratamiento general del “no” no daría resultados correctos en todos los casos. En el estudio realizado lo que se hace –adhiriéndose a Liu (2010)– es invertir la polaridad del tuit si aparece la partícula “no”. Si el tuit es positivo (sin tomar en cuenta la palabra no), si tiene un “no” se invierte la polaridad y se considera negativo. Si el tuit es negativo (sin tomar en cuenta la palabra no), si tiene un “no”, se considera positivo.

Subjuntivo

En español el uso del subjuntivo indica que lo que se predica no es un hecho, puede ser una afirmación hipotética o un deseo y está caracterizado por el rasgo irrealis. Las oraciones que hacen uso del subjuntivo tienen grandes probabilidades de expresar un sentimiento.

Locuciones

Las locuciones son ampliamente usadas en el lenguaje coloquial y especialmente en Twitter. Las locuciones que pueden tener una carga mayor de opinión son las locuciones adjetivales, adverbiales y verbales. Al parecer, el caso menos explorado es el uso de las locuciones verbales, que como parte de la expresión coloquial aparecen detentando polaridad como en el caso de “meter la pata”, “hacerse el oso”, “comérselo en dos panes”, “estar de fiesta”, “costar un ojo de la cara”, “cortarle la cara”, etc.

Resulta de interés incorporar a los lexicones de polaridad algunos verbos que denotan sentimiento y locuciones verbales que también tienen una carga de polaridad semántica importante y que son muy locales y difícilmente se comprenden fuera del ámbito de uso.

Cláusulas subordinadas

Liu (2010) sostiene que el “pero” en inglés cambia la orientación de la cláusula anterior y que en español funciona igual, planteando una situación que se opone o que parcialmente contradice lo anterior, como en los casos siguientes:

La comida es buena pero hay que esperar mucho.
Hay que esperar mucho pero la comida es buena.

El “aunque” funciona de manera análoga:

La casa está en un buen punto, aunque es un poco chica.

En los casos de ejemplo hay dos opiniones, una en la oración independiente y otra en la subordinada, aunque es difícil extraer una regla sobre la opinión general porque incluso manualmente no está clara la opinión global que dependerá del peso que se otorgue en forma particularizada a cada una de las predicciones en cierto sentido antagónicas que plantea la oración.

En lo que sigue se utiliza Freeling,⁸ aplicación de licencia libre (Affero GPL) mantenida por la Universidad Politécnica de Cataluña para el análisis de lenguaje que incluye distintos análisis: morfológico, anotación de PoS, *parsers* de gramáticas libres de contexto y dependencias, entre otros, para una variedad de idiomas y en particular el español.

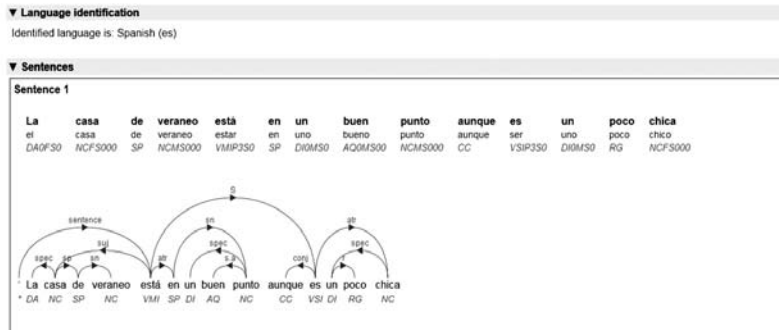


Figura 1. Freeling usando *dependency parser*

En la *Figura 1* se muestra el análisis de Freeling para la oración “La casa de verano está en un buen punto aunque es un poco chica” utilizando el *parser* de dependencias, que funciona correctamente indicando las dos oraciones. En la primera el adjetivo “buen” (denotado por Freeling como adjetivo calificativo) se aplica a “punto” de tipo nombre. Si se anotara manualmente un lexicón dando contexto se pediría la polaridad de “buen” aplicada a “punto” lo que se presume como positivo. De esta forma la primera oración sería positiva. En la segunda oración la palabra que presumiblemente tiene la carga semántica es el adverbio (poco) utilizada con el verbo (es) como contexto, podría anotarse manualmente como negativo, con lo cual la segunda oración subordinada sería negativa. Si se realizara la suma de las dos polaridades para la oración completa el resultado sería neutro, una parte positiva y una parte negativa.

Una línea para profundizar sería el análisis de que en casos como el de esta oración la carga semántica está en la etiqueta *atr* (*attribute*) y el contexto está en el sujeto de la primera oración. Entonces para la oración subordinada la anotación de polaridad con el contexto sería para *chica* con el contexto *casa*, que igualmente sería negativo pero mucho más ajustado. Otra cuestión

8 <http://nlp.lsi.upc.edu/freeling/node/1> Sitio de un conjunto de herramientas de acceso abierto para análisis de lenguaje, desarrolladas y mantenidas por la Universidad Abierta de Cataluña.

interesante respecto a la coordinación la plantea Hatzivassiloglou y McKeown (1997), quienes realizaron un estudio cuando hay dos adjetivos unidos por una conjunción coordinante. Sostienen que cuando dos palabras aparecen vinculadas por “y” siendo ambas adjetivos existe una alta probabilidad de que tengan la misma polaridad, en cambio cuando la vinculación es con “pero” la polaridad tiende a ser opuesta. Por ejemplo, “bonito y barato”, “complicado pero interesante”.

Contexto a través del uso de las características sintácticas de las palabras a través de distintos enfoques

El desafío de mejorar la clasificación automática de la polaridad del tuit pasa por establecer contexto para determinar la polaridad de las palabras. Si bien se adhiere a que existe una polaridad primordial para algunas palabras y que esta polaridad se aplica a la mayoría de los dominios, en muchos casos las palabras utilizadas tienen una polaridad que se relativiza de acuerdo al contexto en que se usan.

Las palabras con polaridad consideradas son los adjetivos y los adverbios, y aunque también los verbos y las locuciones verbales se presumen de carga de polaridad no se han considerado para este estudio. La cuestión es analizar, dado un adjetivo, a qué se aplica, a qué adjetiva; lo mismo con el adverbio.

Rojo (1975) hace una distinción entre adjetivos restrictivos y no restrictivos y analiza su ubicación en la frase nominal estableciendo que en general el adjetivo restrictivo va luego del sustantivo y el no restrictivo antes. En el ejemplo que sigue el adjetivo restrictivo modifica al sustantivo (político) y el no restrictivo a la agrupación de sustantivo y adjetivo (personaje político):

conocido **personaje político**
terrible **accidente aéreo**
evidente **crisis económica**

En estos casos la agrupación sustantivo-adjetivo restrictivo funciona en forma similar a una locución adjetival, lo que se califica es el “accidente aéreo” y la carga de polaridad está entonces en el adjetivo no restrictivo (terrible).

Cuando se usa el analizador sintáctico de Freeling, éste no realiza distinción entre los adjetivos. En el caso de la oración:

El terrible accidente aéreo ocurrió cerca de la ciudad.

El *parsing* PoS en formato CONLL (<https://nlp.stanford.edu/nlp/javadoc/javanlp/edu/stanford/nlp/sequences/CoNLLDocumentReaderAndWriter.html>):

```

1 El      el      DA0MS0 DA pos=determiner|type=article|gen=mascu-
singular
2 terrible terrible AQ0CS00 AQ pos=adjectivetype=qualificative|gen=com-
mon|num=singular
3 accidente accidente NCMS000 NC pos=noun|type=common|gen=mascu-
linelnum=singular
4 aéreo aéreo AQ0MS00 AQ pos=adjectivetype=qualificative|gen=mascu-
linelnum=singular
5 ocurrió ocurrir VMIS3S0 VMI pos=verb|type=main|mood=indicative|ten-
se=past|person=3|num=singular
6 cerca cerca RG RG pos=adverb|type=general
7 de de SP SP pos=adposition|type=preposition
8 la el DA0FS0 DA pos=determiner|type=article|gen=femininelnum=sin-
gular
9 ciudad ciudad NCFS000 NC pos=noun|type=common|gen=femininelnum=-
singular

```

El *parsing* de dependencias en formato CONLL tampoco distingue la etiqueta:

```

1 El      el      DA0MS0 DA pos=determiner|type=article|gen=mascu-
singular      --- 3 spec  --  -
2 terrible terrible AQ0CS00 AQ pos=adjectivetype=qualificative|gen=com-
mon|num=singular      --- 3 s.a  --  -
3 accidente accidente NCMS000 NC pos=noun|type=common|gen=mascu-
linelnum=singular      --- 5 suj  --  A1
4 aéreo aéreo AQ0MS00 AQ pos=adjectivetype=qualificative|gen=mascu-
linelnum=singular      --- 3 s.a  --  -
5 ocurrió ocurrir VMIS3S0 VMI pos=verb|type=main|mood=indicative|ten-
se=past|person=3|num=singular --- 0 sentence - ocurrir.00 -
6 cerca cerca RG RG pos=adverb|type=general      --- 5
cc  --  AM-TMP
7 de de SP SP pos=adposition|type=preposition      --- 6
sp  --  -
8 la el DA0FS0 DA pos=determiner|type=article|gen=femininelnum=sin-
gular      --- 9 spec  --  -
9 ciudad ciudad NCFS000 NC pos=noun|type=common|gen=femininelnum=-
singular      --- 7 sn  --  -
10. . Fp Fp pos=punctuation|type=period

```

Freeling no distingue entre terrible y aéreo para el caso y los considera adjetivos calificativos. Si se pudiera establecer esta distinción se podría especificar el contexto (accidente aéreo) para el adjetivo que tiene la carga de polaridad (terrible).

El posicionamiento de los adjetivos en torno de los sustantivos adquiere distintos formatos por lo cual el posicionamiento no sería una indicación para la determinación de los tipos de adjetivos. Pueden darse posicionamientos diversos como en los siguientes casos:

Sustantivo + adjetivo restrictivo + adjetivo restrictivo Novelas policíaca inglesas
Sustantivo + adjetivo restrictivo + adjetivo no restrictivo Crisis económica
manifiesta

Si utilizamos gramáticas independientes de contexto probabilísticas (GLCP) para analizar esta situación y definimos un grupo adjetival restrictivo (GAR) que instancie aquellos casos en que los adjetivos sean de tipo restrictivos para las oraciones 1) y 2) tenemos:

- 1) Las *novelas policíacas inglesas* están en el estante superior.
- 2) La *crisis económica* manifiesta afecta a los ciudadanos.

O -> GN GV

GN -> det nombre adjetivo adjetivo | *det GAR* | *det GAR* adjetivo | det nombre adjetivo

GAr -> nombre adjetivo | nombre adjetivo adjetivo

GV -> verbo GP

GP -> prep GN

En el caso 1) se puede aplicar GN-> det GAR (GAR -> nombre adjetivo adjetivo), pero en el caso 2) GN -> det GAR adjetivo (GAR-> nombre adjetivo)

Si se midiera la probabilidad de las reglas en un *corpus* podría estimarse cuál es la que tiene mayor probabilidad. La cuestión es que en ningún caso se toman en cuenta los aspectos léxicos que en este caso son los más significativos para determinar la característica de restrictivo de un adjetivo y por tanto las gramáticas libres de contexto probabilísticas no darían una buena solución.

En ese sentido la idea de que el núcleo en un grupo selecciona los complementos en función de sus propiedades léxicas es un punto a tomar en cuenta y parecería que un tratamiento con gramáticas HPSG (*head-driven*

phrase structure grammar) fuera más apropiado para capturar el contexto de los adjetivos y los adverbios. En HPSG el adjetivo, núcleo de la frase adjetival, tiene la característica de que modifica a un elemento que en este caso es un nombre. Puede verse como ejemplo en “acto programado”:

```
programado := word &
[ ORTH “programado”,
  HEAD adj_mas_sg,
  SPR < >,
  COMPS < >,
  MOD < [HEAD noun_mas_sg]>].
```

En el caso del adverbio, que es el núcleo de la frase adverbial, modifica a un elemento que es el verbo. Puede verse como ejemplo en “llegar tarde”:

```
tarde := word &
[ ORTH “tarde”,
  HEAD adv,
  MOD <[head verb]> ].
```

Puede plantearse entonces que el contexto para un adverbio es el verbo y para un adjetivo es el nombre, y de este modo realizar una anotación manual no ya sólo de adjetivos, sino de adjetivos en su contexto (nombre) y de adverbios en su contexto (verbos).

Enfoque Ciencia de la Información: temática

Según Barité *et al.* (2013) la organización del conocimiento es el

Área del conocimiento de formación reciente, que estudia las leyes, los principios y los procedimientos por los cuales se estructura el conocimiento especializado en cualquier disciplina, con la finalidad de representar temáticamente y recuperar la información contenida en documentos de cualquier índole, por medios eficientes que den respuesta rápida a las necesidades de los usuarios.

La Organización del Conocimiento se nutre de los aportes recibidos de la Informática, la Lingüística, la Terminología y la Ciencia de la Información.

[...]

El objeto de estudio de la Organización del Conocimiento es el conocimiento

socializado o registrado, y en lo que hace a Bibliotecología y Documentación, da cuenta del desarrollo teórico-práctico para la construcción, la gestión, el uso y la evaluación de clasificaciones, taxonomías, nomenclaturas, ontologías temáticas y lenguajes documentales. Asimismo, ampara el conjunto de conocimientos vinculados al análisis de información en general, considerando aspectos semánticos, cognitivos e informáticos.

La organización del conocimiento genera productos para categorizar los recursos a describir con distintos grado de estructuración y riqueza como las taxonomías, las clasificaciones, los tesauros y las ontologías. Estos productos son utilizados para categorizar los recursos informacionales de modo que el usuario pueda encontrar los recursos adecuados a su consulta.

Hjørland (2009) discute la “teoría del concepto”, el “átomo” en la descripción del conocimiento que utilizan estos productos y concluye que los conceptos deben ser comprendidos como significados negociados socialmente identificados estudiando los discursos más que como principios *a priori*. Afirma:

Cuando las palabras son colocadas en bases de datos, el contexto que proporciona cierta parte de su significado se pierde en parte. Desde la perspectiva de los paradigmas social e históricamente integrados, la futura mejora de la tecnología de recuperación de información está conectada a las posibilidades de reestablecer los contextos perdidos que determinan el significado de las palabras y los conceptos. (1530, traducción propia)

Realizar una clasificación de tuits por temática o por dominio es una cuestión previa al análisis de opinión porque el lenguaje y las entidades que aparecen están en relación a la temática. Si no se determina la temática existe mucho ruido en la recuperación. Este problema se menciona como trabajo futuro en Selva Castelló (2015) y en Vilares Calvo (2014).

Cuando la búsqueda es por términos específicos, por ejemplo, un modelo de celular, los resultados serán bastante adecuados, pero si se trata de un tópico más general empieza a producirse mucho ruido. Una manera de resolver la cobertura es utilizar sinónimos para la búsqueda.

La precisión está afectada por las palabras que tienen más de un significado y una manera de resolverlo es asociar la palabra al dominio en la que el significado tenderá a ser más estable. Determinar el dominio sobre el que se descarga la muestra de tuits afectará tanto la cobertura como la precisión. En muchos estudios de sentimiento en Twitter se trabaja sobre dominios específicos como cine, etc., partiendo ya de un conjunto de datos clasificados.

Vilares Calvo (2014) recopila algunas iniciativas de clasificación de tópi-

cos desde algunas clasificaciones en 12 tópicos (política, altruísmo, eventos, tecnología, juegos, idiomas, música, personalidad, películas, celebridades, estilo de vida y deportes) hasta 50 a través de la utilización de distintas aproximaciones. En ese panorama el *corpus* TASS ya mencionado provee en español su caracterización propia.

La definición de un tópico en lingüística corresponde a un tema o un asunto principal del que se habla, se explica, se predica o se comunica algo, en una frase o en un discurso y sobre el que se va aportando nueva información.

El ejemplo de un tuit como “arriba defensor...” constituye un caso claro de la necesidad de contexto, de un tema. En Uruguay una persona rápidamente asociaría que la palabra no alude a alguien que defiende sino al equipo uruguayo de fútbol de primera división Defensor. La extracción de tópicos permitiría poner Defensor bajo el tópico “Fútbol” y no bajo “Justicia”.

En el artículo de Gattani *et al.* (2013) se construye una base de conocimientos que tiene una serie de conceptos, subconceptos, un conjunto de instancias y un conjunto de relaciones entre los conceptos. Un planteo muy ingenioso es el uso de Wikipedia para mapear la base de conocimiento (conceptos e instancias). Se señala la importancia que en el caso de Twitter reviste el contar con una base de conocimientos en tiempo real ya que la evolución de los temas, los eventos y las instancias sobre las que versa el tuit es muy dinámica. La propuesta es mapear lo que se denominan “entidades” – que coinciden en gran parte con los sustantivos– a la base de conocimiento para encontrar el tópico.

Esta operación es conocida como *Named Entity Recognition* (NER, <https://nlp.stanford.edu/software/CRF-NER.shtml>) y es el etiquetado de secuencias de palabras en un texto que constituyen nombres de cosas tales como personas, nombres de empresas o proteínas, entre otros. Stanford proporciona un modelo para el inglés con tres categorías (personas, organizaciones, localizaciones) y en el 2014 se generó un modelo en español.

Esta necesidad de categorizar las entidades es cercana a la iniciativa de establecer características de los objetos. Ambas son necesidades de ubicar los objetos, los conceptos o las instancias en un sistema jerárquico, una taxonomía, un tesoro o una base de conocimientos que de alguna forma organiza los elementos en categorías y modela las relaciones entre ellos, fundamentalmente las relaciones jerárquicas.

El *Aspect Based Sentiment Analysis* (ABSA) plantea un análisis en el sentido de lo que Liu (2010) considera una opinión sobre un objeto y sus componentes en una estructura de árbol. Los aspectos (características o componentes) se definen como una combinación de una entidad (por ejemplo, restaurante) y un atributo de esa entidad (por ejemplo, precio) y hay una dife-

rencia entre la opinión sobre una entidad o sobre un aspecto de la entidad o un atributo de la entidad. Puedo tener una opinión positiva del restaurante, pero negativa de un aspecto de éste —el precio—. Por supuesto que la opinión del atributo participa en la opinión de la entidad, pero esa participación no es tan trivial, porque otros atributos también participan y componen la opinión sobre el objeto.

Esta estructuración de los componentes de los objetos en forma jerárquica a través de una relación “se compone de” o la granularización de los conceptos a través de la relación “es una” constituye un primer paso en la categorización, que va en el sentido de delinear un dominio, y la jerarquización, que permite ajustar el enfoque al nivel de detalle que se desee; ambos son pasos incipientes para disminuir el ruido y encaminarse a subir desde el nivel léxico al nivel conceptual en el sentido de un tesoro.

Un tesoro es un vocabulario controlado y estructurado formalmente, constituido por términos que guardan entre sí relaciones semánticas y genéricas de equivalencia, jerárquicas y asociativas. Se trata de un instrumento de control terminológico que permite convertir el lenguaje natural de los documentos en un lenguaje controlado, ya que representa, de manera unívoca, el contenido de estos, con el fin de servir tanto para la indización como para la recuperación de los documentos (Lapiente, 2007).

En el tesoro también se consideran términos compuestos, la estructura sintáctica según Lapiente (2007) es:

- sustantivo + adjetivo: documentos digitales
- sustantivo + sintagma preposicional: documentos de archivo
- sustantivo + sintagma aposicional: documentos RDF/XML

Hay una analogía entre las bases de conocimiento y los tesoros aunque el nivel expresivo de estos últimos es menor. En el tesoro la relación jerárquica impone una taxonomía de conceptos, clases y subclases que se corresponden a términos generales y términos específicos, esta taxonomía es un árbol donde los nodos son los conceptos y los arcos la relación “es una” que mapea a una subclase. Las relaciones de términos modelan asimismo una relación asociativa.

La primera etapa de la elaboración de un tesoro es la recopilación de los términos del dominio. Luego, esos términos empezarán a organizarse de manera que sea posible establecer una taxonomía, relaciones de clase y subclase (término tope y términos específicos), relaciones de véase además (asociativas), y dentro de los términos sinónimos elegir un término tope (el representante de la clase de equivalencia) y los otros términos sinónimos que

remitirán a este.

Las folksonomías parten de un punto distinto: para identificar un texto por su contenido se aportan etiquetas representativas de los temas que son términos sin ningún tipo de normalización y sin establecer relaciones. La riqueza de estas etiquetas o términos descriptivos reside en el hecho de que son aportados en las redes sociales y se reutilizan y comparten. Existen gestores de estos marcadores sociales en los que esta interacción se lleva a cabo,⁹ la cual, junto con la reutilización, puede relacionarse con la idea de la memética y la hipótesis de la replicación de las ideas en otros huéspedes en forma análoga a como sucede en la genética.

Para construir una taxonomía es necesario determinar la categoría de los términos extraídos. Si queremos subir un nivel en la taxonomía es necesario mapear los conceptos extraídos en un sistema jerárquico que vaya delineando un dominio con clases y subclases. Wikipedia es especialmente adecuada porque tiene un sistema en el que cada página –que corresponde a un concepto o a una instancia– tiene etiquetada la categoría a la que pertenece en un sistema jerárquico. Cada categoría en Wikipedia corresponde a su vez a una categoría más amplia. Una primera división es Ciencia, Arte, Naturaleza y Sociedad; luego, por ejemplo, la categoría Ciencia incluye las siguientes 13 subcategorías:

- Científicos (12 cat, 1 pág.)
- Ciencia por año (359 cat)
- Ciencia y tecnología por continente (7 cat)
- Ciencia y tecnología por país (178 cat)
- Wikiproyectos de ciencias (3 cat)
- Anexos: Ciencias (8 cat, 2 págs.)
- Ciencia y sociedad (10 cat, 10 págs.)
- Filosofía de la ciencia (15 cat, 122 págs.)
- Historia de la ciencia (19 cat, 37 págs.)
- Investigación científica (7 cat, 39 págs.)
- Problemas científicos (3 cat, 2 págs.)
- Protociencia (4 cat, 7 págs.)
- Ramas de la ciencia (5 cat, 2 págs.)

Otra herramienta que se puede usar para construir el árbol jerárquico del objeto o del concepto es Wordnet porque incorpora las relaciones semánticas (lemas, sinónimos, hiperónimos, hipónimos). En el caso de un discurso con las características del tuit el uso de un tesoro formal difícilmente se ajustará a esos conceptos socialmente entendidos pero gramatical y sintácticamente transgresores, por lo que se busca desarrollar una tendencia en delinear una

9 <https://del.icio.us/> Sitio que ofrece servicio de marcadores web.

jerarquía que se ajuste con gran flexibilidad.

PROTOTIPO

El prototipo realizado busca determinar la polaridad global de un tuit. No considera la gradación de intensidad. La idea es clasificar un tuit como positivo, negativo o neutro.

Se crea un lexicón que registra la polaridad semántica de algunas palabras (adjetivos y adverbios) a través de una anotación manual utilizando el contexto de los términos sobre los que se expresa el sentimiento (nombre y verbo respectivamente). El lexicón se crea *ad hoc* porque de esta forma se estará tomando en cuenta la localidad y la temporalidad del lenguaje usado en el tema en particular.

El lexicón tendrá como entradas los adjetivos y los adverbios como sujetos pasibles de detentar la polaridad. Los nombres referirán el tema de la opinión y su granularidad con lo cual es posible establecer un mapeo al vector de Liu del objeto de opinión y sus componentes. Algunos verbos tienen también una carga de polaridad y se pueden incluir en el lexicón de sentimientos. En el caso de los verbos en el lexicón se considera el lema ya que en español las variantes de persona y tiempo son muchas.

El lexicón no solamente incluye las palabras que expresan polaridad y su polaridad sino que además registra la palabra a la que están modificando, con el fin de establecer un contexto de aplicación de la misma. Se obtiene entonces un archivo para los adjetivos de la forma Adjetivo(nombre): polaridad, donde polaridad toma los valores 1 (positivo), 0 (neutro), -1 (negativo). Se obtiene un archivo para los adverbios de la forma Adverbio(verbo): polaridad, donde polaridad toma los valores 1 (positivo), 0 (neutro), -1 (negativo).

Los nombres además son utilizados para construir una terminología *ad hoc* o social, que recopila las entidades y sus componentes. Estos nombres son recopilados a ese efecto en un archivo que contiene los nombres que son utilizados en los tuits. Sin llegar a la construcción de un tesoro o una taxonomía se realiza la extracción de términos que ayuden a delimitar el dominio y generar una terminología (no normalizada) pero con la ventaja de tener un conjunto de términos usados en tiempo real. En cierto sentido lo que se modela es una terminología de uso, una terminología folksonómica podríamos decir.

La idea no es crear un tesoro o una base de conocimiento del tópico, sino establecer una recopilación de términos a modo de nube de palabras que sirvan para guiar la búsqueda de tuits y establecer el contexto para la

polaridad de los adjetivos y adverbios. No obstante, como una línea futura de trabajo puede plantearse la construcción no de una terminología *ad hoc*, sino de una taxonomía *ad hoc* mapeando Wikipedia o Wordnet.

La idea que se plantea parte de la hipótesis de que las palabras que se utilizan para expresar opinión son distintas en distintos dominios y en distintos lugares porque el lenguaje y los giros son elementos culturales muy localizados geográficamente y sensibles a los espacios temporales e incluso a lo que podemos denominar “tribus urbanas”. Por lo antedicho la tarea de generar un lexicón que contenga palabras de aplicación global es muy dificultosa y de aplicación muy restringida.

La nube de palabras que se genera *ad hoc* involucra conceptos, instancias y verbos que sirven para dos cosas: la primera, delinear la terminología del tópico; la segunda, contextualizar el aporte de la polaridad de las palabras.

El problema de la ambigüedad en el procesamiento del lenguaje natural es un problema general y en la expresión de sentimiento también se manifiesta. Si un lexicón de polaridad puede advertir sobre el objeto a que se califica tendrá mayor probabilidad de adecuarse, como en el siguiente ejemplo:

desestructurante (jugador) positivo
desestructurante (masaje) neutro
desestructurante (relación) negativo

El prototipo se generó como software libre y está disponible en *Sourceforge* (<https://sourceforge.net/projects/pln-polaridad-en-twitter/>). Debe señalarse que lo que se diseña es un prototipo pero no resulta apropiado para realizar un análisis porque hay un problema metodológico que proviene de extraer la polaridad de los términos en un conjunto de tuits y sobre el mismo conjunto asignar la polaridad con los términos ya relevados. Para poder extraer conclusiones se debería realizar la extracción de términos en un conjunto y realizar el análisis de opinión en otro conjunto. Para lograr un resultado plausible se debería trabajar con un volumen apropiado para verificar una cobertura adecuada. Otra posibilidad es incorporar el lexicón como diccionario complementario en los servicios como *Meaning Cloud*.

CONCLUSIONES

El lenguaje constituye un rasgo importante de la identidad cultural de las personas, e incluso dentro de una misma zona geográfica varía de acuerdo al

grupo social que lo utiliza, el cual tendrá su propia jerga, sus modismos, etc.

El uso del lenguaje en las redes sociales es coloquial, en formato muy informal, sin cuidado de la ortografía, la sintaxis o la gramática pero constituye un formato expresivo de amplio uso. La expresión en Twitter es particularmente fragmentada, constituye una expresión transmedia, hipervinculada, signo de un discurso que se fragmenta: empieza en un medio, continúa en otro, alude a un contexto que está no sólo en otro lugar, en otro medio, sino que además está expresado en formatos diversos, texto, imagen, video, etc., utilizando signos que se continúan en otros lugares y que dan pistas de un camino expresivo laberíntico.

La ciencia de la información tiene un amplio desarrollo en la categorización temática de los documentos y la categorización por sentimientos aparece como una línea de trabajo emergente que es de interés. En este sentido cabe señalar la existencia de un área de interdisciplina entre el procesamiento automático de lenguaje natural y la anotación manual de la polaridad semántica o la extracción de entidades por los profesionales de la información. La anotación o revisión manual aparece como una necesidad ya sea para la anotación de un corpus como para el diseño de distintas estrategias para contextualizar el sentimiento y es un área que puede desarrollarse en conjunto. Modelar la terminología de un dominio, establecer una base de conocimiento de un dominio, elaborar una taxonomía o un tesoro o incluso delinear los atributos de los objetos son operaciones muy cercanas y que también están en esa zona común.

La categorización del sentimiento reviste mayor complejidad que la categorización por tópicos y en particular con los métodos automáticos, ya que el sentimiento involucra elementos como el contexto, la ambigüedad, la ironía, que son fácilmente detectables por los humanos pero no sucede lo mismo en forma automática.

El mayor desafío es la adecuación en la determinación del sentimiento de las palabras y está vinculado con el contexto de las mismas. El contexto incluye los giros idiomáticos, la ironía, el sarcasmo y el lenguaje particular que se usa que puede tener distintos significados dependiendo del tema, el sitio o el emisor.

El aporte de contexto asignando el aspecto calificado (nombre) al calificador (adjetivo) –o adverbio (verbo)– es una iniciativa de mejoramiento y sería posible explorar en trabajos futuros si la polaridad varía en función del objeto al que se aplica o están involucrados otros aspectos no considerados. Otro trabajo a futuro sería abordar la polaridad de los verbos y sobre todo las locuciones verbales que tienen una alta carga de polaridad.

La propuesta de contextualizar se funda en la hipótesis de que un lexicón de polaridad para redes sociales no puede ser genérico, pues está vinculado

no sólo al tema, al lugar, al tiempo, al idioma, sino que varía rápidamente porque constituye una expresión de un idioma cambiante que es un organismo vivo en continua transformación.

Si bien es posible que exista una polaridad general o primordial como sostienen algunos autores, ésta es seguramente sobrepasada por las particularidades de la instancia concreta en un alto porcentaje. Lo que se propone como estrategia es la utilización de lexicones generales en combinación con la generación *ad hoc* de lexicones para cada situación en particular, lexicones que se entrenen con una herramienta como el prototipo que se desarrolló y se usen ajustando, entallando para lograr una adecuación a la situación concreta. Esta propuesta es practicable puesto que un servicio como *Meaning Cloud* permite combinar diccionarios propios.

Existe por otra parte la necesidad de modelar el objeto y sus atributos para ajustar los distintos elementos de la opinión sobre un objeto en concreto estableciendo el árbol de componentes de los objetos para realizar un modelado de los aspectos (ABSA), lo cual desde otro punto de vista constituye una taxonomía o un tesoro a microescala. Este modelado es un análisis que recoge tanto conceptualizaciones como terminología de uso en las redes sociales y permite un análisis más relevante de los objetos.

El análisis de sentimientos se puede complementar con otros análisis como los de la ubicación del emisor en la red de difusión. Identificar el emisor en un grafo permite interpretar la importancia de las opiniones, ya que la opinión de un personaje influyente –con muchos seguidores– llega a muchos otros nodos y resulta más significativa.

La extracción de entidades o terminología de los tuits así como su polaridad constituyen indicadores significativos que pueden ser tomados en cuenta para el desarrollo y mejora de productos, políticas o acciones.

REFERENCIAS

- Agarwal, A., B. Xie, I. Vovsha, O. Rambow y R. Passonneau. 2011. "Sentiment analysis of twitter data". *Proceedings of the workshop on languages in social media* (Association for Computational Linguistics): 30-38.
- Barité, M. et al. 2013. *Diccionario de organización del conocimiento: clasificación, indicación, terminología*, 5a. ed. Montevideo: PRODIC. <http://archivos.liccom.edu.uy/diccionario/Diccionario%20Definitivo%20%5B3-11-13%5D.html>
- Capurro, Rafael. 1987. "La hermenéutica y el fenómeno de la información". *Cuadernos de psicoanálisis freudiano* 8: 107-120. <http://www.capurro.de/herminf.html>
- Cervantes, Ofelia. 2016. *Curso Análisis semántico de redes sociales*. <https://eva.fing>.

- edu.uy/course/view.php?id=939
- Gattani, A., D. S. Lamba, N. Garera, M. Tiwari, X. Chai, S. Das, S. Subramaniam, A. Rajaraman, V. Harinarayan y A. Doan. 2013. "Entity extraction, linking, classification, and tagging for social media: a wikipedia-based approach". *Proceedings of the VLDB Endowment* 6 (11): 1126-1137. <http://www.vldb.org/pvldb/vol6/p1126-gattani.pdf>
- Hatzivassiloglou, V. y K. R. McKeown. 1997. "Predicting the semantic orientation of adjectives". *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*, 174-181. Association for Computational Linguistics.
- Hjørland, Birger. 2009. "Concept theory". *Journal of the American Society for Information Science and Technology* 60: 1519-1536.
- Hu, X., J. Tang, H. Gao y H. Liu. 2013. "Unsupervised sentiment analysis with emotional signals". *Proceedings of the 22nd international conference on World Wide Web*, 607-618. ACM.
- Kouloumpis, E., T. Wilson y J. D. Moore. 2011. "Twitter sentiment analysis: The good the bad and the omg!". *Icwsn* 11: 538-541.
- Lapiente, M. J. L. 2007. Hipertexto: *El nuevo concepto de documento en la cultura de la imagen*. http://www.hipertexto.info/documentos/web_tecnolog.htm
- Liu, C. 2010. *NLP Handbook*. Chicago: University of Illinois.
- Montejo Ráez, A., E. Martínez Cámara, M. T. Martín Valdivia y L. A. Ureña López. 2014. "Ranked wordnet graph for sentiment polarity classification in Twitter". *Computer Speech & Language* 28 (1): 93-107.
- Pang, B., L. Lee y S. Vaithyanathan. 2002. "Thumbs up?: sentiment classification using machine learning techniques". *Proceedings of the ACL-02 conference on Empirical methods in natural language processing* 10: 79-86. Association for Computational Linguistics.
- Pang, B. y L. Lee. 2008. "Opinion mining and sentiment analysis". *Foundations and trends in information retrieval* 2 (1-2): 1-135.
- Rojo, G. 1975. "Sobre la coordinación de adjetivos en la frase nominal y cuestiones conexas". *Verba* 2: 193-224.
- Saif, H., M. Fernandez, Y. He y H. Alani. 2014. "On Stopwords, Filtering and Data Sparsity for Sentiment Analysis of Twitter". *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. <https://pdfs.semanticscholar.org/68ff/581ba7162f5c2e809c72823bf418a58d4ff9.pdf>
- Selva Castelló, Javier. 2015. "Desarrollo de un sistema de análisis de sentimiento sobre Twitter". Tesis de doctorado.
- Taboada, M., J. Brooke, M. Tofloski, K. Voll y M. Stede. 2011. "Lexicon-based methods for sentiment analysis". *Computational linguistics* 37 (2): 267-307.
- Urizar, X. S. e I. S. V. Roncal. 2013. "Elhuyar at TASS 2013". *Proceedings of the Workshop on Sentiment Analysis at SEPLN (TASS 2013)*, 143-150. <http://www.sepln.org/workshops/tass/2013/papers/tass2013-submission3-Elhuyar.pdf>
- Vilares, D., M. A. Alonso y C. Gómez Rodríguez. 2015. "On the usefulness of lexical and syntactic processing in polarity classification of Twitter messages". *Journal of the Association for Information Science and Technology* 66 (9): 1799-1816.
- Vilares Calvo, David. 2014. *Análisis de contenidos en Twitter: clasificación de mensajes e identificación de la tendencia política de los usuarios*. España: Universidad de Coruña, Facultad de Informática.

Para citar este texto:

Temesio-Vizoso, Silvana Grazia. 2018. “La opinión en textos con un enfoque interdisciplinar: propuesta de contexto y léxico *ad hoc*”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 73-99.

<http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57863>

DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57863>

Gestor de contenidos de código abierto para archivos digitales sonoros que preservan materiales de investigación

Perla Olivia Rodríguez-Reséndiz*

Joséphine Simonnot**

Dafne Citlalli Abad Martínez*

Artículo recibido:

9 de agosto de 2018

Artículo aceptado:

14 de septiembre de 2018

Artículo de investigación

RESUMEN

Los institutos y centros de investigación generan una gran cantidad de documentos sonoros de origen digital como resultado de proyectos de investigación en ciencias sociales y humanidades. La acumulación de estos documentos digitales es un problema porque se carece de herramientas de preservación y el riesgo de pérdida es alto. Como una alternativa para preservar este tipo de documentos se presentan los alcances y funcionalidades de Telemeta, gestor de contenidos, para la preservación de contenidos sonoros.

* Instituto de Investigaciones Bibliotecológicas y de la Información
Universidad Nacional Autónoma de México, México perlaolivia@gmail.com
dafne@iibi.unam.mx

** Centre de Recherche en Ethnomusicologie, París, Francia
josimonnot@gmail.com

Palabras clave: Archivo Digital; Preservación Digital, Archivo Sonoro

Open source content management for digital sound archives that preserve research materials

Perla Olivia Rodríguez-Reséndiz, Joséphine Simonnot and Dafne Citlalli Abad-Martínez

ABSTRACT

Institutes and research centers, through research projects in social sciences and humanities, generate a large number of digital sound documents. The accretion of these digital documents come with high risk of loss because of a lack of preservation tools. This paper describes the scope and functionalities of the Telemeta content manager as an alternative for preservation of sound files.

Keywords: Digital Archive; Digital Preservation; Sound Archive

INTRODUCCIÓN

Se ha privilegiado el uso del término *repositorio digital* para denominar al sistema informático en el cual se almacena y organiza información de publicaciones científicas. El movimiento de Acceso Abierto impulsó la creación de repositorios digitales (Keefer, 2008) de materiales impresos (revistas, libros y documentos administrativos); tema que en la última década ha sido ampliamente documentado (Barnard, Delgado y Voutssas, 2014; Alcaráz *et al.*, 2014; Lara, Serralde y Maniega, 2013, y Tramullas y Garrido, 2006). En contraste, las aportaciones para el diseño y creación de archivos digitales que preserven materiales sonoros son incipientes, si bien es evidente la necesidad de contar con herramientas de preservación digital de este tipo de documentos, que son frágiles y cuyo riesgo de pérdida es alto (IASA, 2011).

Las fonotecas, archivos sonoros de medios de comunicación, empresas de producción sonora, discográficas, así como centros de investigación, entre

otros lugares, requieren garantizar la permanencia de contenidos digitales obtenidos como resultado de la digitalización o bien cuyo origen es digital y, para ello, deben contar con archivos digitales. La creación de archivos digitales para preservar colecciones sonoras es uno de los desafíos más complejos que deben encarar las instituciones de la memoria, sobre todo aquellas que no cuentan con infraestructura, personal ni los recursos económicos de forma permanente.

Hace más de una década Bradley (2007) reconoció que era necesario desarrollar sistemas de almacenamiento y preservación en código abierto para que las instituciones pequeñas también preservaran sus colecciones. Se propuso crear un sistema simple y sustentable de código abierto como una alternativa de preservación digital (Bradley, 2007). Desde entonces, la comunidad científica y profesional ha mostrado interés en el uso del software en código abierto (*open source*) como una alternativa para preservación digital. A través del gestor de contenidos se controla el ciclo de vida de los objetos digitales (Crockett, 2016) y se articulan los procesos de identificación, organización, conservación y acceso en un archivo digital. La puesta en marcha de gestores de contenido es la primera etapa de creación de un archivo digital.

Este artículo presenta los alcances y funcionalidades de Telemeta como gestor de contenidos para la preservación de contenidos sonoros. Inicia con la metodología utilizada en la investigación y después se diserta sobre la relevancia de los archivos sonoros en la investigación científica. Más adelante, se examina el término *archivo* en los dominios analógico y digital. Continúa con la definición y descripción de los modelos de gestores de contenidos utilizados en los archivos sonoros y se presentan los resultados observados durante la primera etapa de instalación de Telemeta como gestor de contenidos. Para terminar, se ofrecen las conclusiones.

METODOLOGÍA

Ante la acumulación de contenidos sonoros obtenidos como resultado de proyectos de investigación realizados en el Instituto de Investigaciones Bibliotecológicas y de la Información (IIBI) de la Universidad Nacional Autónoma de México (UNAM), se propuso la creación de un archivo digital para la preservación de estos materiales. Este proyecto se desarrolla en el marco de la colaboración de investigadores de la UNAM con el Centre de Recherche en Ethnomusicologie (CREM) de Francia y como parte del Proyecto PAPIIT IT-400118 “Creación y desarrollo de archivos digitales multimedios (sonoros, audiovisuales y fotográficos) con *open source*. Una propuesta de transferencia

tecnológica para la preservación digital de las colecciones de los pueblos originarios de México de la DGAPA UNAM”.

Este artículo presenta los alcances y funcionalidades observados durante la primera etapa de instalación del gestor de contenidos de código abierto a partir del cual se establecerán las bases para la creación de un archivo digital que preserve las grabaciones de los pueblos originarios de México.

La metodología de este trabajo es teórica y aplicada. En cuanto a la investigación teórica se identificaron y analizaron publicaciones científicas para establecer el marco conceptual e histórico de los gestores de contenido como componentes de archivos digitales sonoros. Se describieron los modelos de gestores de contenidos digitales y se eligió trabajar con soluciones tecnológicas basadas en código abierto. La investigación aplicada se basó en observar, analizar y describir los alcances y funcionalidades del gestor de contenidos durante la primera etapa de instalación.

Los fondos iniciales incorporados en el gestor de contenidos materiales pertenecen a grabaciones de campo de dos grupos indígenas de México: los rarámuris y los purépechas.¹ Estas grabaciones se obtuvieron de los proyectos PAPIIT-IN402016 “Preservación digital de documentos sonoros y audiovisuales de origen digital” y PAPIIT- IG400417 “Fortalecimiento, organización y preservación de la información originaria: bases para construir un modelo de biblioteca indígena de la comunidad purépecha del Municipio de Cherán”, ambos auspiciados por la DGAPA (Dirección General de Asuntos del Personal Académico) de la UNAM.

LOS ARCHIVOS SONOROS EN LA INVESTIGACIÓN CIENTÍFICA

Los investigadores de disciplinas humanísticas y sociales fueron pioneros en utilizar las tecnologías para grabar y reproducir el sonido. Gracias al uso del fonógrafo en la investigación de campo grabaron música, danza, rituales y tradición oral como fuentes de información para el estudio de los fenómenos

1 Los rarámuris, también llamados tarahumaras, viven en condiciones de extrema pobreza entre las montañas de Chihuahua, en la Sierra Madre Occidental de México, en 17 municipios entre los que destacan por su densidad Guachochi, Norogachi, Urique, Batopilas, Balleza, Carichi, Guazapares, Bocoyna, Guerrero, Maguarichi, Nonoaca y Cuauhtémoc (CDI, 2017). Los *p'urbépecha*, también denominados tarascos, viven en la región lacustre y montañosa del estado de Michoacán. En la actualidad los *p'urbépecha* afrontan básicamente tres tipos de problemas: tala inmoderada de bosques, deterioro ambiental y seguridad social (Amézcuea y Sánchez, 2015). Uno de los rasgos más destacados del pueblo *p'urbépecha* fue la instauración del Municipio Autónomo de Cherán en 2012, con lo que se pusieron “a prueba las viejas estructuras de la organización tradicional y permitió la emergencia de nuevos comportamientos organizacionales como la movilización de las mujeres de defensa frente a los embates criminales de la delincuencia organizada” (Amézcuea y Sánchez, 2015: 139).

sociales. Por ello, el desarrollo y florecimiento de la lingüística, etnomusicología y la antropología está vinculado al uso de las tecnologías para la grabación sonora (Schüller, 2008). Como resultado de la acumulación de materiales grabados en trabajos de investigación científica se fundaron los primeros archivos sonoros en el mundo: en 1899 en Viena, en 1900 en París y Berlín y en 1908 en San Petersburgo (Schüller, 2008).

Las posibilidades de grabación de campo se enriquecieron y diversificaron con la incorporación de la tecnología digital. La grabación sonora es, en la actualidad, un medio para generar recursos documentales para la investigación. Además, la grabación digital es una alternativa de salvaguarda frente al riesgo de pérdida de cuando menos el 50 % de los idiomas que se hablan en la actualidad (Unesco, 2003b). Los investigadores de ciencias sociales y humanidades han emprendido, en atención a esta crítica situación, sendos proyectos de investigación para documentar la música, danzas, rituales, narraciones orales y paisaje sonoro, entre otras expresiones, de los pueblos originarios del mundo.

DEL ARCHIVO SONORO ANALÓGICO AL DIGITAL

A partir de 1980 y, en reacción a la “Recomendación sobre la salvaguarda y conservación del patrimonio audiovisual” (Unesco, 1980), los archivos sonoros fueron identificados como patrimonio intangible, legado, producto cultural y testimonio (ICA, 2018). Hasta el siglo pasado, el término *archivo sonoro* se utilizó para denominar el lugar, la institución de resguardo y el soporte en que fue registrado el sonido (Edmondson, 2018). El archivo sonoro como lugar de la memoria es el espacio donde se preservan las grabaciones sonoras, se protege la herencia audible de un pueblo y, coexisten, desde la perspectiva de Derrida (1995), el saber y el poder. Además, con esta expresión se nombra a la institución de la memoria en la cual se resguardan documentos sonoros (Edmondson, 2018). Los principales archivos sonoros son las fonotecas de alcance nacional, los centros de documentación musical, las fonotecas de radio, los archivos sonoros de investigación y especializados, entre otros. El vocablo también ha sido utilizado para hacer referencia al artefacto en el cual ha sido fijado el sonido (cilindros, discos, cintas de carrete abierto, casetes, discos compactos, *digital audio tape*, *digital versatile disc*, Blu-ray, *linear tape open*, disco duro, entre otros).

Los archivos sonoros comenzaron a utilizar soportes digitales en la década de 1980 (Prentice, 2016). El disco compacto fue el primer soporte digital que se empleó en los archivos sonoros como medio de resguardo (Bradley,

2006). En algunos archivos, la noción *archivo digital* fue empleada para denominar a las copias de las grabaciones que se resguardaron en discos compactos.

La locución “archivo digital sonoro” se empleó por primera vez en la década de 1990, en la Conferencia Anual de la IASA (*International Association of Sound and Audiovisual Archives* / Asociación Internacional de Archivos Sonoros y Audiovisuales) organizada en Viena para hacer referencia a las soluciones tecnológicas necesarias para almacenar los contenidos obtenidos con la digitalización de colecciones analógicas (Schüller, 2008).

En los primeros años del siglo XXI, el archivo digital sonoro ha sido definido como un sistema de información y como un conjunto integrado de sistemas (Ojeda Castañeda, 2008) a través de los cuales se articulan los procesos documentales que intervienen durante el ciclo de vida de los objetos digitales, los roles, flujos de trabajo y usuarios que intervienen en la preservación digital (Rodríguez, 2016). El archivo sonoro es un término polisémico cuyos alcances y definiciones han cambiado a lo largo del tiempo. En este trabajo se emplea el término *archivo digital sonoro* para hacer referencia al sistema de información que vincula, a través del uso de herramientas tecnológicas, los procesos documentales, flujos de trabajo y usuarios que intervienen en el ciclo de vida para la preservación digital.

GESTIÓN DE COLECCIONES DIGITALES

El término *gestión de contenidos* tiene diferentes usos y su concepción ha cambiado a lo largo del tiempo. Se empleó para denominar a los programas informáticos de edición colaborativa de sitios web. También se ha utilizado para nombrar a las herramientas, métodos y procesos de la gestión documental tradicional (Éito-Brun, 2013: 378). De forma más reciente se define como la “integración progresiva de las tecnologías para gestionar información y datos no estructurados” (Éito-Brun, 2013: 378).

En la década de 1990 se instalaron los primeros sistemas de gestión de colecciones digitales con funciones de ingesta, selección, descripción, almacenamiento, vinculación, base de datos para catalogar, así como herramientas de búsqueda y recuperación de materiales (Monson, 2017 y Rodríguez Pallares, 2015). Los gestores de contenidos son una nueva herramienta tecnológica que surgió con la digitalización (Muñoz, Meana y Sáez, 2014) y ante la necesidad de manejar y administrar información y datos no estructurados. Los sistemas de gestión se han denominado con diferentes nombres: Sistema de gestión de contenidos, Sistema de gestión de activos digitales,

Repositorio digital y Sistema de gestión y almacenamiento masivo digital. En el ámbito de los archivos sonoros, las soluciones pioneras fueron los Sistemas de gestión y almacenamiento masivo digital (SGAMD). Estos fueron concebidos para almacenar los contenidos obtenidos después de la digitalización, obtener copias automatizadas y apoyar la gestión de contenidos a largo plazo (Murray, 2014); con ello la forma tradicional de trabajo en los archivos sonoros se transformó (Cavaglieri, 2009) en la primera década del siglo XXI. Los SGAMD cohabitaron con el archivo analógico. En un principio, no se dimensionó el alcance que éstos podrían tener; sin embargo, de forma progresiva ganaron importancia dentro de la institución.

El gestor de contenidos es un componente del archivo digital. De acuerdo con el *Open Archival Information System (OAIS)* es el conjunto de servicios y funciones necesarias para generar, mantener y hacer accesibles la información descriptiva, que identifica y documenta los fondos, los datos administrativos del catálogo y el registro estadístico de los datos contenidos (IASA, 2011). “La gestión y administración de datos resulta decisiva en un sistema de archivo sostenible, puesto que garantiza que los ficheros preservados y accesibles puedan ser adecuadamente hallados e identificados” (IASA, 2011: 126).

Monson (2017) ha señalado que existen tres modelos de gestores de contenidos digitales: diseño propio, propietario o comercial y de código abierto. El diseño propio es el desarrollo basado en las características y necesidades de las instituciones. Las instituciones que cuentan con un amplio presupuesto y un sólido equipo de tecnologías de información pueden poner en marcha un sistema de gestión de contenidos de esta naturaleza.

Los sistemas propietarios o comerciales se asientan en la venta de licencias, el pago sistemático de servicios de mantenimiento y actualizaciones del uso de software. Los clientes cuentan con soporte técnico. Esta alternativa es relativamente fácil de poner en marcha en instituciones que no disponen de personal en el área de tecnología. El riesgo más grande de esta modalidad es que se interrumpa el suministro de recursos económicos y con ello se limite la continuidad del servicio. Este tipo de software supone “gastos continuos y el peligro de quedar encasillados en un sistema propietario del que es difícil escapar” (IASA, 2011:133).

Las soluciones de código abierto generalmente están disponibles en internet de forma gratuita. Este tipo de opciones “tienen a su favor que la adhesión a estándares y marcos abiertos que permiten la extracción o migración del contenido” (IASA, 2011: 133). Bradley (2007) fue pionero en proponer el uso del software de código abierto, de bajo costo, para crear sistemas de archivo digital abierto que conserven objetos en diferentes formatos (audio,

texto, imagen en movimiento y fotografía). Este tipo de software es utilizado, sobre todo, en el sector cultural y educativo. Dependiendo de la comunidad de usuarios, el software puede ser modificado, enriquecido y contar con nuevos desarrollos y funcionalidades (Monson, 2017). La desventaja es que, a pesar de contar con ayuda de las comunidades de código abierto, el mantenimiento depende del usuario. Para afrontar esta situación pueden contratarse proveedores comerciales que ofrezcan asistencia técnica (IASA, 2011). En la *Tabla 1* se ofrece una relación de gestores de contenidos digitales de código abierto cuya configuración admite la incorporación de formatos de preservación (abiertos y sin compresión) de documentos digitales sonoros.

En los últimos años, la línea entre los sistemas propietarios y de código abierto ha empezado a ser menos clara. “Una tercera parte de los vendedores ahora ofrecen soporte comercial a las instituciones que usan *open source*. El rango de servicios que ofrecen comprende desde soporte técnico por hora, hasta un paquete de servicios que incluye almacenamiento en la nube y mantenimiento. Con ello, es posible combinar los beneficios de las soluciones propietarias con el *open source* en una solución de software” (Monson, 2017: 113).

| Sistemas <i>open source</i> | Año de creación |
|-----------------------------|------------------------------------|
| Greenstone | Finales 1990 |
| Dspace | 2002 |
| Fedora | Empezó en 1997 y se liberó en 2003 |
| Islandora | 2006 |
| Hydra | 2008 |
| Telemeta | Telemeta |

Tabla 1. Gestores de contenidos digitales de código abierto

Fuente: elaboración propia

RESULTADOS

El archivo digital, como se ha señalado, es un sistema de información cuya arquitectura tecnológica se sustenta en el gestor de contenidos y el almacenamiento digital distribuido. En este trabajo se aborda solamente el gestor de contenidos como herramienta para la organización, identificación, conservación, acceso y reaprovechamiento documental. Se eligió trabajar con Telemeta por ser un software creado en código abierto y porque fue concebido para que los investigadores de las ciencias sociales, sobre todo los antropólogos,

etnomusicólogos y lingüistas, puedan preservar los materiales que graban como resultado de sus trabajos de investigación de campo.

El desarrollo de Telemeta comenzó en 2007, como parte de la colaboración entre la compañía Parisson y el CREM y se puso en marcha en 2011 (Fillon y Pellerin, 2017). Este software está disponible bajo la Licencia de Software Libre GNU Affero General Public License v3.0 y es una herramienta para gestión de contenidos multimedia. Telemeta ha sido definida como una plataforma de audio web escalable, que permite realizar copias de seguridad, indexar, transcodificar, analizar, compartir y visualizar cualquier formato de archivo de audio o video digital de estándares abiertos de la web (Fillon *et al.*, 2014).

Derivado del proceso de instalación de Telemeta se han identificado los siguientes alcances y funcionalidades para la construcción de un archivo digital que preserve grabaciones sonoras de pueblos originarios de México:

- a) Conservación. El gestor de contenidos fue creado para conservar y dar acceso a materiales sonoros creados en formatos de preservación, es decir, abiertos y sin compresión. Esta funcionalidad es la primera que fue considerada en la puesta en marcha de esta herramienta; no obstante, también se pueden almacenar materiales audiovisuales, fotográficos e iconográficos (imágenes digitalizadas, escaneos de folletos y notas de campo, etc.), hipervínculos e información biográfica sobre el recolector (Fillon *et al.*, 2014 y Khoury y Simonnot, 2014). Esta alternativa es de gran apoyo para el trabajo de investigación puesto que, además de grabaciones sonoras, durante el trabajo de campo se recopilan grabaciones audiovisuales, fotografías y notas de campo que también deben ser preservadas.
- b) Acceso y consulta. Los materiales pueden ser consultados y descargados en formatos sin compresión como WAV, o bien de compresión, como FLAT, OGG, MP3. El acceso para consulta o descarga de los materiales se configura de acuerdo con los permisos que tienen los usuarios y con base en los derechos de autor de cada una de las grabaciones. Es preciso señalar que en la creación del archivo digital se busca privilegiar el acceso abierto a partir de contar con el consentimiento de los pueblos originarios. En la *Figura 1* se muestra la interfaz de acceso al gestor de contenidos.



Figura 1. Interfaz de acceso de Telemeta

- c) A través del gestor de contenidos se puede ofrecer el acceso *in situ* o bien en línea a través de la página web. Para el acceso a las colecciones, Telemeta contiene reproductores de audio y video a través de Timeside (*audio processing framework*). Este procesador puede mostrar el sonido gráficamente en forma de onda y como espectrograma.
- d) Identificación. La estructura de los metadatos es interoperable, se basa en los protocolos Dublin Core y OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*). Se pueden incluir detalles sobre el lugar donde se realizó la grabación, los instrumentos musicales, la población, el título de la pieza musical, los elementos culturales relacionados con el elemento musical, el depositante, el coleccionista, el año de la grabación y el año de publicación de documentos que describen el trabajo (Fillon *et al.*, 2014). Si bien la estructura de Telemeta permite mantener los metadatos separados de sus objetos digitales, éstos se resguardan de manera predeterminada en una misma ubicación física. Para tener un almacenamiento distribuido, es necesario reconfigurar el sistema.
- e) El gestor de contenidos tiene herramientas de edición social. Es una plataforma colaborativa (Fillon y Pellerin, 2017) a través de la cual usuarios especializados como los investigadores pueden establecer marcas y anotaciones en las grabaciones y compartirlas con sus colegas. Estas anotaciones son accesibles desde la página web del archivo de sonido y están indexadas a través de la base de datos (Fillon *et al.*, 2014). Con ello, se pueden enriquecer los metadatos obtenidos durante la grabación de campo.

- f) Organización. La organización intelectual de los materiales tiene cuatro niveles: fondos, *corpus*, colecciones e ítems. Estos niveles se pueden adecuar dependiendo de las necesidades de cada archivo. Los fondos iniciales que se incorporaron son Rarámuris y Purépechas. El fondo Rarámuri está formado por dos *corpus*: Continente Rojo y Semana Santa Rarámuri. Continente Rojo agrupa un conjunto de colecciones de las fiestas tradicionales, danzas, música, rituales y entrevistas a diversos pobladores de la comunidad rarámuri que, desde el año 2010, grabó la antropóloga francesa Sylvie Marchand. Los materiales sonoros (entrevistas, paisajes sonoros, danzas, música y rituales) fueron registrados con tecnología digital en formato WAV de 24 bits/ 48 khz. Además, se cuenta con grabaciones audiovisuales (entrevistas, fiestas tradicionales, rituales, danzas, etc) grabadas en formato AVCHD y MOV. Las producciones poseen alta calidad artística y técnica. Los materiales grabados por Marchand son documentos de origen digital, que hasta ahora no han sido preservados en alguna institución de la memoria en México. Por ello, se le propuso a la investigadora participar en este proyecto de investigación destinado a crear estrategias de salvaguarda de esta herencia documental. Por su parte, el corpus Semana Santa Rarámuri reúne grabaciones sonoras, audiovisuales y fotografías del trabajo de campo realizado por investigadores del IIBI en la comunidad de Norogachi, Chihuahua, en 2016.
- En el fondo Purépecha se incorporó el corpus “Cien sonidos del paisaje sonoro purépecha” que reúne grabaciones sonoras de personajes emblemáticos que dan cuenta de la historia y tradiciones de Cherán, fiestas tradicionales, narración de tradiciones, sonoridad de las lenguas habladas en la comunidad, entornos naturales, oficios tradicionales (zapateros, vendedores de frutas, etc), música tradicional, entre otros sonidos de relevancia para la comunidad. Los sonidos grabados fueron identificados por parte de la comunidad e incorporados al archivo digital para su resguardo. Estos materiales a su vez formarán parte de la biblioteca modelo de Cherán.
- g) Para la instalación del gestor de contenidos digitales se requieren conocimientos de administración de sistemas operativos en Linux. El gestor, desde su versión 1.6, incorpora un paquete de Docker que asocia todas las aplicaciones y módulos necesarios para llevar a cabo la instalación. La estructura del software es Modelo-Vista-Controlador que separa la interfaz del usuario, la gestión de las solicitudes y la administración de los datos contenidos. Utiliza como lenguaje de programación Python, el framework Django e incorpora MySQL como

motor de la base de datos. El diseño de Telemeta incluye tecnologías compatibles con HTML5, característica que lo hace funcional para diferentes navegadores web.

CONCLUSIONES

Los científicos sociales y de humanidades registran mediante tecnologías digitales materiales sonoros que contribuyen al estudio y a la solución de problemas sociales. Una de estas vertientes es la documentación de las lenguas originarias que se encuentran en riesgo de desaparecer. La acumulación de estas grabaciones trae consigo grandes desafíos. Se necesita conservar y dar acceso a grabaciones sonoras y materiales asociados (videos, fotografías y notas de campo), y para ello es necesario crear archivos digitales.

El archivo digital sonoro es el sistema de información que vincula, a través del uso de herramientas tecnológicas, los procesos documentales, flujos de trabajo y usuarios que intervienen en el ciclo de vida para la preservación digital. El gestor de contenidos es la herramienta base para la preservación digital a largo plazo. La instalación de gestores de contenido de código abierto es una alternativa para la preservación de los materiales que han sido grabados como resultado de la actividad científica.

La puesta en marcha de Telemeta como gestor de contenidos de código abierto sienta las bases para la creación de un archivo digital sonoro de acceso abierto y es resultado del proyecto de investigación e innovación tecnológica basado en la colaboración de investigadores de México y de Francia.

La colaboración es necesaria en la preservación digital. La tecnología compartida, el almacenamiento distribuido y las estrategias de preservación conjunta son acciones para salvaguardar contenidos digitales, especialmente para regiones como América Latina, donde los recursos económicos son escasos y no se asignan de forma sistemática, son muy pocas las investigaciones que se realizan en este campo y donde aún es necesario formar y actualizar a los profesionales que trabajan en los archivos sonoros.

REFERENCIAS

- Alcaráz, Rubén, Sergi Montes, Lluís Vicente, Merche Blanco, Elena Álvaro, Laia Amorós, Pablo Casas, Laura Nieto y Marc Reig. 2014. "El Archivo de la Palabra: contexto y proyecto del repositorio audiovisual del Ateneu Barcelonés". *BID Textos Universitaris de biblioteconomia i documentació* 33 (Barcelona, España).

- Amézcu, J. y G. Sánchez. 2015. *Pueblos indígenas de México en el siglo XXI. P'urhépecha*. México: Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CDI).
- Barnard, A., A. Delgado y J. Voutssas. 2014. *Los archivos digitales. Una visión integradora*. México: Archivo Histórico Benemérita Universidad Autónoma de Puebla.
- Bradley, K. 2006. *Riesgos asociados con el uso de los discos compactos (CDs) y videodiscos (DVDs) como medios confiables de almacenamiento para colecciones de archivo*. México: Fonoteca Nacional.
- Bradley, K. 2007. *Hacia un sistema de almacenamiento y preservación en código abierto. Recomendaciones respecto a la implementación de un sistema de preservación de archivos digitales y temas en torno al desarrollo de software*. México: Unesco.
- CDI (Comisión Nacional para el Desarrollo de los Pueblos Indígenas). 2017. *Indicadores socioeconómicos de los pueblos indígenas de México*. México. Fecha de consulta: 12 de marzo de 2018 <https://www.gob.mx/cdi/articulos/indicadores-socioeconomicos-de-los-pueblos-indigenas-de-mexico-2015-116128?idiom=es>
- Cavaglieri, S. 2009. "Criteria to consider in the definition of digital mass storage systems", en *La salvaguarda del patrimonio sonoro y audiovisual: un reto mundial. Memorias del Cuarto Seminario Internacional de Archivos Sonoros y Audiovisuales*, Perla Rodríguez (coord.), 151-164. México: CONACULTA, Fonoteca Nacional de México.
- Crockett, Margaret. 2016. *The no-nonsense guide to archives and recordkeeping*. Londres: Facet Publishing.
- Derrida, J. 1995. *Archive Fever*. Chicago: The University of Chicago Press.
- Edmondson, R. 2018. *Filosofía y principios de los archivos audiovisuales*, 3a. ed. México, Unesco, UNAM-UASLP.
- Éito-Brun, Ricardo. 2013. "Madurez de la gestión de contenidos, ¿sinónimo de desgaste o de oportunidades?". *El profesional de la información* 22 (5) (septiembre-octubre): 377-380.
- Fillon, Thomas, Josephine Simonnot, Marie-France Mifune, Stéphanie Khoury, Guillaume Pellerin, Maxime Le Coz, Estelle Amy De la Bretèque, David Doukhan y Dominique Fourer. 2014. "Telemeta: An open-source web framework for ethnomusicological audio archives management and automatic analysis", *Proceedings of the 1st International Workshop on Digital Libraries for Musicology*, ACM: 1-8.
- Fillon, T. y G. Pellerin. 2017. "A collaborative web platform for sound archives management and analysis", *Proceedings of 3rd Web Audio Conference*, London.
- ICA (International Council on Archives). 2018. *¿Qué es un archivo?* Fecha de consulta: 22 de marzo de 2018. <https://www.ica.org/es/%C2%BFqu%C3%A9-son-los-archivos>
- IASA (International Association of Sound and Audiovisual Archives). 2011. *Directrices para la producción y preservación de objetos digitales de audio*. IASA-TC04. España: AEDOM.
- Keefer, A. 2008. "Los repositorios digitales universitarios y los autores". *Revista Analés de Documentación* (España): 205-214.
- Khoury, S. y J. Simonnot. 2014. "Applications and implications of digital audio databases for the field of ethnomusicology: A discussion of the CNRS — Musée de l'Homme sound archives". *First Monday* 19 (10-6). doi: <http://dx.doi.org/10.5210/fm.v19i10.5554>

- Lara, P., E. Serralde y E. Maniega. 2013. "Evolución de los repositorios documentales. El caso SOCIALNET". *El Profesional de la Información* 22 (5) (septiembre-octubre).
- Ojeda Castañeda, G. 2008. *Los archivos audiovisuales en las redes digitales de comunicación para la educación y la cultura*. Informe de Investigación y Documentación Analítica. Serie de Informes CNICE. Ministerio de Educación y Ciencia de España. Fecha de consulta: 20 de mayo de 2018. <http://ares.cnice.mec.es/informes/13/contenido/34.htm>
- Prentice, Will. 2016. "Digital preservation of audio content", en *Managing digital cultural objects*. Chicago: Neal Schuman.
- Monson, J. 2017. *Getting started with digital collections. Scaling to fit your organization*. Chicago: American Library Association.
- Muñoz, P., S. Meana y S. Sáez. 2014. "Cinco años de experiencia digital en los Servicios Informativos de TVE: una nueva gestión de contenidos". *El profesional de la información* 23 (1) (enero- febrero): 72-79.
- Murray, K. 2014. *Audio for Eternity: Schüller and Häfner Look Back at 25 Years of Change*. Fecha de consulta: 20 de marzo de 2018. <https://blogs.loc.gov/thesignal/2014/11/audio-for-eternity-schuller-and-hafner-look-back-at-25-years-of-change/>
- Rodríguez Pallares, M. 2015. "Integración y visibilidad de la gestión de contenidos, dos retos por alcanzar en el modelo de comunicación radiofónica digital. Análisis de los casos de SER, COPE y RNE". *Revista General de Información y Documentación* 25 (1): 191-213. doi: http://dx.doi.org/10.5209/rev_RGID.2015.v25.n1.48987
- Rodríguez, P. 2016. "El OAIS en la preservación digital de archivos sonoros". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 30 (70): 197-220. <http://dx.doi.org/10.1016/j.ibbai.2016.10.009>
- Schüller, D. 2008. *Audiovisual research collections and their preservation*. Amsterdam European Comission on Preservation and Acces.
- Tramullas, J. y P. Garrido. 2006. "Software libre para repositorios institucionales: propuestas para un modelo de evaluación de prestaciones". *El profesional de la información* 15 (3) (mayo-junio).
- Unesco. 1980. *Recomendación sobre la salvaguardia y la conservación de las imágenes en movimiento*. Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Fecha de consulta: 30 de mayo de 2018. http://portal.unesco.org/es/ev.php-URL_ID=13139&URL_DO=DO_TOPIC&URL_SECTION=201.html
- Unesco. 2003a. *Directrices para la preservación del patrimonio digital*. París. Fecha de consulta: 16 de marzo de 2018. <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>
- Unesco. 2003b. *Vitalidad y peligro de desaparición de las lenguas*. París: Unesco, Grupo Especial de Expertos sobre Lenguas en Peligro de Extinción.

Agradecimientos

Proyecto PAPIIT IT- 400118 “Creación y desarrollo de archivos digitales multimedios (sonoros, audiovisuales y fotográficos) con *open source*. Una propuesta de transferencia tecnológica para la preservación digital de las colecciones de los pueblos originarios de México de la DGPA UNAM”.

Para citar este texto:

Rodríguez-Reséndiz, Perla Olivia, Joséphine Simonnot y Dafne Citalli Abad-Martínez. 2018. “Gestor de contenidos de código abierto para archivos digitales sonoros que preservan materiales de investigación”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 101-115.
<http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.58005>

The impact of open access citation and social media on leading top Information Science journals

Paulo Roberto Cintra*

Ariadne Chloe Furnival**

Douglas Henrique Milanez***

Artículo recibido:
20 de noviembre de 2017

Artículo aceptado:
16 de agosto de 2018

Artículo de investigación

ABSTRACT

This study assesses possible advantages of open access can bring to the area of Information Science, using bibliometric citation indicators and Altmetric scores to analyze sixteen hybrid scientific journals, filtered by respective Impact Factors and included in the *Journal Citation Reports*. Data collection was performed using Web of Science, Google Scholar, Altmetric.com and Mendeley. The verification was performed in two periods in order to examine whether open access exerted influence over time. Results indicate that open access boost the number of citations and mentions in social

* Universidade Estadual de Campinas, Brasil

paulocntr@yahoo.com

** DCI, Universidade Federal de São Carlos, Brasil

chloe@ufscar.br

*** NIT-Materiais, Universidade Federal de São Carlos, Brasil

dougmilanez@gmail.com

media for the sample of papers drawn from journals under study. Moreover, this boost is greater when authors upon publishing pay the processing charge to guarantee the immediate open access availability of the paper. In conclusion, while open access can boost the number of citations of a paper, the quality of the paper itself drives most citations.

Keywords: Open Access; Journals; Information Science; Bibliometrics; Altmetrics

Impacto del acceso abierto en citaciones y medios sociales de las principales revistas de Ciencia de la Información

Paulo Roberto Cintra, Ariadne Chloe-Furnival y Douglas Henrique-Martínez

RESUMEN

El objetivo de esta investigación fue verificar las posibles ventajas que el acceso abierto puede ofrecer al área de la Ciencia de la Información. Para ello, se analizaron los indicadores bibliométricos de citas y los datos de altmetría en 16 revistas científicas híbridas, seleccionadas mediante el *Journal Citation Reports* y filtradas con base en sus respectivos factores de impacto. La recolección de datos fue realizada en Web of Science, Google Scholar, Altmetric.com y Mendeley. Esta verificación se realizó en dos periodos de tiempo diferentes para examinar si hubo alguna influencia del acceso abierto en el tiempo. Los resultados indican que el acceso abierto puede ofrecer una ventaja en el número de citas y menciones en las redes sociales para el conjunto de artículos de las revistas analizadas aquí, y que esta ventaja es mayor para los casos en que los autores pagaron el cargo por procesamiento del artículo a garantizar la disponibilidad inmediata del artículo en acceso abierto al momento de la publicación. En la conclusión se afirma que no es sólo el acceso abierto el que provoca una mayor cantidad de citas a un artículo, aunque ayuda, sino la calidad del propio artículo.

Palabras clave: Acceso Abierto; Revistas Científicas; Ciencia de la Información; Bibliometría; Altmetría

INTRODUCTION

Making information available and accessible to anyone is an intrinsic part of the work of Library and Information Science professionals, a principle that, according to Way (2010), was already identified by Ranganathan in *The Five Laws of Library Science*. What is observed, however, is that most scientists and researchers tend not to make their work available in open access (OA). In fact, as reported by Gargouri *et al.* (2010), only 15 to 20% of articles published worldwide are self-archived. According to the researchers, even if institutional strategies are adopted to encourage researchers to self-archive, many authors would only self-archive when required by institutions where they are they employed or by funding agencies.

Given the fact that the impact of scientific research is usually gauged by the number of citations garnered by a scientific paper, a much-debated question is whether works available through OA are more frequently cited than those available only through non-OA¹. This hypothesis, known as the “Open Access Citation Advantage” (OACA), argues that the ease of access can increase visibility and, consequently, potentiate citation frequency. The pioneering work in this area by Lawrence (2001) is often cited in this respect. He analyzed citation patterns of conference papers from the field of Computer Science and related areas, verifying that the most recent articles were more likely to be available online. This ease of access contributed, in turn, to the increase in the number of citations these papers receive.

Since the publication of Lawrence’s (2001) results, as Swan (2010) has shown, numerous other studies have been performed with the aim of examining the existence of the OACA in several other knowledge fields (Antelman, 2004; Harnad and Brody, 2004; Atchinson and Bull, 2015). Archambault *et al.* (2013), for example, verified OACA in 22 knowledge fields. Swan (2010), however, emphasized that there is also research showing there is no advantage in the number of citations for certain scientific fields, such as Economics (Frandsen, 2009 in Swan, 2010).

For the specific case of the Library and Information Science, Xia, Myers and Wilhoite (2011) conducted research to verify the existence of the OACA in 20 journals selected from *Ulrich’s Periodicals Directory* and the *Journal Citation Report* (JCR). As a result, the authors found there is a positive and statistically significant correlation between the number of citations and OA availability. Furthermore, the researchers verified that the number of citations

1 The expression non-open access (non-OA) is used for papers with subscription access only via the journals, without an open access version available on the web.

also increases hand in hand with number of OA copies available on the web, which is more likely to occur when the paper has multiple authors, who either self-archive the paper in their respective institutional repository, or who receive help from librarians or student assistants with the archiving process in the repository.

Nevertheless, according to Gargouri *et al.* (2010), criticisms of the OA-CA hypothesis point out that these OA citation advantages may merely reflect self-selection bias. In other words, scientists will make only those papers available in OA that they themselves consider to be of higher quality and, therefore, potentially more likely to be cited by the scientific community. Harnad (2005), for example, sought to identify the reasons why publications made available in OA would have a greater advantage in the number of citations. Besides the self-selection bias, the author also indicates: i) the advantage bias of anticipation, i.e., papers whose results are made available in OA from the pre-print stage have an advantage over those self-archived later, since they may be cited in advance of publication; ii) use advantage bias, in which OA papers tend to be downloaded more than non-OA and, therefore, tend to be more frequently cited; and iii) quality advantage bias, when papers are genuinely of high quality and, consequently, more frequently cited than the others. In this case, unlike self-selection bias, it is not only the author who judges the value of the work highly, it is also peers. Harnad (2005) and Gargouri *et al.* (2010) argue that the criticism of the self-selection bias is no longer valid, since 100% of a researcher's publications are in OA.

In contrast, Haustein *et al.* (2013) and Priem, Groth and Taraborelli (2012) believe that the assessment of a given paper's impact should move beyond analysis of formal citation. The relatively new sub-field of research metrics, Altmetrics, has emerged to accompany the growing diversity of channels mentioning scientific research outputs, such as posts on social media, or participation in academic collaboration networks (Galligan and Dyas-Correia, 2013; Kousha and Thelwall, 2007).

For Wang *et al.* (2015), Altmetrics data can also be used as a supplementary indicator in the research of OACA. These authors propose making a comparison of OA papers published in *Nature Communications* against those published in non-OA journals, in terms of the average of citations and views received, and the discussions the works prompt on social networks such as Twitter and Facebook. Their analysis confirms the citation advantage hypothesis for journal papers available in OA, verifying that this advantage extends also to Altmetrics data. That is, OA papers tend to receive more attention on social networks than non-OA papers.

Poplašen and Grgić (2016) sought to determine the existence of an “OA Altmetrics Advantage.” To do so, they used the list of top 100 papers on Altmetric.com for the year 2014, comparing non-OA with OA papers, as identified on the website. For their analysis, the authors verified the average, minimum, maximum and median values of each series for open and non-open access papers, finding that “OA and non-OA articles have similar results – average altmetric score is slightly higher for OA articles, but the median slightly higher for non-OA articles” (p. 457). For this reason, the authors, on the basis their sample, were unable to determine whether or not publishing an article in OA afforded any advantage.

In this paper, we report the results of research on the existence of an “OA Altmetrics Advantage” (OACA) in the area of Information Science, by examining whether the free and open availability of articles to the scientific community contributes to frequency of citation in other scientific papers and/or mentions on social networks.

METHOD

Bibliometric and Altmetrics indicators were developed using the procedure shown in *Figure 1*.

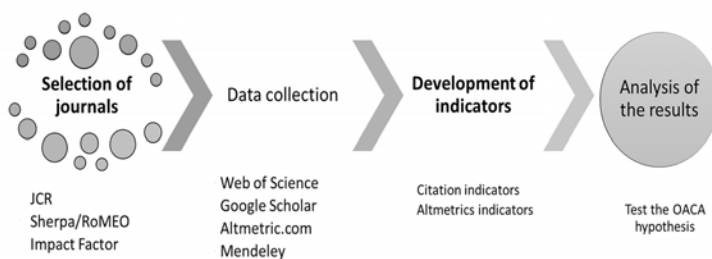


Figure1. Procedure adopted for elaboration and analysis of Bibliometric and Altmetrics indicators

The first stage involved the selection of the journals to be studied. We opted to use the journals listed under the category *Information Science & Library Science* in 2015 Social Sciences edition of the *Journal Citation Reports* (JCR), published by Thomson Reuters. This choice was motivated by the fact that JCR is linked to the Web of Science bibliographic database, which is acknowledged for indexing scientific journals of scholarly relevance. Moreover, it is traditionally used in bibliometric studies. From this list, and using the

Sherpa-RoMEO database (<http://www.sherpa.ac.uk/romeo/index.php>), we noted journal policies with regard to permission to authors to provide OA to their work in an institutional repository or on a personal webpage (the so-called “green road” to OA). We were thereby able to identify those journals distributed exclusively through subscription access, those which allow OA and those using a hybrid publishing model.

Given that we were unable to find a free means of automating the Altmetrics data collection, we set up filters for selecting a manageable number of journals for data collection. The first selection made was from those journals classified as hybrid, that is, subscription access journals that also offer authors the choice for immediate publication in OA if they pay the article processing charges (APCs). We also decided to limit our collection to articles published in 2013, the time needed for article citations to be indexed and subsequently entered in the indexing database.

It should be noted that depending on the policies of each journal, publishers can authorize the deposit of different versions of the paper submitted. In addition to this, distinct embargo periods can be stipulated, after which authors can self-archive their papers in repositories. Therefore, we accessed the website of each journal to check the embargo period between the publication date and online deposit in an OA institutional repository. In this way, we chose journals whose embargo periods varied from 0 and 12 months.

Finally, we applied a journal selection filter, choosing those only those journals with an impact factor (IF) of 1.000 or above. *Table 1* details the characteristics of the set of journals analyzed in this study, showing the IF, required embargo period before self-archiving, the total number of articles and the percentage of papers available in OA to April 21, 2017.

| Journal | Impact Factor | Embargo period (months) | Total no. of published articles | Nº of articles OA journals (APC) | Nº of green road OA articles | % OA Total (APC+green road) |
|------------------------------------------------|---------------|-------------------------|---------------------------------|----------------------------------|------------------------------|-----------------------------|
| <i>Aslib Proceedings*</i> | 1.147 | 0 | 33 | 0 | 12 | 36.4 % |
| <i>European Journal of Information Systems</i> | 2.892 | 12 | 35 | 0 | 10 | 28.6 % |
| <i>Information Society</i> | 1.333 | 0 | 22 | 0 | 6 | 27.3 % |
| <i>Information Systems Journal</i> | 2.522 | 12 | 26 | 0 | 10 | 38.5 % |
| <i>Information Systems Research</i> | 3.047 | 12 | 58 | 0 | 30 | 51.7 % |
| <i>Information, Technology & People</i> | 1.150 | 0 | 18 | 1 | 10 | 61.1 % |

| | | | | | | |
|-------------------------------------------------------------|-------|----|------------|-----------|------------|---------------|
| <i>Int. Journal of Computer-Supported Collab. Learning</i> | 2.200 | 12 | 19 | 0 | 7 | 36.8 % |
| <i>International Journal of Geographical Inf. Science</i> | 2.065 | 12 | 122 | 1 | 39 | 32.8 % |
| <i>Journal of Academic Librarianship</i> | 1.150 | 12 | 71 | 1 | 37 | 53.5 % |
| <i>Journal of Documentation</i> | 1.063 | 0 | 40 | 0 | 20 | 50.0 % |
| <i>Journal of Librarianship and Information Science</i> | 1.239 | 12 | 23 | 0 | 16 | 69.6 % |
| <i>Library & Information Science Research</i> | 1.230 | 12 | 35 | 0 | 15 | 42.9 % |
| <i>Online Information Review</i> | 1.152 | 0 | 48 | 0 | 18 | 37.5 % |
| <i>Program - Electronic Library and Information Systems</i> | 1.000 | 0 | 22 | 1 | 13 | 63.6 % |
| <i>Scientometrics</i> | 2.084 | 12 | 249 | 9 | 126 | 54.2 % |
| <i>Social Science Computer Review</i> | 1.525 | 12 | 51 | 3 | 31 | 66.7 % |
| Total | - | - | 872 | 16 | 400 | 47.7 % |
| Source: Data from the research | | | | | | |
| *Currently: <i>Aslib Journal of Information Management</i> | | | | | | |

Table 1. Details of selected journals

Bibliographic data collection of published articles for these journals was performed in the Web of Science's Core Collection on July 27, 2016 and on April 21, 2017. The number of citations each article garnered was checked in the *Citation Report*, accessed through the same database. We also collected citation data from three other sources: Google Scholar, Altmetric.com and Mendeley, in order to assess if any of these sources best captures the impact that these open access articles have on social networks. To that end, we attached the Altmetric Bookmarklet, a free app provided by the Altmetric.com website, to each article. For data from Mendeley, we noted the number of readers who had downloaded the reference data of the articles under analysis. We first carried out the Altmetrics data collection between July 27 and 30, 2016. The second data collection was carried out between April 21 and 22, 2017.

We accessed each journal site individually to verify which articles were available in OA. In line with Antelman (2004), in order to ascertain whether articles are accessible via the green road, we performed a search in Google Scholar for articles deposited in OA repositories, on personal webpages of authors and in other sources in order to determine the extent of full-text and barrier-free access to these papers. Interestingly, at the time of the second data collection, there were 58 green road OA articles not originally available in OA at the time of the first data collection. Evidently, these articles were archived in a repository at some point between the two data collection moments.

For articles located in OA, we logged data on the version available, i.e., the pre-print or the post-print. We should emphasize that if copies of the article existed on more than one website, we counted only the first result brought up by the search engine. Thus, there might have been cases in which the article was available simultaneously in the institutional repository and on the author's personal webpage. We did not check each case, because at this point in the research, we were interested in whether the article was openly and freely available. In this sense, we were taking the cue from Willinsky (2006) who observed that an alternative to the creation and maintenance of digital repositories by institutions is, indeed, the posting of publications on researchers' personal webpages, on webpages within the university or website operated by the research group. For this reason, our study employs the term *Institutional website* to webpages hosting the freely downloadable article, whether an institutional repository or webpage on a university website.

Lastly, in order to verify the OACA hypothesis and in view of observations by Craig *et al.* (2007) regarding potential methodological issues, we followed the method used by Wang *et al.* (2015) and Poplašen and Grgić (2016) to calculate the average number of citations and mentions on the social web garnered by OA articles and compare these against those obtained by the non-OA articles. The aim of this stage of the data collection is to determine whether the OACA hypothesis can be confirmed for the Altmetrics data collected from sources with a much wider reach than the Web of Science. Moreover, this analysis was performed individually for each data collection set, and the two analyses were compared against each other in order to verify whether OA contributes to a faster increase in the average citation value compared to non-OA articles. We also separately assessed open access offered through the payment of APC and through the self-archiving of articles in OA by the authors.

RESULTS AND DISCUSSION

As seen in *Figure 2*, the post-print versions of the articles (61.3%) were those versions most frequently found on Google Scholar. This may be the consequence of the journal's open access archiving policies, allowing either self-archiving of the pre-print or post-print version of the published paper. A trend was also observed regarding the use of institutional websites and academic social networks for self-archiving, which confirms the relevance of institutional repositories for the dissemination of knowledge generated in a given institution and leveraging scientific progress (Harnad, 2007; Swan, 2010; Suber, 2012). *Figure 2* also shows that only 3.8% of the total number of OA articles analyzed are immediately available in open access in journals through the payment of the APCs.

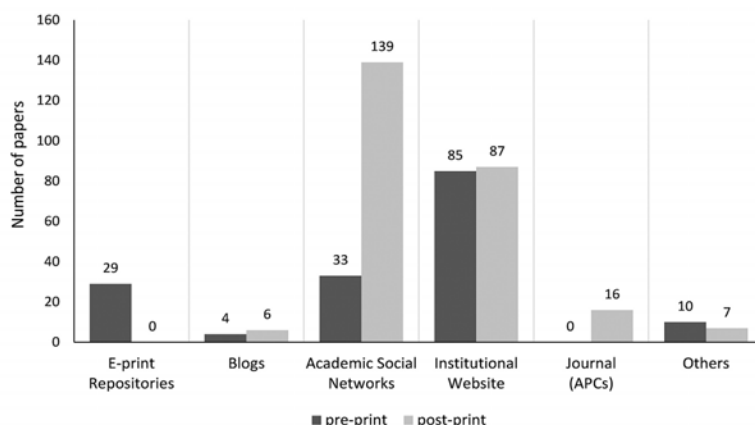


Figure 2. Locations of articles available in open access

The results show of the 416 open access (OA) articles in the second data collection, 385 had at least one citation, which represents a *citedness*² of 92.5%. For non-OA articles, the citedness rate is 90.8% (414 out of 456). This shows, at least for the number of citations, that the two values are extremely close. For comparison purposes, the rate of mentions presented by the Altmetric.com variable is only 48.3% for OA articles and 35.1% for non-OA articles. Although this indicator shows OA seems to achieve at least one social web reference for a greater number of articles, when we compare this data against

2 According to Wang *et al.* (2015), citedness can be calculated by dividing the number of articles with at least one citation by the total number of articles analyzed.

the citation rate, we find there are only a few articles that exhibit this behavior. *Tables 2 and 3* show the results identified from the calculations carried out for the verification of the OACA hypothesis for the first and second data collections.

| Information source | Function | Non-open access | Open access | |
|-----------------------------------------------|----------|-----------------|-------------|------------|
| | | | APC | Green road |
| Web of Science (average no. of citations) | average | 3.95 | 6.94 | 4.99 |
| | min | 0 | 0 | 0 |
| | max | 42 | 25 | 39 |
| | median | 3 | 5 | 3 |
| Google Scholar (average no. of citations) | average | 12.58 | 20.38 | 16.79 |
| | min | 0 | 1 | 0 |
| | max | 222 | 96 | 232 |
| | median | 8 | 15 | 11 |
| Mendeley (average no. of readers) | average | 29.70 | 46.69 | 35.18 |
| | min | 1 | 12 | 2 |
| | max | 184 | 104 | 214 |
| | median | 24 | 44 | 27 |
| Altmetric.com (average no. of donut score) | average | 1.13 | 4.31 | 2.47 |
| | min | 0 | 0 | 0 |
| | max | 47 | 25 | 61 |
| | median | 0 | 1 | 0 |

Table 2. Statistical averages for the first dataset

| Information source | Function | Non-open access | Open access | |
|----------------------------------------------|----------|-----------------|-------------|------------|
| | | | APC | Green road |
| Web of Science (average no. of citations) | average | 6.07 | 8.69 | 7.39 |
| | min | 0 | 0 | 0 |
| | max | 82 | 40 | 74 |
| | median | 4 | 6.5 | 5 |

| | | | | |
|-----------------------------------------------|---------|-------|-------|-------|
| Google Scholar (average no. of citations) | average | 17.16 | 22.00 | 22.40 |
| | min | 0 | 5 | 0 |
| | max | 347 | 122 | 332 |
| | median | 12 | 15.5 | 15 |
| Mendeley (average no. of readers) | average | 35.44 | 53.38 | 43.16 |
| | min | 1 | 14 | 2 |
| | max | 229 | 127 | 244 |
| | median | 29 | 46.5 | 33 |
| Altmetric.com (average no. of donut score) | average | 1.24 | 4.75 | 2.43 |
| | min | 0 | 0 | 0 |
| | max | 47 | 25 | 57 |
| | median | 0 | 2 | 0 |

Table 3. Statistical averages for the second collected dataset

According to *Tables 2 and 3*, was can infer that OA offers some advantage in the number of citations for journals in the field of Information Science, since the averages of OA articles located in both Web of Science and Google Scholar are higher than those for non-OA articles. This information is valid for the two collection periods. It can also be observed in the similar maximum and median values exhibited for green road OA articles and non-OA papers.

Additionally, the Altmetrics data also corroborate the OACA hypothesis. For Altmetric.com and the Mendeley data, we observe that OA articles via APC are those that show the greatest proportional advantage. This suggest the impact of an article on social webs can be improved by making it immediately available. i.e., providing OA, on the website of the journal where it was published. *Figures 3 and 4* illustrate the mean values obtained for each of the information sources analyzed.

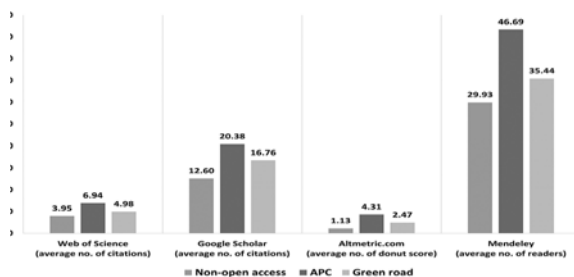


Figure 3. Analysis of OACA in diverse information sources for the first data collection

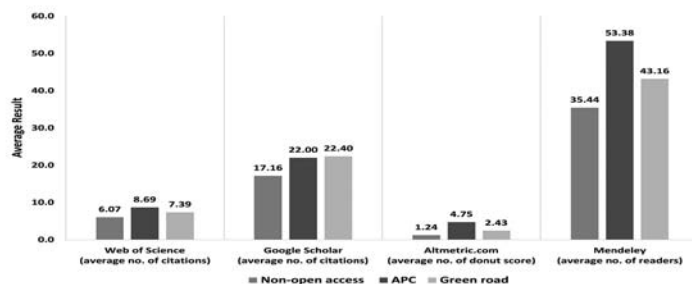


Figure 4. Analysis of OACA in different information sources for the second data collection

Figures 3 and 4 show that for the dataset analyzed for the field of Information Science, open access via the payment of article processing charges (APCs), in most cases, leads to more citations, a greater number of mentions on the social web and a greater number of readers on Mendeley when compared to papers self-archived by authors.

Table 4, on the other hand, shows the percentage growth rate of the means of the variables analyzed in this study. As can be observed, non-open access articles exhibited the largest percentage of growth in citation averages on Web of Science and Google Scholar. It should also be noted that the average Altmetric Score for green road open access articles declines from one data set to another. This could be related to the fact that in the second collection some of the articles did not obtain as many scores as in the first collection. On this point, Haustein, Bowman and Costas (2015) explain that there is uncertainty regarding the consistency of the Altmetrics data, because they are linked to platforms that are constantly changing and whose users are so diverse they have not yet been completely understood.

| Information source | Non-open access | Open access | |
|--------------------|-----------------|-------------|------------|
| | | APC | Green road |
| Web of Science | 53.8% | 25.2% | 48.1% |
| Google Scholar | 36.4% | 8.0% | 33.4% |
| Mendeley | 19.3% | 14.3% | 22.7% |
| Altmetric.com | 9.5% | 10.1% | -1.9% |

Table 4. Average percentage variation between the two data collection

Our results showed that regardless of whether an article is OA or non-OA, the largest increases in absolute numbers of the total citations obtained from the Web of Science were attained by articles that had reached a large number of citations in the first data collection. This also occurred with Google Scholar citation values and the number of readers in Mendeley. This phenomenon, in which success seems to lead to more success, was also observed by Solla Price (1976), who called it the “Cumulative Advantage Distribution.” Analogous to the “Matthew Effect” described by Merton (1968), Solla Price (1976) argues that articles with garnering a large number of citations are more likely to be cited versus those with few citations. Similarly, a journal that is constantly used as a reference for a particular subject area tends to be consulted more often than those less frequently used of those that are just emerging.

In the case of Altmetric.com, as only 76 articles (8.7%) exhibited changes in scores. Moreover, it was not possible to identify a pattern to justify greater or smaller increases or decreases in the donut score. The low number of articles exhibiting changes in Altmetric Score shows that this indicator, obtained from the mentions on social networks, seems to be more stable than the others, insofar as it captures only the most immediate impact of a publication. Therefore, after a period immediately following its publication, an article is unlikely to be mentioned again on a social network.

FINAL CONSIDERATIONS

Through the elaboration and analysis of bibliometric and altmetrics indicators, the aim of this research is to examine the impact of open access (OA) of journal articles in the area of Information Science. The study allows us to infer that archiving an article in open access seems to contribute to an advantage in the potential number of citations and mentions in social media. The study, therefore, corroborates results found regarding the Open Access Citation Advantage (OACA) for other areas of knowledge, while showing that these advantages also exist in terms of mentions in social networks. As such, the viability of an “OA Altmetrics Advantage” is supported.

In this way, we see that OA can contribute to a greater diffusion of knowledge within the scientific community and potentiate the number of citations scientists obtain, something that can positively influence the impact of their research groups or institutions. “OA is not just about public access rights or the general dissemination of knowledge: It is about increasing the impact and thereby the progress of research itself” (Gargouri *et al.*, 2010: 1).

Although articles made available in OA immediately by journal – via the payment of the article processing charge (APC) – exhibited the greatest percentage advantage regarding their number of citations and social network mentions, we cannot indisputably assert that this model is the best in terms of returns to the authors in the form of citations; since only 16 of the 872 articles were made available in OA in this way. Additionally, the literature in this area suggests OACA exists largely because scientists tend to make their better-quality articles available in open access. As such, these articles already tend to receive more citations. This phenomenon was also demonstrated in our research, since the articles in the second sample with the highest percentage citation increases were also those that had already reached a large number of citations in the first data collection. Therefore, it is important to acknowledge that while OA drives a higher citation numbers, the quality of the article is also a significant factor.

In addition to testing the OACA hypothesis, we verified its existence in two time periods in order to examine OA influence of over time. Despite the lower average values, we found the number of citations increased more rapidly for non-OA articles. Furthermore, the Altmetrics data remained practically stable between the two data collections, regardless of which set was analyzed. This demonstrates that we must examine what altmetric indicators actually measure. We hope we have contributed to this discussion by showing that altmetric indicators on the whole actually tend to reflect the more immediate and short-term repercussion of publications: and that it is unlikely that an article will be mentioned again in social networks after the period immediately following publication. Despite this, it should be noted that this study set about the task of obtaining the Altmetrics data. Since an automated, free method of obtaining this information was not available, it was difficult to perform the study on a large sample, something that limits the research to providing only a partial view of Altmetrics in the field of Information Science. Further research could verify the existence of a correlation of the data with shorter time intervals, whereby its could be ascertain whether articles with a substantial number of social media mentions today might lead to a large number of citations in the future.

Finally, we emphasize that more than half of the articles in our sample were not OA available either by APC or by the green road. This aligns with Gargouri *et al.* (2010), who observed that not all authors self-archive their publications after expiration of the embargo period, which means subscription to the journal is still required to access a large portion of the literature. From this standpoint, our results could provide support to decision-makers at government and funding agencies to act more effectively in formulating

future science and technology policies that encourage making scientific information available to any interested party; and, in this way, contribute to the achievement of the goals pursued by OA. Such policies might include, for example, the establishment broad institutional repositories providing access to more than theses and dissertations, or automatic OA publication of the articles after expiration of embargo periods. Such measures would do much to contribute to the consolidation of OA in the field of Information Science.

Acknowledgements

We are grateful to the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for the Master's grant awarded to the first author (process n°. 2015/16872-2) and the Post-Doctorate grant awarded to the third author (process n°. 2015/18878-8).

REFERENCES

- Antelman, Kristin. 2004. "Do open-access articles have a greater research impact?" *College & Research Libraries* 65 (5): 372-382.
- Archambault, Eric, Didier Amyot, Phillippe Deschamps, Aurore Nicol, Lise Rebout, and Guillaume Roberge. 2013. *Proportion of open access peer-reviewed papers at the European and world levels - 2004-2011*. Quebec (QC): Science-Metrix; Report No.: RTD-B6-PP-2011-2. Accessed 19 May 2016. http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf
- Atchinson, Amy, and Jonathan Bull. 2015. "Will open access get me cited? An analysis of the efficacy of open access publishing in political science." *PS: Political Science & Politics* 48 (1): 129-137.
- Craig, Iain, Andrew M. Plume, Marie. E. McVeigh, James Pringle, and Mayur Amin. 2007. "Do open access articles have greater citation impact? A critical review of the literature." *Journal of Informetrics* 1 (3): 239-248.
- Galligan, Finbar, and Sharon Dyas-Correia. 2013. "Altmetrics: rethinking the way we measure." *Serials Review* 39 (1): 56-61.
- Gargouri, Yassine, Chawki Hajjem, Vincent Larivière, Yves Gingras, Les Carr, Tim Brody, and Stevan Harnad. 2010. "Self-selected or mandated, open access increases citation impact for higher quality research." *PLoS ONE* 5 (10).
- Harnad, Stevan. 2005. *OA impact advantage = EA + (AA) + (QB) + QA + (CA) + UA*. School of Electronics & Computer Science, University of Southampton. Accessed 27 June 2018. <http://eprints.soton.ac.uk/262085/>
- Harnad, Stevan. 2007. "The green road to open access: a leveraged transition." In *The culture of periodicals from the perspective of the electronic age*, A. Gacs, ed., 99-105. Paris: L'Harmattan.

- Harnad, Stevan, and Tim Brody. 2004. "Comparing the impact of open access (OA) vs. non-OA articles in the same journals." *D-Lib Magazine* 10 (6).
- Haustein, Stephanie, Timothy Bowman, and Rodrigo Costas. 2015. "Interpreting "altmetrics": viewing acts on social media through the lens of citation and social theories." arXiv.org [preprint]. Accessed 16 April 2017. <https://arxiv.org/abs/1502.05701>
- Haustein, Stephanie, Isabella Peters, Judit Bar-Ilan, Jason Priem, Hadas Shema, and Jens Terliesner. 2013. "Coverage and adoption of altmetrics sources in the bibliometric community." arXiv.org [preprint]. Accessed 02 May 2016. <http://arxiv.org/ftp/arxiv/papers/1304/1304.7300.pdf>
- Kousha, Kayvan, and Mike Thelwall. 2007. "The web impact of open access social science research." *Library & Information Science Research* 29 (4): 495-507.
- Lawrence, Steve. 2001. "Free online availability substantially increases a paper's impact [Internet]." *Nature* 411 (6837): 521-521.
- Merton, Robert K. 1968. "The Matthew effect in science: the reward and communication systems of science are considered." *Science* 159 (3810): 56-63.
- Poplašen, L. M., and I. H. Grgič. 2016. "Altmetric and bibliometric scores: does Open Access matter?" *Qualitative and Quantitative Methods in Libraries* 5: 451-460.
- Priem, Jason, Paul Groth, and Dario Taraborelli. 2012. "The altmetrics collection." *PLOS ONE* 7 (11).
- Solla Price, Derek. 1976. "A general theory of bibliometric and other cumulative advantage processes." *Journal of the American Society for Information Science* 27 (5): 292-306.
- Suber, Peter. 2012. *Open access*. Cambridge (MA): MIT Press.
- Swan, Alma. 2010. *The Open Access citation advantage: studies and results to date [report]*. Southampton (GB): University of Southampton Institutional Repository. Accessed 02 May 2015. <http://eprints.ecs.soton.ac.uk/18516>
- Wang, X., C. Liu, W. Mao, and Z. Fang. 2015. "The open access advantage considering citation, article usage and social media attention." *Scientometrics* 103 (2): 555-564.
- Way, Doug. 2010. "The open access availability of library and information science literature." *College & Research Libraries* 71: 302-309.
- Willinsky, John. 2006. *The access principle: the case for open access to research and scholarship*. Cambridge (MA): MIT Press.
- Xia, Jingfeng, Rebekah Myers, and Sara Kay Wilhoite. 2011. "Multiple open access availability and citation impact." *Journal of Information Science* 37 (1): 19-28.

Para citar este texto:

Cintra, Paulo Roberto, Ariadne Chloe Furnival y Douglas Henrique Milanez. 2018. "The impact of open access citation and social media on leading top Information Science journals". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 117-132. <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57874>

Tendencias de la innovación tecnológica en Colombia 1991-2013 a partir del análisis de patentes

Rosa Elena Gómez Hurtado*

Artículo recibido:

19 de octubre de 2017

Artículo aceptado:

20 de septiembre de 2018

Artículo de investigación

RESUMEN

Se presenta un análisis de las patentes concedidas a instituciones o inventores de origen colombiano registradas en las bases de datos de las oficinas de patentes de Estados Unidos, Europa, WIPO, Alemania y Colombia. Se observan las tendencias de innovación tecnológica que rigen los procesos y productos en Colombia durante el periodo y se comprueba que efectivamente las patentes son indicadores de innovación tecnológica. Se evidencia que las invenciones de colombianos no son necesariamente patrocinadas por instituciones colombianas y que la figura del inventor solitario tiende a desaparecer para dar paso a las innovaciones industriales en manos de los equipos

* Facultad de Ingeniería y Nuevas Tecnologías
Fundación Universitaria UNINPAHU, Colombia rhgomez02@uninpahu.edu.co

de investigación de las grandes empresas, convirtiéndose en un mecanismo de control de la producción y comercialización mundiales.

Palabras clave: Innovación Tecnológica; Patentes; Patentometría

Patent analysis of Colombian technological innovation 1991 - 2013

Rosa Elena Gómez-Hurtado

ABSTRACT

This study offers an analysis of patents granted to Colombian institutions or inventors and registered in patent databases in the US, Europe, WIPO, Germany and Colombia. The trends regarding processes and products of technological innovation in Colombia during the period are observed showing that patents are in fact indicators of technological innovation. Moreover, Colombian inventions are not necessarily sponsored by Colombian institutions, and the phenomenon of the solitary inventor is giving way to research teams working for large companies, which exercise control over production and marketing.

Keywords: Technological Innovation; Patents; Patentometrics

INTRODUCCIÓN

El análisis de patentes es una de las herramientas para entender y direccionar el desarrollo tecnológico y la innovación, pero por sí solas las patentes no desencadenan el proceso de desarrollo tecnológico, pues forman parte de un sistema dinámico en el cual los esfuerzos que se le dediquen a la I+D serán la materia prima del sistema científico tecnológico (Montes Hincapié *et al.*, 2009). Esta es una de las razones por las que se monitorean las actividades de ciencia, tecnología e innovación (ACTI) al igual que por la incidencia que el patentamiento tiene en las estrategias de I+D en el desarrollo de los países.

El periodo analizado (1991-2013) tiene características determinantes que es necesario tener presentes por su incidencia en la decadencia de áreas que para los inventores eran muy importantes, pero que con el paso del tiempo van perdiendo interés y son reemplazadas por otras que se convierten en estratégicas; es allí donde se concentra el nivel de I+D+I. Para el patentamiento, la globalización de la economía, caracterizada por la complejidad y un alto nivel de incertidumbre, obligó a las empresas a ser más competitivas, a centrarse en su saber hacer y a establecer las alianzas estratégicas necesarias para tener presencia en los mercados globales, con el propósito de reducir la incertidumbre.

Otro hito importante a considerar radica en los cambios tecnológicos que desde décadas pasadas se venían preparando para alcanzar el mayor grado de madurez, a partir de las siguientes características:

- el proceso de fabricación flexible presiona sobre la innovación tecnológica,
- se da una visión orgánica de la empresa, y
- comienzan a desarrollarse estrategias tecnológicas.

La innovación tecnológica es entendida como “el producto, proceso o metodología que aparece en un mercado determinado y que es aceptado por el mismo” (Barceló, Solé Parellada y Valls, 1992: 127). Sin embargo, es conveniente aclarar que innovación no es sinónimo de invención, sino un concepto o proceso más amplio (Ortiz y Pedroza Zapata, 2006: 66). En este contexto, las actividades de innovación tecnológica son el conjunto de etapas científicas, tecnológicas, organizativas, financieras y comerciales, incluyendo las inversiones en nuevos conocimientos, que llevan o intentan llevar a la implementación de productos y de procesos nuevos o mejorados (OCDE, 2003: 17). Estas actividades son la adquisición de la tecnología y de *know-how*, la puesta a punto de herramientas y la ingeniería industrial, entre otras. Numerosos estudios han demostrado que las actividades de I+D se realizan en cooperación con investigadores individuales, equipos de investigadores y unidades de investigación.

En la actividad inventiva el indicador más elemental y generalmente más utilizado es el simple cómputo de las patentes, aunque también se pueden contabilizar aquellas registradas por los inventores de un país determinado, bien sea en forma global o por sectores de actividad. Este indicador debe ser manipulado con precaución, puesto que las estrategias de adopción de patentes dependen de los sectores, de las compañías y de las legislaciones nacionales (Callon, Courtial y Penan, 1995).

En Colombia, el modelo lineal de la ciencia (insumo-producto) que sigue las orientaciones del Manual de Frascati evalúa el gasto que tienen las empresas en ACTI a partir de los resultados de las encuestas de desarrollo e innovación tecnológica (EDIT), distinguiendo aquellas actividades dedicadas a la innovación (Lucio *et al.*, 2013), teniendo en cuenta que están atadas al fortalecimiento del Sistema Nacional de Ciencia, Tecnología e Innovación (SNCTI) (CONPES, 2008) a través de acciones como la mejora de la innovación tecnológica con énfasis en los sectores de clase mundial y la promoción de la productividad y la competitividad teniendo como plataforma la ciencia, la tecnología y la innovación.

En este sentido, el análisis de la política de innovación colombiana, que debe ir más allá del desarrollo científico y tecnológico, muestra que las empresas colombianas están siendo cada vez menos innovadoras y el gasto en actividades de I+D es demasiado bajo (70 % del PIB) (BID, 2014: 9).

Es así que la innovación tecnológica se constituye en la base de la competitividad y en una mayor integración de los productos, los procesos y el mercado, lo que determina la dinámica de patentación de los inventos de los colombianos en el periodo de estudio, principalmente a mediados de ese lapso, cuando la dinámica de patentación de innovación tecnológica cambia.

El artículo se estructura en tres partes: en la primera, se introduce el tema a través de un marco general que permite entender y ubicar el objeto de estudio; en la segunda, se presenta la metodología utilizada para el desarrollo de la investigación; en tercer lugar, se aborda el análisis de los resultados, la discusión y las conclusiones.

METODOLOGÍA

La investigación, de corte descriptivo, utilizó métodos patentométricos para explorar la innovación tecnológica en Colombia durante el periodo 1991-2013, con el propósito de mostrar el impacto de las patentes como indicadores de la innovación tecnológica. Estos métodos han sido utilizados para mostrar el rendimiento de la investigación de un país y su competitividad (Huang, Chang y Chen, 2012). Su objetivo consistió en identificar las tendencias de innovación tecnológica en Colombia mediante el análisis de la información contenida en las patentes concedidas a empresas o innovadores colombianos, registrada en los bancos de patentes, durante el periodo señalado.

Se analizaron 5 001 patentes concedidas a empresas o a inventores colombianos, independientemente de si la patente fue concedida en el país o en el exterior. La estrategia consistió en recuperar las patentes donde al menos uno

de los inventores fuera de origen colombiano, utilizando el software Aclai-
mIP y la base de datos *Thomson Innovation* que incluye los bancos de patentes
de Estados Unidos, Europa, WIPO, Alemania y Colombia, recurriendo a di-
ferentes filtros; al igual que el software Vantage Point. Se trabajó con la clasi-
ficación internacional de patentes (IPC) la cual se divide en 8 secciones, cada
una de las cuales es designada por una letra mayúscula y se constituye en el
nivel jerárquico más alto de la misma; 120 clases y 628 subclases, con el fin de
facilitar al investigador su búsqueda y la recuperación de los contenidos de las
patentes.

En la primera etapa de la metodología se seleccionaron las fuentes de los
datos originales y la descarga de la información se realizó a una hoja de Ex-
cel, observando los títulos recuperados y los resúmenes disponibles de cada
patente para su posterior tratamiento importando los datos al software Van-
tage Point, que permite realizar *Tech Mining*, una forma de analizar conte-
nidos mediante la cual se extrae información útil de varias formas de datos,
con énfasis en el análisis de los datos numéricos y explotando especialmente
las fuentes de datos de texto estructurado en campos como autor, fecha de
publicación y palabras clave, con moderada facilidad en fuentes de texto se-
miestructuradas (etiquetadas) y utilidad limitada con textos no estructura-
dos (Zhou *et al.*, 2014).

Para localizar los trabajos publicados en el periodo de análisis entre cu-
yos innovadores, empresas o personas al menos uno fuera colombiano, se
efectuó una estrategia de búsqueda global debido a la gran variedad de nom-
bres bajo los cuales aparecen las instituciones en la base de datos. Utilizando
la búsqueda avanzada, se estructuró la estrategia de búsqueda de la siguiente
manera:

CU = (Colombia OR Columbia) AND PY = 1991-2013

donde CU corresponde al país (se incluye la palabra Columbia porque otros
estudios anteriores han detectado instituciones colombianas con el país es-
crito en esa forma. En este sentido, si bien es cierto que se ha ido corrigiendo
progresivamente este error, aún se muestra al país escrito como Columbia en
muchos registros) y PY se refiere al año de publicación.

- Una vez subida la información a Vantage Point, se procedió a crear un
tesauro que facilitara el tratamiento y filtrado de los datos para superar
las posibles discrepancias de los nombres, utilizando rutinas de coin-
cidencia difusa a través de reglas de correlación, como la coocurrencia
de términos que indican una posible relación entre los mismos, para fi-

nalmente consolidar las variaciones en el nombre de las organizaciones o autores, eliminar los registros duplicados, normalización de campos, eliminación o inclusión de caracteres, desagregación de ciertos campos como las direcciones y las categorías de materia, y la unificación de variantes en nombres de instituciones, ciudades, países, entre otros.

- El análisis de los conglomerados, por otra parte, se pudo realizar mediante la agrupación de términos.
- La aplicación del mapeo se basó en una serie de pasos, entre los que se contempló la homogeneización y depuración de la información (confección de un diccionario de términos, eliminación de registros duplicados), obtención de matrices de coocurrencia (palabras, empresas, innovadores, códigos de clasificación de patentes, palabras claves, etc.) y la obtención de mapas tecnológicos. La representación de la información con la utilización del mapeo tecnológico tuvo como fin identificar áreas de conocimiento y obtener una visión de la estructura de relaciones existentes. La distancia en el mapa entre dos palabras (referentes a tecnologías, productos, autores, empresas) señala la mayor o menor relación entre ellas, según van Raan (citado por Escorsa, Maspons y Llibre, 2001).
- El análisis de los resultados se realizó adaptando la metodología propuesta por Goodman y Lawless (1994), ya aplicada por Castellanos *et al.* (2005). En este punto, el análisis de las fechas de presentación y aprobación de las patentes permitió identificar las tendencias respecto a las áreas de conocimiento, quiénes trabajan en ella y con qué (*Tech Mining*) (Zhou *et al.*, 2014)

PATENTES

Para efectos de este trabajo, entendemos como patente “una autorización legal que concede a su titular el derecho exclusivo de controlar el uso de una invención, según se la define en sus reivindicaciones, dentro de un área geográfica y en un plazo limitado impidiendo que terceros, entre otras actividades, fabriquen, usen, vendan u ofrezcan en venta la invención sin la autorización pertinente” (OMPI, 2017: 6).

Los documentos de patentes ofrecen diversa información que puede considerarse de dos tipos: 1) la documentación administrativa, que distingue elementos como a) los datos relativos al solicitante y de su representante, b) la documentación aportada, c) datos sobre la vida administrativa de la patente y d) los pagos realizados, y 2) la información técnica integrada por

tres partes: a) la descripción: plantea el problema técnico que la invención pretende solucionar, a la vez que se detalla el contenido y las características técnicas de la misma, b) las reivindicaciones: se hace referencia a la segunda parte del documento de la patente, y c) los dibujos, que no tienen por qué incluirse siempre, aunque permiten tener una concepción visual más completa de la invención.

Las patentes se constituyen en una de las fuentes de información tecnológica más completas, accesibles, confiables, actualizadas y que abarcan un gran número de sectores, y es por ello que, sobre la base de las reivindicaciones, los examinadores de las oficinas de patentes las catalogan en una categoría de clasificación. Existen varias clasificaciones de patentes y una de ellas es la clasificación internacional de patentes basada en el tratado multilateral internacional administrado por la OMPI (Arreglo de Estrasburgo, 1971), que entró en vigor en 1987 y que en la práctica es utilizada por las oficinas de propiedad industrial de más de 100 países en virtud del Tratado de Cooperación en Materia de Patentes. Sin embargo, aunque sea de aplicación en todo el mundo, algunas oficinas adoptan clasificaciones propias como es el caso de la UPSTO (Oficina de Patentes y Marcas de Estados Unidos) y de la Oficina Europea de Patentes (EPAT).

En el manual de Frascati se indica que ya las estadísticas de I+D no son suficientes para evaluar sus resultados y que en el contexto de la economía del conocimiento los datos deben ser examinados en un marco conceptual que permita relacionarlos con otros medios disponibles y con los resultados derivados de las actividades de I+D de que se trate (OCDE, 2003). Es así que el tema de los indicadores para evaluar la ciencia y tecnología surge con la aparición de la tecnometría, desarrollada por Jacob Schmookler en 1950, quien construyó e interpretó por primera vez las estadísticas de patentes para la medición de las actividades científicas y tecnológicas (Orozco y Chavarro, 2010) que se consolidaron en la década de 1960 con la National Science Foundation (NSF) y la OCDE.

Los indicadores de patentes proveen información sobre los resultados de la actividad científica en una institución, país o región del mundo. Tradicionalmente se han clasificado como sigue: a) indicadores de actividad, que proporcionan datos acerca del volumen e impacto de las actividades de investigación y cuyas variables más importantes son los *inputs* (recursos puestos en circulación) y los *outputs* o productos que se elaboran, entre los que se encuentran las patentes, y b) indicadores de relación, que rastrean los lazos y las interacciones entre investigadores y campos.

En la actividad inventiva, la patente es examinada por la oficina en don-

de se presenta la petición teniendo en cuenta la novedad y el contenido de la descripción propuesta, es decir, la reproductibilidad; con lo cual queda claro que la patente describe la invención, pero no la innovación ya que ésta supone la comercialización con éxito. Autores como Spinak (1998), Callon, Courtial y Penan (1995) y Camps (2008) señalan que los indicadores más utilizados son:

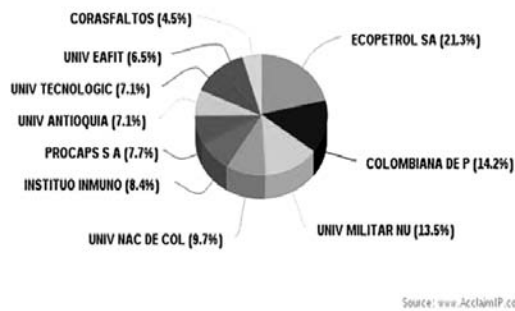
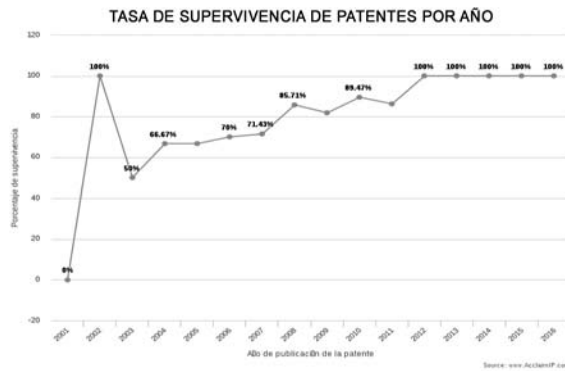
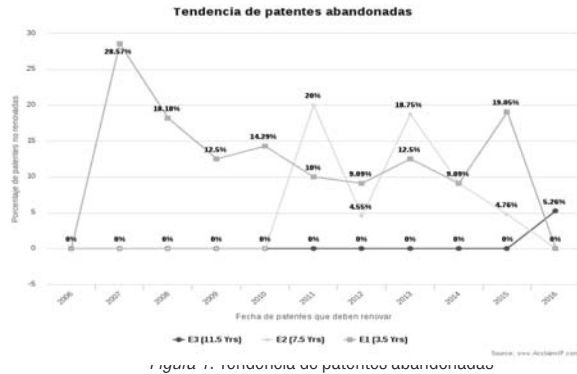
- Indicadores de actividad (número de patentes, distribución, países, etc.),
- Indicadores relacionales de primera generación (colaboraciones o vínculos entre inventores, empresas, etc.),
- Indicadores relacionales de segunda generación (coocurrencia de textos, resumen, título, etcétera),
- Indicadores de tercera generación (clasificación de la información mediante clusters, mapas tecnológicos),
- Familia de patentes (permite medir el nivel de actividad tecnológica),
- Identificación de autores, entidades, países y sectores líderes, así como su impacto,
- Identificación de competidores, posibles socios, etc., evolución de la técnica en el sector, caracterización de las empresas involucradas en el sector teniendo en cuenta su comportamiento,
- Impacto y análisis de las citaciones y cocitaciones, y
- Determinación de las principales tendencias tecnológicas e identificación de señales débiles.

PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

El abandono de una patente lleva a que ésta sea de dominio público y, por lo tanto, resulte de libre disposición para cualquier interesado, situación atractiva para aquellas empresas con menos recursos, las cuales pueden detectar tecnología de dominio público que seguramente les puede ser de utilidad.

Sobre esa base, los resultados muestran que el mayor número de patentes no renovadas se presenta en el año 2007 seguido del 2011 y 2013.

En 2003 se observa una disminución de casi el 50 % de las patentes renovadas, abandono que va disminuyendo paulatinamente a medida que avanzan los años. A diferencia de la tendencia observada en la *Figura 1* la supervivencia de patentes presenta una curva ascendente principalmente del 2004 al 2012, y en adelante su comportamiento es estable.



Se observa que las principales organizaciones inventoras colombianas son

Ecopetrol SA, Colombiana de Petróleos, Universidad Militar Nueva Granada, Universidad Nacional de Colombia, Instituto Inmunología de la Universidad del Valle, PROPCAPS S.A., Universidad de Antioquia y Universidad Tecnológica de Pereira, Universidad EAFIT y CORASFALTOS. Es probable que las instituciones con mayor número de patentes también tengan líneas de investigación muy específicas.

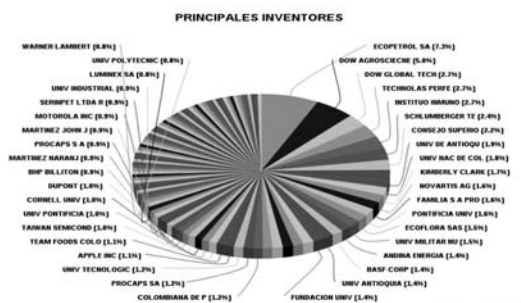


Figura 4. Organizaciones / inventores líderes

En la Figura 4 se aprecia que ECOPETROL ostenta el liderazgo de la producción total de patentes del país con 145 patentes. Tiene tres patentes conjuntas con la Universidad de Antioquia, tres con EAFIT y cinco con la Universidad Pontificia Bolivariana. Le siguen las empresas Dow Agrosciences LLC con 99 patentes y Dow Global Technology LLC con 49. Con un porcentaje menos significativo se encuentran la empresa Schlumberger Technology, la Universidad de Antioquia, Kimberly Clark y la Pontificia Universidad Javeriana. El 91.1 % restante está compuesto por instituciones o personas naturales, que si bien es cierto tienen menor número de patentes, no por ello son menos importantes.

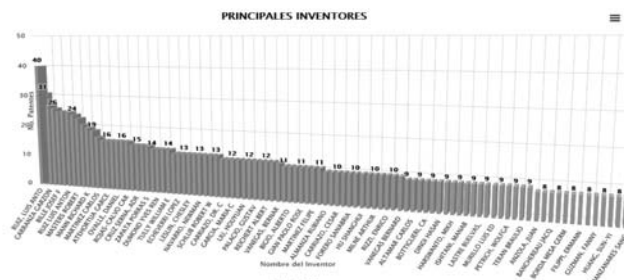


Figura 5. Principales inventores

La dinámica de la productividad de los inventores permite visualizar que Luis Antonio Ruíz, Josef Bille, Jose Orlando Giraldo, Nelson Carranza Garzón, Robert Master y Richard Mann son quienes mayor cantidad de patentes registran. El promedio de publicación de patentes por las principales clases, según la clasificación internacional de patentes (*Tabla 1*), muestra una apuesta explícita hacia el sector de las tecnologías para el manejo de residuos sólidos, preparaciones medicinales que contienen ingredientes orgánicos activos y generación de energía a través de fuentes de energía renovables, que son de las que mayor preponderancia tienen para el país en sus dos vertientes: la generación de energía y las fuentes de generación de la misma. En menor medida se encuentran las patentes sobre tratamiento de agua residual o de alcantarilla, alimentos o productos alimenticios, procesamiento de hidrocarburos, tecnologías para la producción de combustible de origen no fósil, y café: substitutos y preparaciones, estas últimas patentes basadas en un producto en el que tradicionalmente el país ha sido líder como productor.

| CLASE | DESCRIPCIÓN |
|----------|----------------------------------------------------------------------------------------------------------------|
| Y02W30 | Tecnologías para el manejo de residuos sólidos |
| A23L1 | Alimentos o productos alimenticios. Preparación o tratamiento |
| A61K31 | Preparaciones medicinales que contienen ingredientes activos orgánicos |
| Y02E10 | La generación de energía a través de fuentes de energía renovables |
| A61K9 | Preparaciones medicinales caracterizadas por la forma física especial |
| C02F1 | El tratamiento de agua, agua residual o de alcantarilla |
| Y02E50 | Tecnologías para la producción de combustible de origen no fósil |
| A61K45 | Preparaciones medicinales que contienen ingredientes activos no previstos en los grupos A61K31/00 to A61K41/00 |
| C10G2300 | Aspectos relacionados con el procesamiento de hidrocarburos cubiertos por los grupos C10G1/00 - C10G99/00 |

Tabla 1. Principales clases de patentes (grupo principal)

Es importante aclarar que a pesar de que la fecha de presentación y la fecha de prioridad se usan indistintamente, no son lo mismo. La fecha de presentación es cuando una solicitud de patente se presenta por primera vez en una oficina de patentes. La fecha de prioridad establece la novedad y/o la evidencia de una invención en particular en relación con otras y puede ser anterior a la fecha de presentación efectiva de la solicitud.

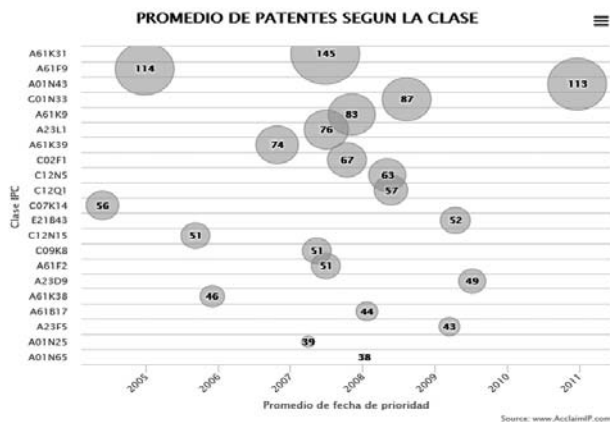


Figura 6. Tendencias tecnológicas según la clase

Ahora bien, si se observa la *Figura 6*, la tendencia dominante en la producción de patentes se presenta en el 2007 en el campo de preparaciones medicinales que contienen activos orgánicos, mientras que en el año 2004 se orienta a métodos o dispositivos para tratamiento ocular: lentes de contacto, dispositivos para corregir la visión, aparatos para guiar a los ciegos y dispositivos para proteger los ojos. Finalmente, en el 2010 las invenciones giran en torno a las biocidas, repelentes de plagas y reguladores del crecimiento de las plantas que contienen compuestos heterocíclicos.

En este periodo y los subsiguientes se pone en evidencia el intento de los inventores por ser más competitivos, por esta razón, la cooperación universidad-empresa se presenta con muchas posibilidades. Igualmente, el patentamiento y posicionamiento de los productos de patentes en el mercado tienen un costo relacionado con la adquisición de licencias, tecnologías o vigilancia tecnológica, entre otras, y es posible que estos asuntos no estén al alcance de los inventores, por lo cual la alternativa es la realización de alianzas estratégicas con otras empresas para efectuar desarrollos conjuntos como parte de la estrategia empresarial.

En consecuencia, al analizar las subclases se encuentra que a medida que avanzan los años la producción de patentes se concentra en nuevas tecnologías o áreas que bien podrían denominarse emergentes y que antes del 2009 no eran objeto de patentación por parte de los inventores. Surge a partir de ese año el interés de los inventores por inventos relacionados con productos y procedimientos químicos, como preparaciones para propósitos de tocador, aprovechando quizá el esfuerzo propio de I+D y la capacidad investigadora de centros externos constituidos por empresas o instituciones, entre las que

se encuentran las universidades. A la par, son importantes las patentes relacionadas con alimentos, productos alimenticios o bebidas no alcohólicas y biocidas, repelentes de plagas y reguladores del crecimiento de las plantas que contienen compuestos heterocíclicos.



Figura 7. Tendencias y áreas emergentes por subclases

Red de colaboración

La dinámica de generación de la innovación tecnológica se puede evaluar a partir de las redes de colaboración entre instituciones e innovadores, quienes están unidos por vínculos laborales y académicos. El componente gigante de la red de colaboración durante el periodo 1991-2013 está formado por 226 solicitantes, todos interconectados de algún modo.

El análisis se efectúa a partir de los conceptos de centralidad y densidad, que resaltan la contribución de los diferentes agregados a la estructuración de la red general. Desde esta perspectiva, se observa que entre más numerosas y fuertes sean las relaciones con otros agregados, más se detecta un conjunto de temas de innovación tecnológica considerado crucial para la comunidad de científicos o tecnólogos (Callon, Courtial y Penan, 1995: 76).

El mayor nodo que se observa en esta red de colaboración es el de Eco-petrol, que ocupa una posición estratégica manifiesta en su alto grado de desarrollo e integración. Los agregados de esta red probablemente están bajo la tutela de un grupo de investigadores bien estructurado; se detecta entre los agregados de este nodo una fuerte conexión. Los principales innovadores de esta red son Jorge Enrique Forero Sanabria, Gonzalo Leal Díaz, Guillermo Latorre Cortés y Olga Patricia Ortiz Cansino, quienes constituyen el centro de la misma. El tamaño de los enlaces indica una mayor cosolicitud

de patentes. El nodo de la Pontificia Universidad Javeriana muestra un interesante nivel de patentamiento a través de la colaboración entre y con los innovadores. Se resalta el subnodo a partir de uno de sus agregados y la patentación con la Fundación Universidad del Norte, principalmente. El patentamiento a través de la colaboración de la Universidad Nacional de Colombia se caracteriza por la debilidad de las relaciones con sus agregados.

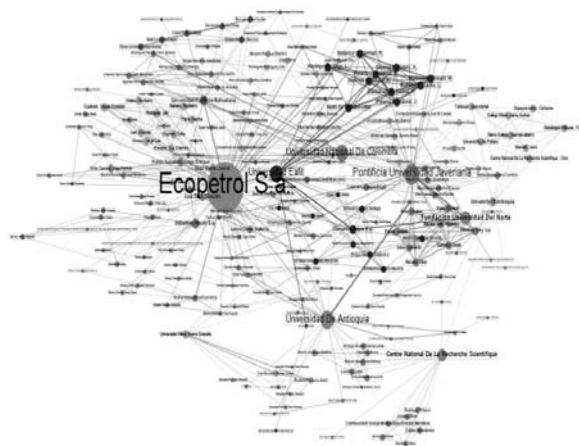


Figura 8. Redes de colaboración

En la red general se detectan agregados que se podrían denominar periféricos o podría tratarse de agregados que aún siguen siendo objeto de inversiones importantes en materia de innovación, y que se han visto marginados progresivamente llamando cada vez menos interés a su alrededor. Aparecen como redes que actúan débilmente con respecto a otras subredes y representan los márgenes de la red. Es el caso de la Universidad Militar Nueva Granada, cuya mayor densidad es con Wilmer Carrillo León y Astrid Rubiano Fonseca, y de Industrias HACEB S.A. y la subred compuesta por Luis Echeverría Afanador, Jorge Luis Granados, Humberto Videla, Claudia E. Cuadrado, Lilia Rodríguez, Martha Parra y Erika Guzmán, cuya densidad de las relaciones se muestra muy débil con respecto a otras subredes.

Se identifica, en el mismo sentido, el nodo de la Universidad de Antioquia entre cuyos agregados se encuentra el Centre National de la Recherche Scientifique; esta es una muestra del patentamiento colombiano en el exterior y de colaboración.

DISCUSIÓN

Se detecta una tendencia creciente de la actividad de patentación principalmente al final del periodo estudiado; no obstante, a pesar de que cada vez se incrementa el número de patentes en las bases de datos de patentes, los resultados generales del país no son los más convenientes frente a otros países, confirmando lo anotado por Morales y Sifontes (2013) y Sánchez, Medina y León (2007).

Los resultados de la investigación son concordantes con los estudios de la OCDE (2016) según los cuales las empresas colombianas están siendo cada vez menos innovadoras y el gasto en actividades de I+D es demasiado bajo (70 % del PIB).

Se pone de manifiesto la teoría de la productividad patentométrica sesgada (Sánchez Guzmán, 1999: 42), según la cual la productividad de los inventores, bien sea personas naturales o instituciones, no es la misma. Parece que un innovador o empresa con mayor número de patentes puede tener mayor facilidad para producir otras y seguramente se le asignará mayor cantidad de recursos para desarrollar nuevas patentes; por consiguiente, siempre va a existir un pequeño grupo de innovadores y empresas muy productivas y otras que no lo serán de la misma manera.

Se evidencia el intento de los inventores por ser más competitivos y la cooperación universidad-empresa se presenta con muchas perspectivas; sin embargo, no se puede desconocer que el patentamiento y posicionamiento de los productos de patentes en el mercado tienen un costo relacionado con la adquisición de licencias o tecnologías, entre otros aspectos, que es posible que no estén al alcance de los inventores, por lo cual la alternativa es la realización de alianzas estratégicas con otras empresas para efectuar desarrollos conjuntos, como parte de la estrategia empresarial. Igualmente, no se puede desconocer que en Colombia no se pueden patentar los seres vivos ni los segundos usos de los medicamentos ya patentados, lo cual podría influir en la cantidad de patentes registradas. Tampoco se puede patentar el software, aunque sí puede ser registrado según la legislación de derechos de autor. La base de esta reglamentación está contemplada en la Decisión 486 de la Comunidad Andina de Naciones (CAN) y en el Decreto 2591 de 2000.

CONCLUSIONES Y RECOMENDACIONES

Se advierte la importancia del salto que Colombia ha dado en la inversión en actividades de ciencia y tecnología y su incidencia en el desarrollo y la competitividad del país y, por tanto, en la disminución de la brecha tecnológica

con relación a otros países de la región, aun cuando no en la medida que se esperaría. La innovación tecnológica se constituye en la base de la competitividad del país y en una mayor integración de los productos, procesos y mercado, lo que determina la dinámica de patentación de los inventos de los colombianos en el periodo de estudio, principalmente a mediados cuando la dinámica cambia.

Se encontró que un porcentaje importante de patentes pertenecen a universidades como la Javeriana (8 %), la Universidad de Antioquia (6.5 %), la Universidad Nacional de Colombia (7.4 %), la Universidad Eafit (2.5 %) y la Universidad Militar (7.0 %). Esta realidad pone de manifiesto el avance en la relación empresa-universidad para la generación de alianzas estratégicas que benefician el progreso tecnológico del país.

El progreso tecnológico a través del patentamiento muestra una apuesta explícita hacia el sector de las tecnologías para el manejo de residuos (141 patentes), preparaciones medicinales que contienen ingredientes orgánicos activos (145 patentes) y las caracterizadas por la forma física especial y generación de energía a través de fuentes de energía renovables en sus dos vertientes: la generación de energía y las fuentes de generación de la misma (87 patentes). Sin embargo, las patentes orientadas hacia el campo de preparaciones medicinales que contienen activos orgánicos, biocidas, repelentes de plagas y atra-yentes o reguladores del crecimiento de las plantas que contienen compuestos heterocíclicos tienen gran importancia en el país (214 patentes).

Se detectaron áreas emergentes de patentación. A partir de 2009 las organizaciones o inventores se interesaron por inventos relacionados con preparaciones para tocador y patentamiento aislado en ingeniería genética, potenciando sus propias capacidades en I+D y las de otras instituciones como las universidades.

La red de colaboración muestra que son numerosos los esfuerzos de patentamiento individual, más que de colaboración, que se realizan en el periodo analizado. Es importante resaltar que las invenciones de los colombianos no son necesariamente patrocinadas por empresas o instituciones de origen colombiano. En un buen número de los casos se trata de instituciones de carácter internacional como el Centre National de la Recherche Scientifique, perteneciente al nodo de la Universidad de Antioquia.

La autora quiere agradecer a Sergio Cuellar Parrado y Matt Troyer por su colaboración en la investigación.

Artículo basado en los resultados del proyecto “Tendencias de innovación tecnológica que rigen los procesos y productos en Colombia durante el periodo 1991 – 2013 mediante el análisis de las patentes solicitadas y concedidas a colombianos” INP-2016-008218.

REFERENCIAS

- Barceló, Miquel, Francesc Solé Parellada y Jaume Valls. 1992. “Tecnología y PME en España”. *Economía Industrial* 284: 127-136.
- BID (Banco Interamericano de Desarrollo). 2014. “¿Cómo Repensar El Desarrollo Productivo?”, en *Políticas e Instituciones Sólidas Para La Transformación Económica*, Gustavo Crespi, Eduardo Fernández-Arias y Ernesto Stein (ed.), Washington: BID. <https://publications.iadb.org/handle/11319/6634>
- Callon, M., J. P. Courtial y H. Penan. 1995. *Cienciometría. El estudio cuantitativo de la actividad científica: de la bibliometría a la vigilancia tecnológica*. Gijón, España: Ediciones Trea.
- Camps, Diego. 2008. “Limitaciones de los indicadores bibliométricos en la evaluación de la actividad científica biomédica”. *Colombia Médica* 39 (1): 74-79.
- Castellanos, Oscar Fernando, Andrés Mauricio León, Víctor Mauricio Montañez, Claudia Nelcy Jiménez y Freddy Abel Vargas. 2005. “El análisis de patentes como instrumento para la generación de conocimiento y de estrategia tecnológica”. *Revista GTI* 4 (9): 71-82.
- CONPES (Consejo Nacional de Política Económica y Social). 2008. *Política Nacional De Competitividad y Productividad*.
- Escorsa, Pere, Ramón Maspons y Joan Llibre. 2001. *De la vigilancia tecnológica a la inteligencia competitiva*. Madrid: Prentice Hall.
- Goodman, Richard Arthur y Michael W. Lawless. 1994. *Technology and Strategy: Conceptual Models and Diagnostics*. Oxford University Press on Demand.
- Huang, Cui, Jun Su, Xiang Xie, Xuanning Ye, Zhang Li, Alan Porter y Jiang Li. 2015. “A Bibliometric Study of China’s Science and Technology Policies: 1949–2010”. *Scientometrics* 102 (2): 1521-1539.
- Huang, Mu-Hsuan, Han-Wen Chang y Dar-Zen Chen. 2012. “The Trend of Concentration in Scientific Research and Technological Innovation: A Reduction of the Predominant Role of the US in World Research & Technology”. *Journal of Informetrics* 6 (4): 457-468.
- Lucio, Jorge, Diana Lucio-Arias, Mónica Salazar, Cristhian Ruiz, Edgar Bueno, Jose Montes, Juliana Velandia et al. 2013. *Observando el sistema colombiano de ciencia, tecnología e innovación: sus actores y sus productos*. Colombia: Observatorio Colombiano de Ciencia y Tecnología.

- Montes Hincapié, Juan Manuel, Joaquín Lloveras Macià, Pere Escorsa i Castells y Elicet Cruz Jiménez. 2009. "Análisis de patentes en el sector de los aceites esenciales como estrategia para impulsar el desarrollo regional en Antioquia". XIII Seminario Latino-iberoamericano de Gestión Tecnológica ALTEC, Cartagena de Indias
- Morales, R., y D. Sifontes. 2013. "Reporte de la Actividad Innovadora de América Latina: Un Estudio de Patentes". VIII Congreso de Indicadores de Ciencia y Tecnología. OCDE (Organización para la Cooperación y el Desarrollo Económicos). 2016. *Estudios de la OCDE de las Políticas de Innovación: Colombia*. OECD Publishing. <https://www.oecd.org/sti/inno/colombia-innovation-review-assessment-and-recommendations-spanish.pdf>.
- OCDE. 2003. *Manual de Frascati. Propuesta de norma práctica para encuestas de investigación y desarrollo experimental*. París: Fundación Española para la Ciencia y la Tecnología (FECYT).
- OMPI (Organización Mundial de la Propiedad Intelectual). 2017. *Manual de la OMPI de redacción de solicitudes de patente*. Ginebra: OMPI.
- Orozco, Luis Antonio y Diego Andrés Chavarro. 2010. "Robert K. Merton (1910-2003). La ciencia como institución". *Revista de Estudios Sociales* (37): 143-162.
- Ortiz, Sara y Álvaro R. Pedroza Zapata. 2006. "¿Qué es la gestión de la innovación y la tecnología (GInnT)?" *Journal of Technology Management & Innovation* 1 (2): 64-82.
- Porter, Alan L. y Scott W. Cunningham. 2004. *Tech Mining: Exploiting New Technologies for Competitive Advantage*. Hoboken, New Jersey: John Wiley & Sons.
- Sánchez Guzmán, M. V. 1999. "Patentometría. Herramienta para el análisis de oportunidades tecnológicas". Tesis, Doutorado em Gerência de Informação Tecnológica, Facultad de Economía, Universidad de La Habana, Cuba.
- Sánchez, Jenny Marcela, Javier E. Medina y Andrés Mauricio León. 2007. "Publicación internacional de patentes por organizaciones inventores de origen colombiano". *Cuadernos de Economía* 26 (47): 247-270.
- Spinak, Ernesto. 1998. "Indicadores científicos". *Ciência da Informação* 27 (2).
- Trejos, Alberto, Graciana Rucci, Marina Bassi, Alessandro Maffioli, Jorge Cornick, Alejandro Rasteletti, Carlo Pietrobelli, Fernando de Olloqui, Gabriel Casaburi y Manuel R. Agosin. 2014. *¿Cómo repensar el desarrollo productivo? Políticas e instituciones sólidas para la transformación económica*. Washington: Inter-American Development Bank. <https://publications.iadb.org/handle/11319/6634>
- Zhou, Xiao, Yi Zhang, Alan L. Porter, Ying Guo y Donghua Zhu. 2014. "A Patent Analysis Method to Trace Technology Evolutionary Pathways". *Scientometrics* 100 (3): 705-721.

Para citar este texto:

Gómez Hurtado, Rosa Elena. 2018. "Tendencias de la innovación tecnológica en Colombia 1991-2013 a partir del análisis de patentes". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 133-150.
<http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57859>

Redimensionamento do espaço e do tempo: a incerteza e os desafios da sociedade, da informação e do planejamento do marketing

Roberto Vilmar Satur*

Júlio Afonso Sá de Pinho Neto**

Emeide Nóbrega Duarte**

Artículo recibido:
12 de abril de 2018
Artículo aceptado:
21 de septiembre de 2018
Artículo de revisión

RESUMEN

Este artículo tiene el objetivo de contextualizar lo que la posmodernidad reserva para la sociedad y para lo social, desde la perspectiva de la información y del marketing. La información, que ejerce la función esencial de alimentar las estrategias de marketing, cambia en conjunto con el ambiente, en consecuencia, si el ambiente es posmoderno, la información tendrá esas características también. Como el marketing actúa en el escenario vigente, necesita actualizarse para ese nuevo contexto adecuando todas sus formas de actuación. Considerando un ambiente donde el tiempo se ha transformado en el ahora, y el espacio es el aquí, queda una interrogación muy grande sobre si es posible todavía planificar en ese

* Departamento de Mediações Interculturais, Universidade Federal da Paraíba, Brasil
robertosatur@yahoo.com.br

** Programa de Pós-Graduação em Ciência da Informação, Universidade Federal da Paraíba, Brasil
sadepinho@gmail.com emeide@hotmail.com

contexto. Los cambios tecnológicos han contribuido a acortar la perspectiva del tiempo y el espacio, de modo que se tiende a que todo sea pensado y hecho en ese contexto, incluso el marketing y su necesario uso de la información. Este artículo estudia algunas tendencias de ese ambiente posmoderno, de esos cambios y sus impactos en la perspectiva del marketing y de la información. Finaliza construyendo la perspectiva acerca de cómo debe ser concebida la planificación estratégica del marketing y el uso de la información en ese nuevo contexto social, en una construcción reflexiva que no descarta una de las principales características de la posmodernidad: la incertidumbre.

Palabras clave: Información; Sociedad Posmoderna; Espacio-tiempo; Incertidumbre; Marketing

Re-dimensioning space and time: The uncertainty and challenges in society, information and marketing planning

Roberto Vilmar Satur, Júlio Afonso Sá de Pinho-Neto and Emeide Nóbrega-Duarte

ABSTRACT

This article aims to contextualize what postmodernity holds for society in terms of information and marketing planning. The information, which plays an essential role in development of marketing strategies, changes hand in hand with the environment. As such, if the environment is postmodern, the information will exhibit postmodern features as well. The way marketing works, including strategic planning, must be upgraded in order to operate in this current scenario. In an environment where time and space mean the here and now, we may well ask if planning is even possible. Technological changes have served to telescope both time and space, and this exerts impacts on how marketing exploits information. This paper therefore examines some of the features of this postmodern environment, contextualizing their impact on both marketing and information usage, concluding by offering perspectives on how to conceive of strategic planning of marketing and the use of information in this new context, which does not ignore uncertainty as a major feature of postmodernity.

Keywords: Information; Postmodernity society; Uncertainty; Space-time; Marketing

INTRODUÇÃO

Neste artigo, apresenta-se uma discussão teórica acerca dos novos desafios que o marketing tem em seu objeto, que é a troca, especialmente no elemento que mais alimenta toda a sua estratégia, seu planejamento e sua ação, que é a informação, considerando o ambiente pós-moderno. Embora se saiba que muitos autores não consideram a atualidade como algo pós-moderno, tampouco uma transição, optou-se por chamar esse momento de “transição para a pós-modernidade” visando marcar a diferença atual com a do tradicional estilo moderno (metade do século XX e anterior) e o que está por vir que pode o período pós-moderno. Sem entrar no mérito do uso do termo pós-moderno, o fato é que se está em um ambiente de transição e de grandes mudanças. Nesse novo tempo o principal foco é a nova configuração de tempo e espaço e o papel das tecnologias.

Também é importante lembrar que, apesar do modismo da expressão pós-modernidade, sabe-se que muito do que existe ainda tem traços fortes da modernidade e nem se imagina que a pós-modernidade irá romper com tudo o que é moderno, muito menos descartá-los, partindo do zero. O novo ambiente se reconfigura sobre o atual, reutilizando alguns, descartando outros, mas, principalmente, mudando a forma de uso e as relações.

O marketing é uma das áreas da gestão que mais rapidamente se atualiza aos novos cenários. Isso porque uma de suas missões é estar atento ao que acontece em seu campo de atuação e rapidamente se readequar a ele. Se o ambiente atual é pós-moderno ou tende ao pós-moderno é evidente que a informação disponível, de que tanto o marketing precisa, será fruto desse ambiente. Cabe ao marketing entender e se conectar a esse novo momento para conseguir as informações.

Dentro dessa perspectiva, podem-se destacar algumas mudanças efetivamente ocorridas, entre elas, o tempo e o espaço que mudaram substancialmente, influenciados, sobretudo, pelo avanço da ciência e da tecnologia, que encurtou distâncias e tempo, com melhores recursos computacionais e melhores formas de comunicação e de transmissão de dados, informação e imagem em tempo real. Isso tem feito o virtual ocupar espaços antes exclusivos do real. Aliado a isso surgem mudanças de postura, de valores, de política, de economia, de relacionamento, de crenças e tantas mais.

METODOLOGIA

Este artigo trata-se de um estudo com características de pesquisa qualitativa, de revisão teórica e propositiva, pois usa uma base de informações de

natureza secundária (publicados) para explicar a nova dimensão que as tecnologias causaram no espaço, no tempo e sobre as informações e seu impacto no marketing, no seu planejamento e nas estratégias. E na última etapa do artigo ele é propositivo, pois tratam-se de construtos elaborados pelos autores do artigo de modo a deixar uma contribuição para o assunto em questão visando melhor entender a realidade atual para atuar nela. Com isso os autores dão um passo além do que já se conhece tradicionalmente, considerando, agora, um ambiente transacional para o pós-moderno, com novas dimensões de tempo e espaço, especialmente em relação à busca e ao uso da informação.

Este estudo qualitativo, que envolve pesquisa bibliográfica, pois reúne citações de um conjunto de materiais escritos, já elaborados e publicadas por outros autores, e também elabora proposta de entendimento sobre o caso estudado, pois finaliza com características de pesquisa-ação (propositiva), em que propõe formas de melhor entender a realidade e assim poder agir de forma mais eficaz, foi desenvolvida a partir de 2014 no Programa de Pós-Graduação em Ciência da Informação, e focava como estratégia a discussão sobre as mudanças das tecnologias, especialmente as da informação e da comunicação, sobre a realidade da sociedade atual e seus efeitos sobre a informação no seu contexto do planejamento do marketing, especialmente considerando o redimensionamento do espaço tempo e as incertezas decorrentes de tais mudanças.

Dentro desta estratégia buscou elencar algumas relevantes publicações e autores que abordassem as mudanças do tempo e espaço decorrentes do avanço tecnológico e o ingresso na pós-modernidade, discutindo o efeito disso sobre a informação e, por conseguinte o impacto destas mudanças sobre o planejamento e as estratégias do marketing. Evidentemente que aliou-se a isso reflexões sobre essas mudanças. Por essa razão os tópicos foram estrategicamente desenvolvidos com essa sequência lógica:

- (I) *o marketing e a informação enfrentando ou absorvendo mudanças (o que está acontecendo)*. E para desenvolvê-lo após pesquisa teórica em diversos autores selecionou-se os que seguem por contribuir com os propósitos desse item. Os selecionados para essa etapa foram: McKenna (1992); Santos (1988); Stewart (1998); Hooley, Saunders e Piercy (2001); Sabbag (2007); Sina (2008); Monteiro e Abreu (2009); Trigo, Soares e Quoniam (2012); Keegan e Green (2013); Sousa *et al.* (2013); Bembem, Santos e Pinho Neto (2014); Perrotti e Pieruccini (2014); Satur, Paiva e Duarte (2017) e a *American Marketing Association* (AMA, 2018). Embora esses não sejam os únicos autores da área, mas eles representam bem a discussão sobre o tema;

- (II) *mudanças do marketing em função da informação e do ambiente (os impactos e as repercussões)*. Para buscar responder a esse tema, após a revisão de teóricos com publicações de diversos anos chegou-se selecionou-se os autores que seguem: Freitag (1986); Santos (1988); Chesneaux (1989); Baudrillard (1992); Weissberg (1993); Quéau (1993); Virilio (1993); Serra (1998); Latour (1994) e Latour (2012); Lèvy (1999); Barreto (2005); Pinho Neto (2008); Vargo e Lusch (2011); reforçados com argumentos de Bourdieu (1997) e Sina (2008). O uso de autores a partir do final dos anos 80 e intensamente dos anos 90 para cá não é coincidência. É a contar dessa época que a academia (universidades, centros de pesquisa e pesquisadores) começam a dar maior atenção ao fenômeno oriundo das novas tecnologias e do redimensionamento espaço e tempo;
- (III) *adequando o planejamento de marketing na pós-modernidade (tentativa de apontar o que se pode fazer a partir dessa nossa perspectiva)*. A partir do que os autores apresentaram como tendência os autores do presente artigo tentaram explicar essa perspectiva de uma forma didática e construíram construtos tentando demonstrar a nova realidade que se apresenta. Nessa parte a contribuição é dos autores do artigo, construindo uma proposta didática de entendimento da nova realidade, baseados nas leituras e releituras feitas nos autores.

O MARKETING E A INFORMAÇÃO ENFRENTANDO OU ABSORVENDO MUDANÇAS

A *American Marketing Association* (AMA), tem atualizado constantemente o conceito de marketing acompanhando as mudanças atuais. A última conceituação o define como “a atividade, ou conjunto de instituições e processos para a criação, comunicação, entrega e troca de produtos e serviços que têm valor para os consumidores, clientes, parceiros e a sociedade como um todo” (Keegan e Green, 2013: 6). Mesmo assim, a pós-modernidade sempre o desafiará a continuar se atualizando e sempre haverá déficit no conceito, porque conceito, normalmente, tem características estáticas.

Para evidenciar essas mudanças e, automaticamente, seu impacto no marketing, destacaram-se quatro subgrupos relevantes de mudanças percebidas. O primeiro inclui dez destaques de McKenna (1992); o segundo menciona as mudanças apontadas por Hooley, Saunders e Piercy (2001); o terceiro, um esboço de algumas mudanças destacadas por diversos outros autores.

I) McKenna (1992) já alertava que existe um *in* que afeta diretamente o marketing na atualidade: a *incapacidade* de fazer frente ou de não perceber a habilidade de outro *in*, ou seja, os concorrentes *intangíveis*. A seguir os dez desses concorrentes *in*, citados a seguir de forma adaptada e com um complemento explicativo.

| Item | Concorrente <i>in</i> | Explicação básica |
|------|-----------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Mudança | Costuma aparecer sem perguntar se pode. É ditadora. |
| 2 | Resistência à mudança | Temos mentalidade moderna, gostamos das coisas como estão, ao menos por um bom tempo. |
| 3 | Consumidores bem informados | É fácil lidar com consumidores desinformados, o contrário dá trabalho. E não se quer mais trabalho. |
| 4 | A mente do consumidor | Consumidor não se comporta apenas racionalmente, mas também por conta do afeto, da emoção, do ambiente e do momento. Perceber e medir isso é difícil. |
| 5 | A mentalidade da pasteurização | A massificação da produção é boa para baixar custos, mas, com o passar do tempo, todos farão isso, então não haverá diferenciação, e a única forma de concorrer será se digladiando pelos preços. Assim, os ganhos são absorvidos pela concorrência desenfreada. |
| 6 | A mentalidade de grandeza | Empresas grandes tendem a trilhar o caminho seguro, de menos risco, produção em escala e acabam inibindo a inovação, "atolando-se" na burocracia e na massificação da produção e da venda. |
| 7 | Cadeias rompidas | Muitos esquecem que todas as situações estão inter-relacionadas e acabam priorizando algumas etapas da cadeia, que vai do produtor ao consumidor. No entanto, numa rede de relações, a força da cadeia é medida pelo elo mais fraco dela, e é grande a chance de esse elo ser exatamente o que foi esquecido. O elo mais fraco dará o ritmo de toda a cadeia, independente dos outros serem fortes. |
| 8 | O conceito do produto | O próprio conceito de produto, se não for atualizado, torna-se um concorrente, uma dificuldade. O que se vende na atualidade são serviços, satisfações. O produto é apenas um artefato que carrega em si tais serviços e satisfações. Focar no produto não funciona adequadamente. |
| 9 | Coisas que aparecem da noite para o dia | Foi um <i>iceberg</i> que afundou o Titanic, ele apareceu da noite para o dia. Todo dia aparecem novos " <i>icebergs</i> " nas organizações e no ambiente profissional. |

| | | |
|----|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10 | Você (eu) | Você pode ser o seu concorrente mais difícil. Máquinas e produtos não competem, mas as pessoas competem entre si, ou contra si. Usam produtos e máquinas para isso. Ao subestimar suas ideias, ao não ter autoconfiança ou ao achar que é infalível, que sabe tudo, ao não querer mudar, você concorre contra você mesmo. |
|----|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Quadro 1. Os concorrentes *in* do marketing

Fonte: Adaptado de McKenna (1992), com agregações dos autores

II) Juntam-se a todos esses “*in*” outros fatores que reforçam o ambiente de mudança e de complexidade que a pós-modernidade traz consigo, gera-se uma situação em que não se pode querer mudar, mas é preciso fazê-lo. Hoo-ley, Saunders e Piercy (2001) destacam alguns fatos relevantes nesse sentido:

| Item | Ambiente de mudança |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Enquanto a Revolução Industrial levou 60 anos para dobrar a produtividade por pessoa, a China e a Coreia do Sul conseguiram fazer isso em 10 anos. |
| 2 | O que se gera de inovação e conhecimento novo por ano, na atualidade, fruto de pesquisa e de desenvolvimento, supera todo o conhecimento gerado até os anos 60, ao menos em quantidade. |
| 3 | Os fabricantes de computadores reduzem em 30% os custos por ano para continuar competitivos. |
| 4 | Consumidor mais informado é mais exigente, quer mais pagando menos. Está também querendo que empresas tenham compromissos adicionais (sustentabilidade, direitos humanos). |
| 5 | Consumidores estão migrando para ofertas de valor alternativo, e não, somente, o modo clássico de valor. |
| 6 | Restaram dois tipos de organização: as que têm departamento de marketing (moderna) e estão, aos poucos, desaparecendo e as que têm alma de marketing, que respiram marketing no todo da organização. |

Quadro 2. Mudança não é uma questão de escolha, mas de necessidade e velocidade

Fonte: Adaptado de Hooley, Saunders e Piercy (2001), com agregações dos autores

III) Podem-se destacar, ainda, outros autores que trazem exemplos tão importantes como os anteriores e reforçam esse cenário de mudanças.

| Item | Ambiente de mudança acelerado pela tecnologia |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | A cada dois dias, o Yahoo grava 24 terabytes de informações sobre o que fizeram seus usuários. (Keegan e Green, 2013) |
| 2 | De 2003 a 2010, na China, o número de usuários de internet saltou de 68 bilhões para 450 bilhões. Isso transformou esse país no maior <i>e-commerce</i> do mundo. (Keegan e Green, 2013) |

| | |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3 | Gutenberg transformou todo o mundo em leitor, a Xerox transformou todo o mundo em editora, os PCs transformaram todo o mundo em escritores, e a internet transformou todo o mundo em atores, co-criadores e produtores em redes digitais. (Keegan e Green, 2013) |
| 4 | Em 1981, Bill Gates dizia que 640 kilobytes eram mais do que suficientes para qualquer um. Em 1977, Ken Olson, da Digital Equipment Corp., dizia não ter razões de alguém querer ter um computador em casa. E o presidente da IBM, Thomas Watson, em 1943, disse: "Penso que há, talvez, no mundo, um mercado para cinco computadores". Erraram, todos os seus prognósticos. O sucesso das empresas tecnológicas foi tão significativo que, atualmente, tem uma bolsa de valores exclusiva para elas: Nasdaq. (Sina, 2008: 17-18) |
| 5 | Várias estimativas mostram a ordem de grandeza de vários exabytes (1018 bytes) de nova informação de que, todos os anos, é colocada na web, crescimento acima dos 20% anuais. Isso significa que, mesmo se considerarmos a divisão em disciplinas no conhecimento, a quantidade de informação nova disponível em cada segmento do conhecimento ultrapassa as capacidades de absorção do cérebro humano. (Trigo, Soares e Quoniam, 2012: 67) |
| 6 | Em 1994, tempos, ainda, de início da internet, já se dizia: "[...] na internet, tudo se passa tão rapidamente que um ano civil correspondiam a sete anos de evolução na internet. Essa realidade é, hoje, ainda mais complexa, uma vez que passamos para níveis de velocidade nos quais um ano na internet corresponde a 10, 15 ou, até mesmo, 20 anos civis". (Trigo, Soares e Quoniam, 2012: 67) |
| 7 | "todos os anos, são produzidos aproximadamente 1,5 bilhões de gigabytes, armazenados nos mais variados suportes de informação, o que fornece em média de 250 megabytes de informação para cada ser humano". (Reis, 1997, em Sousa <i>et al.</i> , 2013: 9) |
| 8 | Na pós-modernidade, robôs dançam valsas, temos som de flauta que não vem da flauta, computadores fazem trilhões de operações por segundo, máquinas registram nanossegundos, coisas que a mente humana não daria conta sozinha, nem no tempo e nem no espaço. (Santos, 1988) |
| 9 | "[...] metade do que o calouro de engenharia aprende ao entrar na faculdade estará obsoleto quando ele se formar; a obsolescência do conhecimento em eletrônica é tão rápida que os especialistas em informática usam a expressão 'idade da internet' da mesma forma que as crianças falam em 'idade de cachorro'". (Stewart, 1998: 63) |
| 10 | O tempo entre a descoberta científica e sua transformação em produto tecnológico diminuiu 60% do início do Século XX até o final deste. (McKenna, 1992) |

Quadro 3. Mudança ainda mais velozes com o advento da tecnologia computacional e da internet
Fonte: Adaptado dos diversos autores citados no quadro

Quem entender bem e mais rapidamente tudo isso, se é que é possível usar a palavra 'tudo', terá vantagem, pois pode se antecipar utilizando novas estratégias de marketing ou atualizando as anteriores.

Como já mencionado, o objeto do marketing são as trocas, e seu objetivo é de agregar valor, dentro de sua cadeia, em curto e em longo prazo para todos os envolvidos no processo ou atingidos por ele. Para isso, o marketing

faz gestão de promessa e de expectativas. Promessa se faz com informações, e expectativas se criam pelas promessas feitas. O desafio da gestão do marketing é de garantir que o produto ou o serviço entregue garanta que seus atributos confirmem as expectativas geradas pela promessa, porquanto o consumidor não compra mais produtos e serviços, mas o benefício que ele gera. É um desafio complexo no ambiente descrito acima, com tantas mudanças e com o consumidor cada vez mais informado, exigente e inquieto, digamos, pós-moderno.

No marketing, fala-se em local de origem (espaço) como algo que pode ser utilizado como estratégia. Mas com as novas tecnologias e o ciberespaço, ficou difícil manter isso estático. Todos os lugares fazem trocas, compram e vendem. Com o ciberespaço e a internet, as trocas continuam ocorrendo, mas é difícil dizer o “lugar” da ocorrência. Redefiniu-se o espaço: a empresa pode estar no principal centro comercial físico do mundo ou em um quatinho no fundo de uma casa, em qualquer país que o processo de consulta, acesso e negociação *online* é o mesmo e independe do cliente pode ser o vizinho da empresa ou estar há milhares de quilômetros de distância. Compra-se, e o produto, em muitos casos, será entregue pela própria rede mundial de computadores, via *download* ou *login* com senha exclusiva de acesso *online*. O que continua distinto ainda é o processo de entrega quando ela é física e exige deslocamentos dos produtos ou das pessoas (em caso de viagens e turismo). Mas é inegável que a comercialização ganhou um novo espaço de atuação, o ciberespaço. E se for serviço, o ciberespaço pode ser também espaço de entrega.

Expandir negócios agora não envolve apenas abrir filiais em novos espaços físicos, contratar mais representantes e outros. A expansão pode ser a implantação de um novo site na internet que permita *e-commerce*. A noção de local para instalar a nova filial ganhou um concorrente de peso: o ciberespaço.

A Amazon Books é um exemplo de uma empresa que preferiu o mercado virtual em detrimento do físico, consolidou-se no mercado mundial e continua apostando mais nas indicações do cliente sem investir em publicidade. O sucesso da empresa é inquestionável. O que não se sabe é se, na era digital, ainda poderia se chamar de “boca a boca”. Já o Google busca também lucros no mercado virtual, mas aposta na publicidade como um desses meios. Também é um sucesso contundente (Sina, 2008). Isso para ficar em apenas dois exemplos de muitos outros possíveis.

Com ou sem publicidade, ambas não prescindem do marketing, da tecnologia e da informação como instrumentos e fatores fundamentais à sua atuação, e isso lhes garante o sucesso. Esse processo elimina a necessidade de grandes investimentos físicos, equipe de vendas de plantão à espera de que o

cliente entre pela porta. Também permite eliminar grandes estoques e custos logísticos desnecessários, pois pode fazer a venda e, a partir dela, disparar a ordem de entrega direto do próprio fornecedor. Assim, o produto sai do fornecedor direto para o cliente sem ter passado pela empresa que o vendeu.

Não se atua mais no mesmo mercado, ou melhor, não existe somente o mesmo mercado. Trata-se do espaço (mundo) virtual e do espaço (mundo) real que se complementam. É o ciberespaço, com suas múltiplas formas possíveis de interação e de convivência, que mudou a noção de estar no espaço, de circular por ele, de interagir nele.

Pelas plataformas digitais e a inteligência coletiva, assimila-se a informação, cujo foco é a construção do conhecimento, por meio da cooperação, da troca de informações e de comunicação com a interação em rede. Assim, a sociedade e a tecnologia produzem o indivíduo, que também as produz, pela interação. Por essa razão, a tecnologia não pode ser vista como ameaça, tampouco como salvadora, mas como facilitadora das relações e das interações humanas, nesse novo tempo e espaço do aqui e agora (Bembem, Santos e Pinho Neto, 2014). Nesse novo dimensionamento se criam “vínculos simbólicos entre os referentes, espaços de transição, pontos de convivência que tornam possível o ‘viver juntos’, em especial em época de mudanças que alteram relações de tempo e espaço, via tecnologias virtuais” (Perrotti e Pieruccini, 2014: 11).

Nessa perspectiva, segundo Monteiro e Abreu (2009: 1) ocorrem agenciamento maquínicos, muito característico do pós-moderno. Não se trata de uma oposição homem x máquina mas sim de um contexto “onde homem e máquina e signos se hibridizam nas múltiplas maneiras de organizar e buscar o conhecimento no ciberespaço, seja por meio da etiquetagem social, seja por ontologias ou pelos mecanismos de busca”. Homem e máquina convivem em plena interação, complementam-se e utilizam esses avanços. Evidente que todo o restante foi ao menos perturbado com tais mudanças, pois dizer que tudo mudou seria um pouco pretensioso.

O ciberespaço trouxe o espaço para perto, e a Internet nos inundou de dados e informações nunca antes vistos. “[...] em nenhum outro período da história, as respostas podem antecipar-se às questões, uma vez que os cenários estão realmente diante dos olhos dos executivos, tamanha é a quantidade de informação disponível” (Sina, 2008: 320). Então, se a informação está mais próxima e imediata como nunca antes visto, em quantidade e, de alguma forma, em qualidade, vale a tendência apontada por Harris de que, segundo Sabbag (2007), a relevância da informação na atualidade tem se sobressaído tão grandemente em relação aos demais fatores e variáveis que está tomando o lugar da autoridade nas relações interpessoais, inter e intraorganizacionais.

Mas, junto com a quantidade demasiada de informações disponíveis, é possível detectar um expressivo número de informações inúteis, enganadoras, ambíguas, assimétricas, incompletas, imperfeitas, excessivas, complexas, dúbia, duplas, insinuantes, contraditórias e falsas (Satur, Paiva e Duarte, 2017). A informação, que veio para facilitar as ações de marketing, diante do excesso, também tem exigido melhores critérios e mais empenho para selecioná-la e classificá-la, com filtro, procurando “agulhas dentro de um palheiro”.

MUDANÇAS DO MARKETING EM FUNÇÃO DA INFORMAÇÃO E DO AMBIENTE

No método tradicional, o marketing utiliza um modelo sequencial e lógico de planejamento: primeiro, busca conseguir informações suficientes; depois, procura pensar estrategicamente sobre seu planejamento no tempo e no espaço. Somente depois ocorre a ação, e finalmente, a avaliação dos resultados para readequar o planejamento. Todavia, no ambiente pós-moderno, tudo acontece agora, aqui. O tempo tende a se contrair na instantaneidade. Nessa nova reconfiguração, tudo acontece ao mesmo tempo, num espaço remodelado.

O marketing sempre está em busca de informações e assim continuará, especialmente as emitidas pelos atores (consumidores, *stakeholders*, usuários). Nesse novo cenário, é preciso perceber que esses atores também mudaram seu comportamento e o espaço onde circulam. Além do fato de que as trocas de informações ocorrem com rapidez, independentemente de ela estar perto ou longe, ser do passado, do presente ou uma projeção do futuro também precisa se considerar a busca por informação precisa seguir a lógica de Latour (1994 e 2012), que envolve seguir rastros, suspeitas deixadas pelos atores que liberam a informação. Isso afeta sobremaneira a perspectiva clássica de mensurar, medir e projetar. E afeta ainda mais ao saber que os rastros e as suspeitas não estão mais apenas no espaço físico mas também no ciberespaço.

A complicação ou complexidade aumenta quando o próprio ator não está entendendo muito bem o seu novo espaço. Sobre isso, pode-se reportar a Chesneaux (1989), que afirma que *o homo mundialis modernicus* anda desorientado com essa modernidade. Ele está constrangido com os ajustes temporais ávido de ganhar tempo: a sociedade de consumo consome muito tempo. Ele está frustrado, neurótico e tenso. Está ocupadíssimo, fazendo muitas coisas ao mesmo tempo e clama por tempo livre. Mas, quando tem tempo livre, fica com a impressão de que está perdendo tempo. Trata-se de uma sociedade sincrônica integral. O tempo virou um sistema que permite fugir do real, com o advento do virtual que, em muitos casos, confunde-se com esse

real e até é mais agradável do que ele. E pelo virtual pode-se ir para o passado e para o futuro no mesmo instante. O espaço também mudou, pois, com a virtualidade, é possível viajar pelo mundo acessando um computador. O mundo agora está “aqui”. O tempo e o espaço deram lugar ao prazo (Chesneaux, 1989). E normalmente esse prazo é o agora.

Trata-se de um ambiente imediatista, angustiado e confuso. Como afirma Virilio (1993), a pós-modernidade gerou uma confusão entre o mundo real e o virtual, como tempo e espaço. Fusão de vários lugares ao mesmo tempo, além das aparências sensíveis. Com a virtualidade, não há mais tempo fixo no espaço. Vive-se numa transparência ou trans-aparência transmitida a distância instantaneamente, em que os pixels das imagens são transmitidos e recebidos pela tele-ação, tele-audição, tele-visão.

Relembrando Latour (2012), que recomenda seguir rastros e suspeitas para identificar o comportamento dos atores. Nesse caso, surge uma pergunta: como seguir suspeitas e rastros das informações deixadas pelos atores nesse contexto? Afinal, o rastro e a suspeita podem ser tão somente uma imagem, um simulacro, uma aparência.

O ator que emite informações é um ser que atua nesse novo ambiente com o tempo e o espaço redimensionados e recriados. Todavia não significa que é um ator mais autônomo. Na verdade, a alienação continua e agora é provocada pela Ciência, pela tecnologia, pela aparência e pela imagem. Nela a razão se voltou contra quem a pensou, e a razão abrangente e humanista se atrofiou (Freitag, 1986).

Logo, o planejamento estratégico de marketing abrangente, racional, focado na essência pode ter problemas ao retratar esse contexto. Pois, corre-se o risco de se focar num ator alienado que somente foca na aparência e na imagem.

Ou seja, avanço da tecnologia e da Ciência não garantiu a tão sonhada autonomia do cidadão, mas facilitou a condição para que a sociedade aderisse a uma cultura, a um consumo e a um comportamento de massa, sem posições e reflexões políticas profundas. Não garantiu a libertação, mas fortaleceu a dominação, a alienação e a sedução. Uma espécie de autonomia ao contrário. Nesse contexto, por exemplo, ganha a eleição quem usar as informações de modo a promover o melhor espetáculo, com um show agradável aos olhos e aos ouvidos do eleitor. Assim, os partidos viram “produtoras de publicidade” e os candidatos “mercadorias” a serem vendidas (Serra, 1998).

A ética do marketing tem defendido que as ações de marketing devem evitar explorar públicos vulneráveis. Tradicionalmente se conhecia públicos vulneráveis como sendo crianças, idosos, deficientes, pessoas com déficit educacional etc. Ou seja, pessoas com fragilidades financeiras, psíquicas, emotivas, etc. Isso também envolvia pessoas com estado emocional alterado,

endividados, pessoas sob pressão, carentes de afetividade, ignorantes ou desinformados sobre aquela assunto, com dependência química, compulsivos e outros. O marketing ético defende que não se pode usar da fragilidade das pessoas para as convencer a fazer algo que não seria o melhor para elas. No entanto, na pós-modernidade isso fica ainda mais complexo por que, se considerarmos a alienação como uma fragilidade, praticamente toda a população poderia ser considerada vulnerável ou em estado de vulnerabilidade, ao menos em algumas situações.

Soma-se a isso exclusão do acesso ao saber como um estado de vulnerabilidade, que se intensificou com as novas mídias. Embora as redes informacionais permitam um maior acesso à novas formas de ensino, à novas aprendizagem, à mais informação e à mais conhecimento, ainda assim, a exclusão persiste (Serra, 1998). Pois novas exclusões foram geradas levando em conta que muitos não têm acesso a esses recursos, outros não sabem os usar plenamente e muitos estão conformados em viver na sociedade do espetáculo, da informação superficial, do modelo de consumo em massa. Nessa perspectiva, fica difícil definir até onde vai a vulnerabilidade que o marketing não pode eticamente explorar.

Soma-se a isso o fato de que, na chamada pós-modernidade, o consumo deixou de ser gozo, necessidade e prazer, para ser um fazer gozar por objetos, serviços e especialmente signos. Faz-se de tudo que se quer, manipula-se, controla-se. Comunicar não é mais falar, mas fazer-falar; informação não é saber, mas fazer-saber. O verbo fazer (operacional) e não mais o de ação (crer, gozar, consumir, agir) tomou a dianteira de tudo. Tudo agora é fazer-agir, fazer-crer. E mesmo que o indivíduo não queira, será convencido a se fazer-querer. O fazer por persuasão ou dissuasão. O verbo ação deu lugar definitivo ao verbo operacional (fazer), que é algo produzido, induzido, solicitado, manipulado, tecnificado, midiaticado. A comunicação é, agora, o fazer-comunicação. A informação é o fazer-informação, do contrário, ela não repercutirá (Baudrillard, 1992).

Pode-se dizer, então, que o agir do marketing também passou a ser o fazer-marketing. E se faz com serviços. As pessoas não compram produtos, mas serviços, o benefício que ele proporciona. O objeto é apenas o artefato que carrega em si os atributos que permitem que tal benefício. Segundo Vargo e Lusch (2011), no mercado, tudo agora é B2B (*business to business*) e envolve A2A (ator a ator), num sistema em rede, em que se foca a criação de valor. Está em voga a lógica dominante em serviços (SDL), segundo a qual todos nós somos dependentes das relações e dos serviços que os outros nos prestam e vice-versa. Somos todos seres que servem e são servidos, pelas especialidades. Assim, ficamos todos entrelaçados por um emaranhado de

redes interdependentes. Ao propor o B2B no lugar de B2C (empresa *versus* consumidor), é porque, no B2C, o consumidor tem um papel mais passivo na troca, enquanto no B2B, ambos são “*business*”, atuantes. Já o termo A2A, é mais adequado em tempos de SDL. Na perspectiva A2A e B2B todos os atores estão cocriando valor em marketing, num ambiente em rede e de lógica dominante em serviços (SDL). Trata-se de um ecossistema de serviços e/ou serviços ecopensados, de SDL como evolução orgânica.

Prevalece a interação em rede, as sociedades em rede, a cibercultura, o ciberespaço, o lócus virtual pela noosfera, a convergência de ideias de percepções e memórias (com base na perspectiva de Lèvy, 1999). É a sociedade da informação que usa as TICs, na tecnoeconomia e gera uma nova forma de superar a crise capitalista, permitindo uma nova fase de acumulação e de crescimento. Novas formas de reproduzir o capital (Barreto, 2005).

Também mudou a perspectiva de tamanho, que deixou de ser parâmetro para o real. Para Virilio (1993), tudo, de todos os tamanhos, virtualmente foi reduzido ao tamanho que cabe na imagem, numa tela. As coisas se tornaram planas, próximas e longe, no mesmo plano, sem curvas ou profundidades. A realidade se tornou uma sequência de imagens. O relevo virou uma espécie de alma da pintura, que não dá para tocá-la, ir até ele fisicamente, mas está lá, permite ser imaginado, visto em uma dimensão 3D e até bem observado com esforço mental.

Como captar a informação necessária para o marketing nessa perspectiva? Não dá para pegá-la mesmo ela estando ali. Ela pode ser imaginada e até bem observada num esforço da mente, mas não tem curvas nem profundidade. E mais, se a informação for real, talvez não seja mais um parâmetro da realidade, apenas uma aparência, uma imagem. Não é uma tarefa fácil. A clássica forma de captação e mensuração quantitativa fica cada vez mais irreal considerando-se esses novos atributos. Mas seguir os rastros e suspeitas, também é difícil. Afinal, o que seguiríamos? O real ou o virtual? O próximo ou o distante? A que espaço teríamos que nos ater: ao espaço geográfico ou ao ciberespaço? Tudo ao mesmo tempo, responderiam os autores pós-modernos. Tarefa árdua e espinhosa.

É a complexidade complicando a tarefa do planejador e do pesquisador. Avançam sobremaneira as entidades híbridas. São situações em que o real e o não real parecem ser a mesma coisa. Novos habitantes fazem parte do nosso espaço: ícones, mouses, roteadores levam à emergência de realidades artificiais como forma de apresentar um novo ambiente. A hipertrofia da imagem com comandos informatizados e modelos matemáticos “faz a imagem fazer” coisas que, no real, não faria, faz com que ela não seja mais só representação, mas apresentação com vida própria e tende a não parar nunca mais

de avançar. A representação está assumindo o lugar do objeto representado, numa forma mais ideal que o próprio “objeto” ou “ator”. O virtual como uma nova dimensão do real, e não, seu substituto (Weissberg, 1993). Como já dito, estamos na era do fazer, nesse caso, fazer-parecer com o real, ou fazer-ser melhor que o real.

Imagens infográficas imitam a natureza, traduzem teorias de forma sensível ou nos desconcertam com imagens fragmentadas e, aparentemente, sem sentido. Linguagens simbólicas permitem essas imagens sínteses, essencialmente abstratas, uma aliança entre a forma e o sensível. A numeralização e a virtualização da informação fizeram crescer o papel da tecnologia de representação na sociedade. A imagem deixa de ser metáfora para ser modelo de um caráter mais concreto e experimental à teoria. O modelo pode ser experimentado, testado, confrontado com o real. A imagem síntese é uma versão equivalente ao modelo. A infografia, com sua capacidade de interagir com o espectador, é a grande revolução, que significa imersão interativa. Exemplo: simuladores de voo. A imagem síntese acrescenta uma impressão física forte, envolvente, muito próximo do real, parecendo ser “mais real que o real” (Quéau, 1993).

O marketing, como teoria e ferramental de atuação, precisa considerar todas essas novas perspectivas. É necessário que o marketing desmonte e remonte seu ferramental e suas formas de aplicações, incluindo essa nova perspectiva.

Um *desmontar*, e, posteriormente, um *remontar* do marketing levando em conta que na pós-modernidade, não se vive mais um ambiente de tempo e de espaço harmônico. Vive-se uma nova sensação de irrealidade, com vazio e confusão, onde se cultua o “des” predomin (Santos, 1988): desencanto; desordem; descrença; deserto; desreferencialização do real; desregulamentação da economia; desestelização da arte; desconstrução da filosofia; despolitização da sociedade; dessubstancialização do sujeito; descompromisso; desmantelamento; desregramento e muitos outros “des”. Nesse ambiente, onde o “des” é regra, resulta um ambiente onde a “in” faz parte do dia a dia: a *incerteza* (Santos, 1988).

Numa perspectiva do marketing e da informação, poderíamos dizer que o primeiro terá de perceber que o que aparentemente é informação pode ser desinformação ou apenas informação despolitizada, desordenada, descompromissada, desreferenciada, ou outro ‘des’ qualquer. Isso torna a informação imperfeita e assimétrica por ficar ‘in’: incompleta, incerta e insegura.

Essa obscuridade, ou incerteza, aparenta ser contraditória, se considerarmos Baudrillard (1992), que diz que, na pós-modernidade, estamos transparentes, iluminados por todos os lados a ponto de perder nossa sombra. Estamos superexpostos pelas técnicas, pelas imagens e pelas informações. Como

não temos refratários, absorvemos tudo, branqueando corpos, mentes e memórias. Aparentemente, seria mais fácil ver a “verdadeira informação” mediante tanta “iluminação e transparência”. Mas não é bem assim. Por que isso nos leva a fazermos apenas atividades brancas, sociabilidade branca, vivendo numa assepsia total. “Cirurgia estética” (aparência), “branqueamento” (superficialidade), incerteza radical, simulação indefinida e proibição da negatividade são os caminhos disponíveis. Para isso, “cirurgias amputam” traços negativos do corpo e da mente. Somos remoldados por operação de síntese para ficar mais no estilo ideal, aparentemente mais agradáveis aos olhos e aos ouvidos. Praticamente mata-se o ser humano, e ele se transforma uma redenção geral de signos, aparência, conveniência, estética e imagem. Quase nada sobra do humano anterior que foi sacrificado em nome da geração operacional, das coisas, pela causa, pela meta, pelo objetivo (Baudrillard, 1992).

Nesse caso, a perspectiva do marketing e da informação terá que analisar com quem está se relacionando, para quem está planejando e agindo. Também de quem vem, ou para quem vai, a informação. Estará se direcionando ao ser humano ou a esse outro ser “híbrido remodelado”, redimensionado? Cabe ao marketing reatualizar seu conhecimento, saber, teoria e forma de observar a informação nesse novo contexto de mais aparência e menos essência.

Esse novo ser híbrido faz parte da chamada sociedade da informação, onde surgem os “novos homens do saber” e os “novos saberes” baseados na informação. A função $Y=f(K,L)$ dá lugar a $Y=f(S,I)$, com a riqueza produzida sendo agora função do saber e da informação. Isso deveria permitir uma sociedade que incluía, mais livre, igual e fraterna. Resgataria a ideia original do ciberespaço, em que ‘cyber’ vem do homem do leme. Esse homem do leme daria seu rumo num espaço de novas fronteiras. Todavia, essas promessas utópicas não passaram de sonho. O ciberespaço virou poluição da velocidade, reduziu o mundo a um universo virtual paralelo, criado e sustentado por computadores interligados em redes, num espaço controlável e manipulável. A pergunta que surge é: controlável por quem, de quê, por quê e para quê? Houve evolução tecnológica, mas exercício de poder não mudou. Continua o controle, desigual e excludente (Serra, 1998).

Existe um grande hiato entre o superdesenvolvimento das tecnologias e o desenvolvimento institucional e social. As instituições ainda são modernas, duras, rígidas e, nem sempre, dão conta dessa nova realidade. Muitas teorias também. Quanto ao social, é importante destacar a presença ainda muito forte da exclusão do acesso das mais diversas formas e as várias coisas, inclusive as tais novas tecnologias. O pós-moderno não é ainda o sonhado plenamente incluyente (Pinho Neto, 2008).

Com essa preocupação, também precisam ser contextualizados “o novo marketing” e a “nova informação”. É preciso sempre se perguntar se o que surge de novo é autêntico e espontâneo ou está à serviço do controle, da manipulação e do poder? Se surgiu para incluir ou excluir? Inclusive, a informação captada pode ser pretensamente produzida daquele modo porque interessa ao controlador, ao manipulador e ao poder. Nesse caso, seria algo bem “real” (o poder controlador) controlando um ambiente virtual. Pode-se então dizer que toda riqueza produzida agora predominantemente pelo saber e pela informação é produzida a serviço de interesses, digamos escusos. A exclusão tecnológica também teria este viés. Nesse contexto, surge outra pergunta: é possível acreditar na informação que é gerada? Ela continuaria tendo “somente os velhos defeitos” (incompleta, assimétrica, imperfeita, imprecisa) ou também teria “novos/velhos defeitos” como intencional, branqueada, manipulada, ficcional, superficial, irreal, incerta, incorreta, dúbia e falsa?

Toda essa complexidade não deve levar ao desespero nem a desistência. Esses dois “des” não servem na atuação pós-moderna. Agora um “des” positivo é recomendável: o despertar. Despertar a atenção e o cuidado com as informações que balizarão a atuação do marketing na pós-modernidade.

As inovações tecnológicas que geram o aqui (espaço) e o agora (tempo) mudaram as relações, formaram uma lógica da aceleração, da velocidade e de novos híbridos. Gerando assim três incidências (com interrogações) contraditórias: as redes (promovem o fim dos vínculos territoriais?); a Internet (desvanece os intermediários do espaço público?) e o regime temporal das tecnologias (acelera e retarda?). Fato é que a telepresença é um paradigma cultural pós-moderno, ligado a desterritorialização (Weissberg, 1993).

Trazendo essa argumentação para o marketing, percebe-se que essas três perguntas são relevantes porque mexem em seu mais tradicional modelo, o seu famoso composto de marketing, ferramental conhecido como os 4Ps do marketing (produto, preço, praça e promoção). Esse modelo surgiu e evoluiu num contexto ainda moderno: anos 50 até meados dos anos 70. O modelo dos 4Ps, a partir de então, dominou o cenário dos estudos do marketing e permanece atual, embora já não tão unânime como outrora. Contextualizando do referido modelo com as três incidências contraditórias observadas por Weissberg (1993), vê-se que elas impactam em mudanças internas na forma de pensá-lo, especialmente nos itens promoção (comunicação integrada de marketing) e praça (ponto, local, logística, distribuição), sem isentar os outros dois (produto e preço).

Isso porque, num contexto de desterritorialização, de ciberespaço, de internet e de redes virtuais, como se deve pensar praça e promoção? Onde estaria a praça? No ciberespaço? Se for isso, como localizar o ponto ou o local

de atuação? Na tela do meu computador? E como fazer a logística e a distribuição? Como fica o mercado tradicional? No P de promoção continuaria a comunicação integrada de marketing ou ela passaria a ser desintegrada? Algumas dessas indagações o marketing já respondeu, em outras, inclusive, já se adaptou, mas o processo está distante de ser completo e de se ter certeza de que evoluiu suficientemente. Nada mais pós-moderno do que isso: incompleto e incerto.

ADEQUANDO O PLANEJAMENTO DE MARKETING NA PÓS-MODERNIDADE

Depois de tudo o que foi relatado ainda vale a pena planejar ações de marketing? Evidentemente que essa missão deve continuar, mas agora adequada à nova realidade. Planejar, normalmente, vem antes de agir. Ao menos, era normal até então. Isso porque a nova configuração caótica típica da pós-modernidade, não deixa isso mais tão evidente, sobre o que vem primeiro e depois (ordenação).

Segundo Chesneaux (1989), na pós-modernidade, a partida já é a própria chegada, tudo está presente em toda parte, nada em lugar nenhum, tudo é artificial e virtual. Degradou-se o espaço temporal, aqui, ali, lá, dentro, fora, próximo, distante, antes, agora, depois. Estamos num hiperespaço de acesso direto, instantâneo e generalizado (Chesneaux, 1989).

O planejamento do marketing precisa se adequar a algo que não é simples. Afinal, ver o tempo como o agora quebra um importante pressuposto do planejamento: o ato de pensar no presente sobre o futuro, o pensar para depois agir. E segundo Bourdieu (1997) o pensar, por si só, tem uma pita-da de subversivo, desmonta ideias prontas, faz encadeamentos, apela para a razão, para o conteúdo e para as demonstrações. Para isso precisa de tempo, que agora não se tem mais. Ainda assim, segundo Sina (2008: 20) a teoria do marketing tem acompanhado e evoluído o mercado, “sem deixar de considerar seus fundamentos, conceber, planejar e executar ações capazes de expandir [...], criar valor, agregar valor”.

Se antes, planejar as estratégias do marketing, executar o plano, obter os resultados da ação e avaliar eram ações linearmente e temporalmente pensando, na atualidade, isso não deve ser feito, necessariamente, numa forma escalonada e linear nem com a mesma rigidez e as “certezas” de outrora. Cada vez mais a terminologia “in” toma corpo especialmente na incerteza, e a terminologia “des” predomina em quase todas as situações. Como o planejamento estratégico de marketing é uma forma de se organizar e pensar no futuro em sintonia com o que se vê no presente e no passado, fica explícito

que se trata de uma forma de ordem e organização. Todavia, no ambiente pós-moderno prevalece a desordem e a desorganização. Mas também se sabe que toda desordem causa uma nova ordem, ou novas ordens, raramente idênticas às anteriores. Então, depois da ruptura causada pela “desordem”, um novo cenário se cria (nova ordem) e é a partir disso que deve se pensar o planejamento. Como? Mudando, inclusive, a forma de planejar. Está em xeque a cronologia do planejar e não o planejar em si.

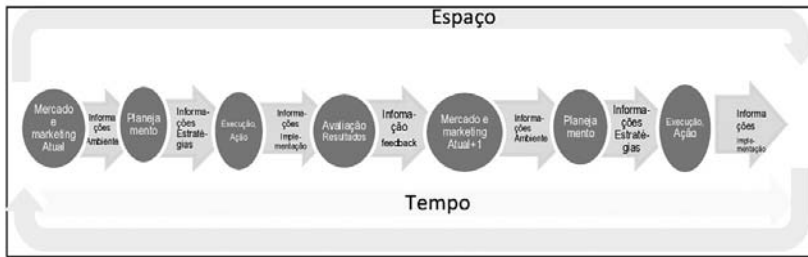


Figura 1. Cronologia Tradicional do Planejamento Estratégico de Marketing
Fonte: Elaboração dos autores

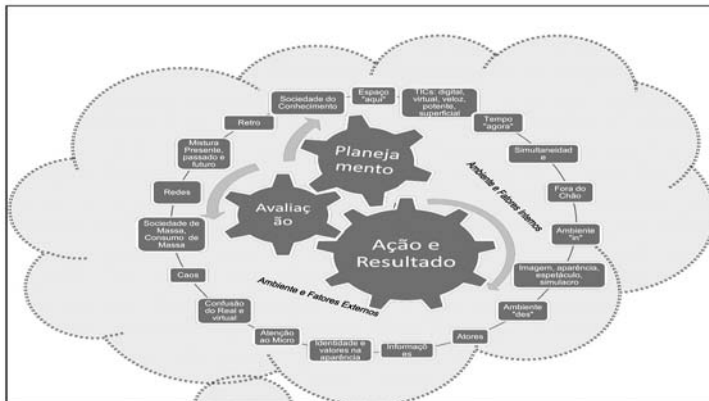


Figura 2. Novos dimensionamentos do planejamento do marketing na pós-modernidade
Fonte: Elaboração dos autores

Nesse novo ambiente do marketing, continua sendo importante o planejamento, a execução, o resultado e a avaliação para repensar o planejamento e executá-lo. A diferença é que tudo acontece ao mesmo tempo. É um *continuum*. Não é um processo encadeado, mas todos se encadeiam ao mesmo tempo, se autoinfluenciam e influenciados tudo em tempo real. O próprio ambiente é instável ao se definir o que é interno ou externo. O formato

externo em forma de nuvem busca demonstrar que não é nem linear nem circular o processo, mas pode tomar várias formas, assim como as nuvens mudam com os ventos. O fato de a área externa estar cercada por linha pontilhada significa que ela é vulnerável e instável tanto ao que vem de fora como ao que vem de dentro.

Assim cronologia, na pós-modernidade, deixa de ser uma linha contínua para ser um rizoma ou, mínimo, uma linha torta.

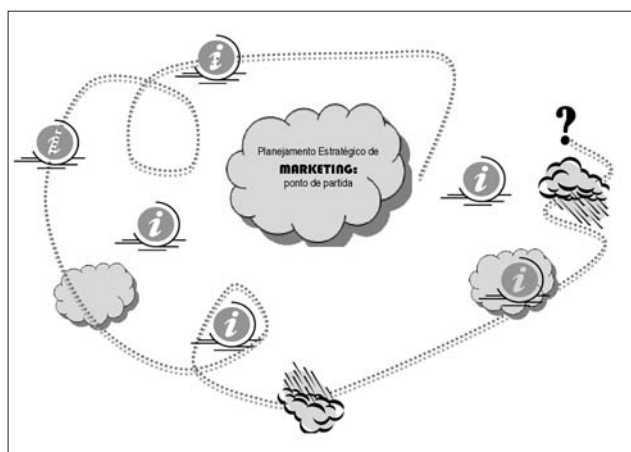


Figura 3. Possível ambiente, caminho e cronologia que encontrará o marketing ao se planejar na Pós-modernidade
Fonte: Elaboração dos autores

Planejar o marketing, nesse novo cenário, exige seguir a direção de onde a informação está e enfrentar diversos cenários. Nem sempre se encontrará a informação necessária, às vezes, se passará ao largo dela, outras vezes, rodeá-la-á, sem conseguir retirar dela o que precisa e, por causa disso, se pagará um preço (situações torrenciais, avessas,...). Ser planejador de marketing, na pós-modernidade, é se adaptar o tempo todo; é seguir o rastro da informação adequada, por mais difícil que seja; é saber que o caminho não é tão bonito como o cenário inicial imaginado (por isso no desenho está pontilhado ou vazado); é admitir que a incerteza faz parte do processo. O importante é nunca parar e ficar o tempo todo, simultaneamente, planejando, agindo, observando resultados, avaliando e retroalimentando. Não existe mais o cidadão que acorda de manhã e diz: “Hoje vou pensar e planejar o futuro”, ou o que afirma: “Hoje vou executar um dos itens previstos no planejamento construído no passado”. O cidadão já acorda sabedor que terá que fazer todas as atividades permanentemente.

CONSIDERAÇÕES FINAIS

Neste artigo, procedeu-se a uma reflexão sobre novos ambientes para se planejar e executar o marketing e os novos pressupostos do uso da informação na pós-modernidade. Se, no passado, fazia-se um planejamento para 10, 15 ou 20, se recentemente se admitia o curto prazo como sendo o tempo de até um ano, médio prazo, de até três anos, e longo prazo, de até cinco anos, com a pós-modernidade, esses prazos não se distinguem mais tão claramente.

Há algumas décadas, as máquinas eram fabricadas para atuarem por até 30 anos. Na atualidade imaginar a vida útil de uma máquina por mais de cinco anos é desnecessário, pois se ela resistir em termos de vida útil, mesmo assim estará obsoleta tecnologicamente. O argumento da qualidade pela resistência e da durabilidade perdeu força para o descartável e o atualizado.

O planejamento estratégico de todas as áreas e do marketing também observam seus prazos se encolhendo. A última trincheira de resistência é essa transição para o aqui e o agora, para o simultâneo. Planejar também é o agora. Acontece, simultaneamente, com executar o plano.

Então, qual o papel da informação nesse planejamento pós-moderno? Continua o mesmo: alimentar o planejamento para ele tenha condições de retratar o que, de fato, precisa ser considerado naquela realidade. A informação continua a principal aliada do planejamento e da execução do marketing, porque dela sempre virão os “rastros” e as “suspeitas” que permitirão importantes antecipações de situações. Mesmo que essa antecipação não seja apenas de dias, horas, minutos ou até milésimos de segundos. Mas já é uma antecipação.

Podemos ter certeza de que tudo funcionará adequadamente como planejado? Não. Na pós-modernidade a única certeza é que teremos incerteza.

REFERÊNCIAS

- AMA (American Marketing Association). 2018. *Definition of Marketing*. <https://www.ama.org/AboutAMA/Pages/Definition-of-Marketing.aspx>
- Barreto, A. M. 2005. “Informação e Conhecimento na Era Digital”. *Transinformação* 17 (2): 111-122.
- Baudrillard, J. 1992. “A brancura operacional”, em *A transparência do mal: ensaios sobre os fenômenos extremos*, 2. ed., 51-57. São Paulo: Papirus.
- Bembem, A. H. C., P. L. V. A. C. Santos e J. A. S. Pinho Neto. 2014. “Tempo do Conhecimento Interativo: Reflexões sobre a Inteligência Coletiva e o Pensamento Complexo”. *Em Questão* 20 (1): 12-31.

- Bourdieu, P. 1997. "O Estúdio e seus Bastidores", em *Sobre a Televisão*, 16-54. São Paulo: Editora Zhar.
- Chesneaux, J. 1989. "Uma Outra Relação com o Espaço e com o Tempo", em *Modernidade-Mundo*, 17-40. Petrópolis: Vozes.
- Freitag, B. 1986. *Teoria crítica: ontem e hoje*. São Paulo: Brasiliense.
- Hooley, Graham J., J. A. Saunders e N. F. Piercy. 2001. *Estratégia de Marketing e Posicionamento Competitivo*, 2. ed. São Paulo: Prentice Hall.
- Keegan, W. J. e M. C. Green. 2013. *Marketing Global*. São Paulo: Saraiva.
- Latour, B. 1994. *Jamais Fomos Modernos: Ensaio de Antropologia Simétrica*. Rio de Janeiro: Editora 34.
- Latour, B. 2012. *Reagregando o social: uma introdução à Teoria Ator-Rede*. Salvador: Edusc.
- Lèvy, P. 1999. *Cibercultura*. São Paulo: Editora 34.
- Matos, O. C. F. 1989. *Paris 1968: as barricadas do desejo*. Tudo é história 9. 3. Ed. São Paulo: Brasiliense.
- McKenna, R. 1992. *Marketing de relacionamento: estratégias bem sucedidas para a era do cliente*. Rio de Janeiro: Campus.
- Monteiro, S. D. e J. G. Abreu. 2009. "O Pós-Moderno e a organização do conhecimento no ciberespaço: agenciamentos maquínicos". *DataGramaZero* 10 (6) (dez.). Acesso em 05 dec. 2014. http://www.dgz.org.br/dez09/Art_05.htm
- Perrotti, E. e I. Pieruccini. 2014. "A mediação cultural como categoria autônoma". *Revista Informação & Informação* 19 (2) (maio/ago.): 1-22. Acesso em 02 oct. 2015. <http://www.uel.br/revistas/informacao/>
- Pinho Neto, J. A. S. 2008. "As Novas Tecnologias de Comunicação e Informação da Transversalidade entre Natureza e Cultura". *Revista Culturas Midiáticas* 1 (1): 1-11.
- Quéau, P. 1993. "O Tempo do Virtual", em *Imagem-Máquina: A era das tecnologias do virtual*, A. Parente (org.), 91-99. Rio de Janeiro: Editora 34.
- Sabbag, P. Y. 2007. *Espiraís do conhecimento: ativando indivíduos, grupos e organizações*. São Paulo: Saraiva.
- Santos, J. F. 1988. *O que é Pós-Moderno*, 5. ed. São Paulo: Brasiliense.
- Satur, R. V., S. B. Paiva e E. N. Duarte. 2017 "Informação imperfeita e seu impacto nas estratégias empresariais". *Brazilian Journal of Information Studies: Research Trends* 11 (2): 07-18. Acesso em 01 jun. 2014. <http://www2.marilia.unesp.br/revistas/index.php/bjis/article/view/5057>
- Serra, J. P. 1998. *A informação como utopia*. Covilhã (PT): Labcom Books.
- Sina, A. 2008. *Marketing global: soluções estratégicas para o mercado brasileiro*. São Paulo: Saraiva.
- Sousa, R. P. M., G. A. Dias, B. Oliveira e M. Sousa. 2013. "O exercício da arquitetura da informação na Web: reflexões relativas à Lei dos Direitos Autorais". *Encontros Bibli* 18 (36) (jan./abr.): 107-128.
- Stewart, T. A. 1998. *Capital Intelectual*. Rio de Janeiro: Campus.
- Trigo, M. R., B. Soares e L. M. Quoniam. 2012 "Inteligência competitiva e inovação estratégica: A IC acompanhando a evolução mundial", em *Gestão da informação, inovação e inteligência competitiva: como transformar a informação em vantagem competitiva nas organizações*, C. Starec (org.), 59-78. São Paulo: Saraiva, pp. 59-78.
- Vargo, S. L. e R. F. Lusch. 2011. "It's all B2B and Beyond: Towards a Systems Perspective of the Market". *Industrial Marketing Management* 40 (2): 181-187.

- Virilio, P. 1993. *O espaço crítico e as perspectivas do tempo real*. Rio de Janeiro: Editora 34.
- Weissberg, J. L. 1993. “Real e virtual”, em *Imagem-Máquina: a era das tecnologias do virtual*, A. Parente (org.), 117-126. Rio de Janeiro: Editora 34.

Para citar este texto:

- Satur, Roberto Vilmar, Júlio Afonso Sá de Pinho Neto y Emeide Nóbrega Duarte. 2018. “Redimensionamento do espaço e do tempo: a incerteza e os desafios da sociedade, da informação e do planejamento do marketing”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 151-173.
- <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57942>

Representación temática y recuperación de los noticieros de televisión en la web: propuesta y aplicación de un modelo metodológico

Catalina Naumis Peña*
Silvano Soto Hernández**

Artículo recibido:
17 de febrero de 2018
Artículo aceptado:
18 de septiembre de 2018
Artículo de investigación

RESUMEN

Los noticieros se trasladan de la televisión a la web para acercarse a los usuarios. La representación textual de su contenido es importante para posicionar la noticia en la web en motores de búsqueda y facilitar su consulta y preservación. El valor económico que adquiere la difusión de los noticieros de la televisión incrementa cuanto más claro es el tema para las tecnologías web. En el presente estudio, basado en las metodologías del tratamiento documental y de la Investigación-Acción, se propone un modelo de representación temática de contenido para optimizar la recuperación a fin de promover el tráfico de usuarios. Se puso el modelo a prueba en un estudio de caso y se presentan los resultados obtenidos.

- * Instituto de Investigaciones Bibliotecológicas y de la Información, UNAM, México
naumis@unam.mx
** IDMS Consulting SC, México
silvano_soto@idms.com.mx

Palabras clave: Noticieros Televisivos en la Web; Representación Temática del Video Periodístico; Posicionamiento Web del Noticiero Televisivo

Subject area representation and retrieval of television news on the web: proposal and application of a methodological model

Catalina Naumis-Peña and Silvano Soto-Hernández

ABSTRACT

Newscasts migrate from television to the web in order to reach more users. The textual representation of news content serves to position the news on web in search engines and facilitate their consultation and preservation. The economic value of television news broadcasts increases with increasing clarity of the enunciation of the subject area for web technologies. Employing documentary management and action research methodologies, this paper proposes a subject area representation model to optimize both retrieval and user traffic. The model is put to the test in a case study and the results obtained are presented.

Keywords: Television News on the Web; Subject Area Representation of Journalistic Video; Web Positioning of Television Newscasts

INTRODUCCIÓN

La representación temática del contenido facilita la visibilidad, recuperación y acceso de los documentos a través de internet. En el caso de las noticias, la recuperabilidad se expresa a través de la consulta de los usuarios en la web o apps y, no menos importante, en la reutilización de los propios medios, sus editores y periodistas (Sobak y Pharo, 2017: 739).

Los teóricos que han estudiado la representación del contenido a través de la indización destacan cómo se orienta cada vez más hacia grupos de usuarios específicos (Hjørland, 2017: 59). En el ámbito de la edición y publicación en la web, la indización se ha especializado y enriquecido a la optimización para

buscadores o SEO (*Search Engine Optimization*). Al respecto, la investigación bibliotecológica especializada es escasa y, todavía más, tratándose de los contenidos audiovisuales (Arlitsch, OBrien y Rossmann, 2013).

Curran (2004) señaló que el SEO influye positivamente en la publicación web a través de la manipulación de elementos utilizados por los buscadores para determinar su relevancia, y propuso la realización de una extensa investigación sobre palabras y términos clave para determinar cuáles son usados en títulos y descripciones. Luego, Curran (2004) y Raisinghani (2005) destacaron que existen oportunidades de mejora para aportar mayor puntaje a la evaluación de los sitios web; por ejemplo, el uso de palabras clave en el código. Tanto Curran como Raisinghani convergen con las recomendaciones de Zhang y Dimitroff (2005), quienes identificaron que al optimizar las etiquetas semánticas se alcanzan posiciones más relevantes en la búsqueda. Descubrieron que este fenómeno se repetía en sitios donde las páginas clave o descriptores temáticos aparecían en las etiquetas de título y en el cuerpo del texto.

Por otro lado, Sen (2005) desarrolló un modelo teórico para comparar los efectos del SEO con las estrategias de posicionamiento pagado *versus* el tráfico orgánico o natural de audiencia y Zing y Lin (2006: 519-529) reconocieron que, aunque el SEO estaba resultando de gran interés en la industria publicitaria, no ocurría lo mismo en el ámbito de la investigación científica y académica; menos en la industria televisiva. Más tarde, Malaga (2007: 68-82) estructuró una concepción sobre la planificación del SEO para incrementar la relevancia y la visibilidad.

El impacto del SEO ha sido tal que, en 2009, la *British Broadcasting Company* (BBC) desarrolló una posición laboral de especialista en SEO Periodismo y, en 2011, *Los Angeles Times* instituyó al *SEO Chief* (Ellis, 2011). Por su parte, para Stabiner (2011) y Bakker (2012), el SEO se puede comprender como un proceso que sirve a los editores web para crear y ajustar los encabezados o los textos de las noticias a fin de mejorar su posicionamiento.

Spyridou y Veglis (2012) aportan que las organizaciones periodísticas, entre otras actividades, redefinen y modifican sus procesos de producción alrededor del SEO para mejorar posicionamiento y tráfico, asunto que luego validaron Giomelakis y Veglis (2016) al concluir que existe una clara correlación entre resultados y optimización. Gonzalo, Codina y Rovira (2015) añaden que las tendencias se enfocan en dos aspectos relacionados con los usuarios: a) perfil del usuario y b) la intención temática de la búsqueda.

El campo de contribución de la bibliotecología y los estudios de la información en el enriquecimiento de la representación temática de los contenidos en la web tiene claras oportunidades en lo que se refiere a la investigación, la elaboración de diagnósticos y el diseño de intervenciones para la

mejora continua y la innovación. Texto, audio, video y animaciones conforman un discurso que, por su formato, demanda de un análisis documental especializado donde se consideren aspectos específicos como el visionado, el análisis del contenido, la elaboración del resumen, los descriptores temáticos con perspectiva semántica para el SEO, así como el etiquetado. En el caso de los noticieros televisivos en la web es necesario realizar un esfuerzo adicional, pues si bien existe un cuerpo de texto que acompaña al recurso multimedia, en realidad la mayor parte del contenido enriquecido se encuentra justo en dicho elemento. De ahí la utilidad de un modelo de representación y recuperación temática (Engerer, 2017: 661).

En Noruega, un estudio realizado en una televisora muestra las dificultades en la indización de contenidos televisivos de *Norwegian Broadcasting Corporation*. Los principales hallazgos revelan una indización incompleta, inconsistente y desigual (Sobak y Pharo, 2017: 739). En México, poco se conoce sobre estos y otros aspectos especializados en el discurso de los noticieros televisivos en la web. Por ello, se tomó como punto de partida el diagnóstico sobre el uso de métodos, técnicas o herramientas de representación temática en televisoras. Posteriormente se diseñó un modelo basado en experiencias de indización, con perspectiva bibliotecológica, que se probó en un entorno de negocio real para observar el efecto en el tráfico de usuarios.

METODOLOGÍA

Se acudió al método conocido como Investigación-Acción (*Action Research*) que Comfort (1985) y Kemmis, McTaggart y Nixon (2014) han estudiado, analizado y sistematizado. Su perspectiva fue tomada como punto de partida para construir la metodología, considerando las cinco etapas que guiaron el presente estudio. Los autores mencionados retoman el desarrollo atribuido a Lewin (1948) en la psicología social, seguido entre otros por Checkland (1981) y Checkland y Holwell (1998), cuyas principales contribuciones facilitaron la revisión del método con el rigor apropiado en la aplicación de la teoría de los sistemas de información y comunicación. Baskerville y Wood-Harper (1996: 235-246) describen la Investigación-Acción como un acercamiento científico basado en diagnósticos, intervenciones y generación de aprendizajes para construir un modelo cíclico que consta de cinco fases: diagnóstico, modelo metodológico, diseño de la intervención para probar el modelo, la implementación y evaluación correspondientes, así como la formulación de aprendizajes.

DIAGNÓSTICO

Para realizar el diagnóstico se establecieron criterios de análisis de la publicación de los noticieros en la web, considerando la interfaz del usuario y el código fuente de contenidos. El diagnóstico se realizó entre febrero y abril de 2015 en 14 televisoras mexicanas: Aprende Televisión Educativa, CNN en español, Canal Once, Efekto TV, Excélsior TV, Foro TV/Televisa, Fuerza Informativa Azteca, Milenio Noticias, MVS Noticias, Noticieros Televisa, Proyecto 40, Telefórmula, Televisión Metropolitana Canal 22 y TV UNAM.

Análisis de la interfaz de usuario. Se revisaron y analizaron las páginas web para identificar presencia o ausencia de metadatos de representación temática para el posicionamiento, la visibilidad y la recuperabilidad: título del noticiero, fecha y horario del noticiero, fecha y hora de posteo, conductores, síntesis del noticiero, palabras clave, nombres propios, temas relacionados, extensión de contenido en redes sociales, el noticiero en versión completa está disponible en vivo o bajo demanda, publicación de guion, escaleta, transcripciones, traducciones o subtítulos y publicación de fragmentos del noticiero bajo demanda.

Hallazgos del diagnóstico en la interfaz de usuario. Las televisoras cuentan con gestores de contenido preparados para asignar, de manera automática, título del noticiero, fecha del noticiero, horario de transmisión, fecha y hora de posteo y conductores. El 34 % no incluye una síntesis que exprese el tema, 86 % no contiene palabras clave o nombres propios como etiquetas complementarias, 79 % no agrega etiquetas sobre temas relacionados, 14 % no añade herramientas para compartir el contenido en redes sociales y 21 % no ofrece los noticieros en vivo o bajo demanda. En el 100 % está ausente la publicación de guiones o escaletas o transcripciones. Sólo un 7 % no ofrece fragmentos de los noticieros.

Análisis del código fuente. Por cada televisora o entidad productora se llevó a cabo la inspección del código fuente de las publicaciones de los noticieros disponibles, en vivo o bajo demanda, así como en cinco clips de video. Se tomaron en consideración los siguientes rasgos: densidad de palabras clave, etiqueta TITLE, etiquetas meta, etiquetas H1 y H2, análisis de palabras clave de la competencia, selección de palabras clave por estacionalidad, variantes idiomáticas y concatenaciones, etiquetado controlado y social, optimización de imágenes, metadescripciones, uso de microdatos y URL.

Hallazgos del diagnóstico del código fuente. Sólo 14 % presentó textos donde la densidad de las palabras clave superó el 15 %, colocándolas dentro del rango. Sin embargo, ninguna presentó una densidad mayor al 30 %. Es decir, los editores web publican “resúmenes periodísticos” pero carecen de

elementos semánticos suficientes para establecer el tema. En 71 % el código fuente no contiene palabras clave en la etiqueta <TITLE>, 57 % no suma enunciados que describan el tema y, en el 43 % restante, incorporan textos periodísticos con baja densidad de palabras clave (no obstante que sí lo hacen). El 100 % de las televisoras utilizan etiquetas <H1>, es decir, que colocan al menos un título en la página web, pero no siempre incluye alguna palabra clave temática. Por otra parte, ninguna utiliza las etiquetas <H2> para elaborar epígrafes, subtítulos o subsecciones en el cuerpo de los resúmenes periodísticos. Asimismo, no aprovechan la oportunidad de estudiar las palabras clave empleadas por sus competidores. En todos los casos se identificó la ausencia del uso de palabras clave seleccionadas por temporalidad; se agregan temas populares (*trending topics*) aunque este recurso no es capitalizado. También hay ausencia en el uso de variantes idiomáticas y concatenaciones.

Un 57 % de las televisoras ejecuta el proceso de asignar etiquetas; no obstante, se deduce la carencia de una estrategia clara para hacerlo, con base en la creación de listas, repertorios, vocabularios o taxonomías. Con un trabajo más estructurado destacan Efekto TV y Televisa que emplean palabras clave con un sentido más cercano a la indización temática. Sólo en el 21 % de los contenidos se encontró que las imágenes fijas o la primera imagen del reproductor de video no están optimizadas. En el 79 % de las televisoras no se identificó el uso de metadescripciones (uso de etiqueta <META = DESCRIPTION>) para mejorar la descripción temática del contenido, a través de un resumen estructurado. Sólo aparece presente con un mayor grado de consistencia en Noticieros Televisa, Foro TV y Efekto TV. Este último caso sobresale porque utiliza una etiqueta específica: <META = ABSTRACT>. En el 100 % de las televisoras se identificó ausencia de microdatos mostrados en los *snippets* de las páginas de resultados arrojadas por motores de búsqueda. Por último, en el 71 % de las televisoras (10 casos) no se registran palabras clave en las URL.

PROPUESTA DE MODELO METODOLÓGICO

Una vez realizado el diagnóstico, en coincidencia con los trabajos de Serrano-Cobos (2015: 11), Malaga (2007: 72), Giomelakis y Veglis (2016, 22-33), así como Codina *et al.* (2016: 7), se presenta la propuesta de un modelo de representación temática de los noticieros en la web (*Figura 1*), tomando como base los factores esenciales para incrementar el valor de la optimización para buscadores (SEO).

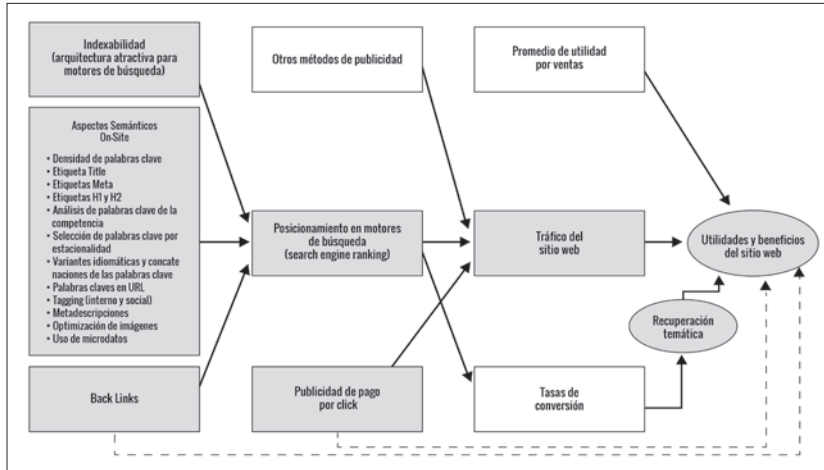


Figura 1. Modelo de representación temática del noticiero televisivo
 Con base en la aplicación de análisis y tratamiento documental del video, la recuperación de otros insumos de la creación del noticiero y aspectos semánticos on-site relevantes para SEO
 Fuente: Soto-Hernández (2017: 94) con base en Malaga (2007), Soto-Hernández (2009), Soto-Hernández y Naumis-Peña (2014) y Serrano-Cobos (2015)

El modelo es una aportación planteada desde la bibliotecología y los estudios de la información hacia el sector periodístico en la web. Tiene como objetivos los siguientes:

- Aplicar un modelo de análisis y tratamiento documental del video para refinar la representación temática (Soto-Hernández, 2009; Soto-Hernández y Naumis-Peña, 2014).
- Incrementar la relevancia de los contenidos digitales para los motores de búsqueda; su visibilidad, posicionamiento y acceso.
- Aprovechar el valor semántico en la web (Serrano-Cobos, 2015: 11).
- Facilitar la representación y recuperación temática mejorando la experiencia del usuario.
- Contribuir a la productividad.

En el modelo se incorporan los procedimientos técnicos del análisis y tratamiento documental del video: visionado, análisis de contenido, resumen, descripción temática y etiquetado. Cada uno de los aspectos semánticos más relevantes presentes en el modelo fueron explicitados en detalle y se presenta aquí la densidad de palabras clave como ejemplo.

- Representa el número de repeticiones en el cuerpo de un texto (no del total del sitio web).
- A través de la densidad de palabras clave, los motores de búsqueda determinan cuál es el tema principal o *aboutness* del contenido (Yablo, 2014: 1). Se recomienda utilizar todas las palabras de un mismo campo semántico incluyendo sinónimos, antónimos, variantes y formas verbales. De este modo aumenta la densidad, sin necesidad de incrementar las repeticiones.
- Es importante equilibrar la densidad de las palabras clave, así como su distribución en todos los espacios donde se puede aumentar la relevancia. De otro modo, el abuso puede ser interpretado por los motores de búsqueda como una mala práctica (*spam*), e incluso, ganarse una penalización en la evaluación de relevancia total del SEO.

De acuerdo con Soto-Hernández (2009: 101), el tratamiento documental del video engloba procedimientos y operaciones que son clave en la representación y recuperación temática de contenidos audiovisuales. Con este proceso técnico se busca satisfacer las demandas de la distribución, comercialización y visibilidad de los noticieros televisivos a través de un esquema multiplataforma (televisión, *web desktop*, *web mobile*, OTT como Netflix, app o medios sociales).

No obstante, las nuevas condiciones del contexto económico, social, empresarial, laboral y técnico, así como la velocidad de la obsolescencia tecnológica, han provocado la incorporación de profesiones emergentes. El etiquetado que pudieran realizar los lectores de las noticias tampoco es una solución. Por lo tanto, hay consenso en que las etiquetas generadas socialmente tienen problemas de calidad asociados con el uso de palabras no declaradas, poliseimia, sinonimia y falta de jerarquía (Guy y Tonkin, 2006; Matusiak, 2006; Lu, Park y Hu, 2010), y la falta de distinción sobre a qué tipo corresponde una etiqueta (Springer *et al.*, 2008: 18). En el caso aún de la indización de imágenes fijas, los problemas existentes para indizar textos se multiplican (Matusiak, 2006: 294), debido a la riqueza semántica y la ambigüedad de la imagen (Melgar *et al.*, 2017: 350).

Cuando las empresas televisoras decidieron difundir contenidos en la web, surgió la necesidad de incorporar otro perfil profesional conocido como editor web cuyos objetivos y competencias son una obra en progreso. Éste es responsable de asegurar la ejecución de tareas que van desde la preparación de los materiales hasta otras vinculadas con el análisis de contenido.

El otro perfil profesional de reciente incorporación en las televisoras que publican noticieros en la web es el *SEO manager*, a cargo de supervisar

que cada contenido digital, como un sitio web, cuente con elementos que los motores de búsqueda exigen a través de sus algoritmos para alcanzar la mejor posición posible en las páginas de resultados.

Cabe preguntarse cuál es el lugar del bibliotecario, el profesional de la información, en este panorama (Arlitsch, OBrien y Rossmann, 2013: 177-188), pues de las diferentes operaciones que deben realizar el editor web y el *SEO manager* hay algunas que, explícitamente, corresponden al ámbito de los procedimientos del campo bibliotecológico. La ejecución guarda apenas una distancia sutil pero relevante (Soto-Hernández, 2017: 101).

Para lograr que la representación temática se genere de forma efectiva, se proponen operaciones que proceden de la tradición bibliotecológica y de los estudios de la información, para los fines de recuperación, incremento de visibilidad, mejora en el posicionamiento en motores de búsqueda, experiencia de usuario y aumento en los indicadores de tráfico: resumen (no periodístico), descripción temática a través de palabras claves y etiquetado.

Las palabras clave son producto del análisis del contenido del video periodístico y extraídas de sus elementos textuales, expresados en el discurso del noticiero o del guion de una nota, reportaje, entrevista, etcétera. A diferencia de la indización especializada que se conoce en la bibliotecología, tratándose de documentos científicos, no resulta tan sencilla la extracción de palabras clave desde el título del noticiero (v. g. “CNN con Carmen Aristegui”) o de una noticia (v. g. “SSPCDMX aplicará operativo vial en Viaducto Tlalpan por obras”). Como lo señala Naumis-Peña (2007: 105), lo ideal es que sean representativas de los contenidos documentales y, en este sentido, cuando la extracción e indización se puede realizar por asignación y con participación humana, resultará más significativa.

En un artículo reciente se argumenta que las técnicas de inteligencia artificial no son lo suficientemente maduras para resolver el problema de la indización multimedia conceptualmente y no serán capaces de reemplazar a los indizadores humanos en el futuro previsible (MacFarlane, 2016: 180). El principal destinatario de las palabras clave es el motor de búsqueda y se persigue que haya la mayor congruencia entre el significado que están representando y la infinita gama de posibilidades cognitivas en una consulta efectuada por un usuario.

DISEÑO DE LA INTERVENCIÓN PARA APLICAR EL MODELO

Tomando como base los datos arrojados por el diagnóstico y el modelo de representación y recuperación propuesto en el apartado anterior, se estableció

contacto con las 14 televisoras analizadas. A todas se les propuso un programa de intervención. Finalmente, sólo una de las empresas aceptó que se ejecutara el proyecto; se reserva el nombre o marca de la organización por acuerdo de confidencialidad. Enseguida se presentan el caso y las acciones ejecutadas.

Reto

- Contrarrestar el efecto negativo en los principales indicadores de tráfico web que en febrero de 2015 disminuyeron 45 % (la aparición de Netflix y otro de tipo de programación fija) con respecto al año inmediato anterior, en la plataforma web y *web mobile* (web en dispositivos móviles).

Objetivos

- Diseñar una estrategia organizacional de tipo editorial y operativa para la representación temática.
- Implementar la estrategia de forma experimental en un ambiente de pruebas.
- Evaluar el impacto.
- Documentar aprendizajes.

La implementación. Etapas

- Establecimiento de un comité de trabajo.
- Diseño de la estrategia e indicadores.
- Implementación en un ambiente de pruebas.
- Análisis de datos.
- Formulación de aprendizajes.

Acciones

- Creación de comité de trabajo / Mayo, 2015.
- Diseño de la estrategia / Mayo, 2015.
- Análisis / Junio, 2015.
- Establecimiento de indicadores / Junio, 2015.
- Implementación en ambiente de pruebas / Julio-diciembre, 2015.
- Análisis de datos / Enero-febrero, 2016.
- Formulación de aprendizajes / Marzo, 2016.

Resultados esperados

- Frenar la caída del tráfico web y, de ser posible, incrementar los indicadores de desempeño.

- Indicadores observados: usuarios únicos absolutos, páginas vistas, videos vistos, tiempo de permanencia, perfiles demográficos.
- La ejecución del experimento se realizó en cinco noticieros que figuran dentro del portal web de la empresa. La intervención se aplicó en dos noticieros mientras que los otros tres sirvieron como medio de control, y se llevó a cabo de septiembre a noviembre de 2015.

EVALUACIÓN DE LOS RESULTADOS

Al concluir la intervención se procedió a la recolección, acopio y análisis de los datos relacionados con los indicadores clave de tráfico web establecidos. Los resultados se describen en la serie de figuras que aparecen a continuación. Como se aprecia en la *Figura 2*, al repunte de usuarios únicos absolutos que se presentó en febrero de 2015 siguió una caída constante que llegó a su nivel más bajo en julio, coincidiendo con el auge de Netflix y otros servicios que ofrecen programación fija. En febrero inició la etapa de implementación de la nueva estrategia de representación temática. Con las medidas, se logró un pico en la curva hacia octubre. Para noviembre, había comenzado a presentarse un comportamiento más estable con una ligera recuperación en diciembre, aun cuando la temporada navideña suele tener un impacto negativo en los indicadores de tráfico web. No se pagó campaña de impulso.

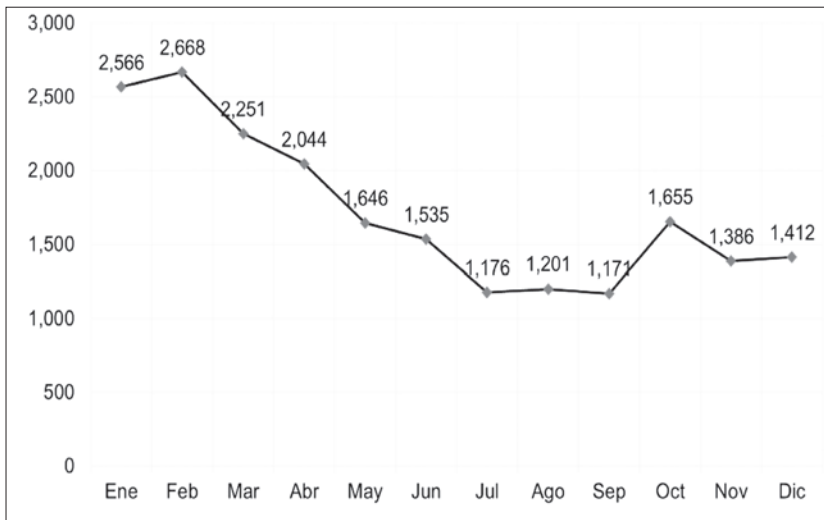


Figura 2. Usuarios únicos absolutos (millares)
Fuente: ComScore MyMetrix, Google Analytics

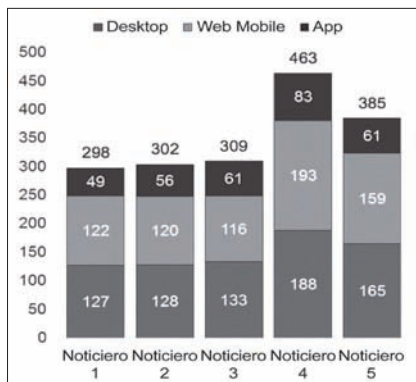


Figura 3. Usuarios únicos absolutos según noticiero (millares)
Fuente: ComScore MyMetrix, Google Analytics

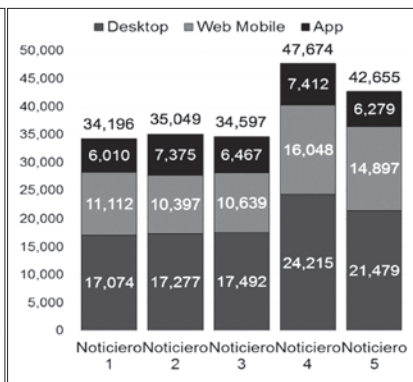


Figura 4. Páginas vistas según noticiero (millares)
Fuente: ComScore MyMetrix, Google Analytics

La Figura 3 expone el comportamiento que tuvieron los cinco noticieros que produce la empresa estudio de caso en el indicador de usuarios únicos absolutos. Sólo en los noticieros 4 y 5 (en los que se aplicó la política de representación) hubo un incremento notable. La diferencia parece sobresaliente en *web desktop* y *web mobile*; esto se debe a que la estrategia no se implementó en la app. En la Figura 4, en correspondencia, el número de páginas vistas también aumentó como resultado del experimento.

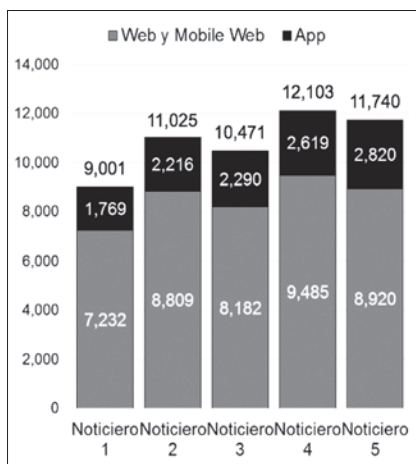


Figura 5. Videos vistos según noticiero (millares)
Fuente: ComScore MyMetrix, Google Analytics

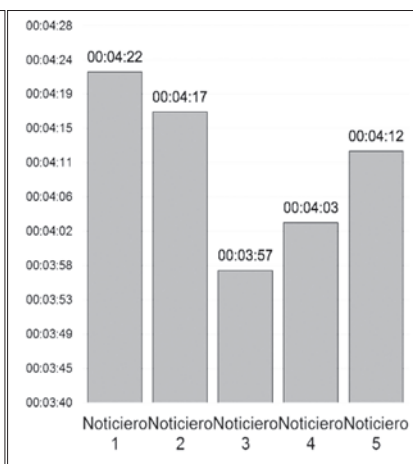


Figura 6. Tiempo de permanencia según noticiero (minutos/segundos)
Fuente: ComScore MyMetrix, Google Analytics

La *Figura 5* expresa la distribución del consumo de video según plataforma. Hubo crecimiento en los noticieros 4 y 5 por encima de los demás. En la *Figura 6* se observa el tiempo de permanencia realizando una actividad (*clicks*, recorrer la página, activar un reproductor de video); el incremento estuvo por arriba de la media comercial que ronda el 1.5 minutos en los noticieros donde se implementó la política de representación temática.

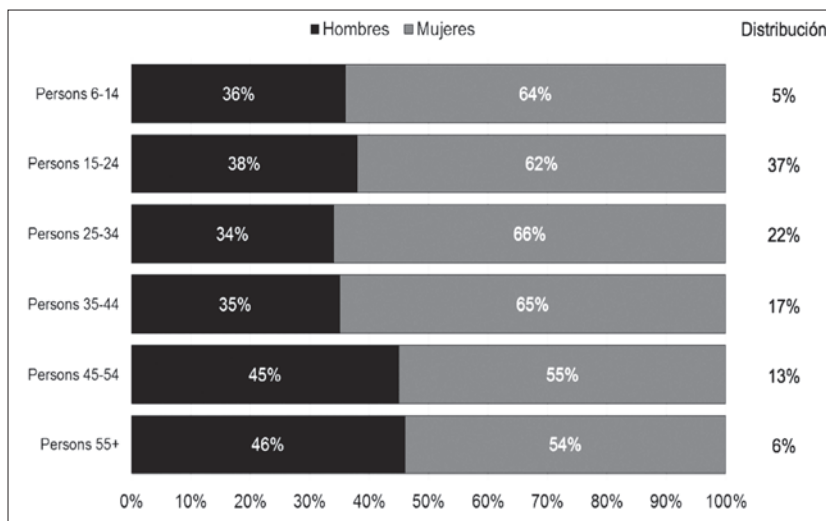


Figura 7. Perfiles demográficos generales (%)

Fuente: ComScore MyMetrix, Google Analytics

En la *Figura 7* se presenta el perfil demográfico de los usuarios. Se incorporó a niños y adultos mayores, lo que significa se alcanzó al llamado *target* comercial 4+ o familiar con el que, por cierto, antes no contaban. Esto aplica para los noticieros 4 y 5 en los que se realizó la intervención.

Formulación de aprendizajes

En la empresa estudio de caso se identificaron grandes dificultades para encontrar editores web con formación específica, profesionalizados y disponibles para llevar a cabo las tareas periodísticas y las de representación temática con rigor y efectividad. Ninguno de los editores web participantes en el proceso cuenta con algún tipo de entrenamiento especializado en biblioteología o gestión de información, o bien, en posicionamiento en motores de búsqueda desde una perspectiva semántica.

CONCLUSIONES Y RECOMENDACIONES

La televisión y la web están inmersas en la convergencia digital en el que se generan diversidad de expresiones creativas y donde se combinan estructuras, géneros, lenguajes y concepciones artísticas y periodísticas para satisfacer las necesidades de los usuarios, los anunciantes y el modelo de negocio.

Al probar el modelo propuesto para la representación de los noticieros en un entorno de negocio real, incluso en condiciones de estrechez presupuestal, se identificó el potencial que puede tener la contribución de la bibliotecología y los estudios de la información. Se pudo constatar que la metodología propuesta fue efectiva en un periodo breve de implementación.

Entre los hallazgos hay dos muy relevantes. El primero de ellos es que las prácticas de optimización para buscadores (SEO) se complementan y enriquecen con la representación temática resultado del tratamiento documental. El valor de la publicación digital en la que se difunden los noticieros televisivos incrementa cuanto más claro es el tema para los motores de búsqueda. Debería ser un objetivo el dotar de mayor recuperabilidad a la publicación web para desempeñar mejor sus funciones de comunicación social y, en su caso, los retos y metas comerciales de negocio que hacen posible la producción y distribución. Hay aquí una amalgama que se convierte en una vía de investigación, innovación y desarrollo para la disciplina bibliotecológica y los estudios de la información.

El segundo hallazgo está relacionado con las nuevas competencias profesionales que se requieren para gestionar el tratamiento documental como una innovación aplicada al sector industrial de la información en el entorno web. Hay demanda de recursos humanos preparados para realizar operaciones transaccionales, analíticas y estratégicas. En una era donde la visibilidad cobra un valor comercial, social, político, jurídico, histórico y cultural de enorme relevancia, hay un lugar importante para el bibliotecólogo como profesionista y para el investigador especializado.

Finalmente, el modelo muestra que el análisis documental y la representación temática y su aplicación al SEO del noticiero televisivo en la web requieren que el personal se actualice y profesionalice, traslapando su perfil con el del editor web en lo particular y con el de los profesionales de las industrias creativas en lo general (Mietzner y Kamprath, 2013: 280-294). Es decir, para el recurso humano implica diseñar y estructurar la publicación considerando la relevancia lingüística y semántica para mejorar la representación, recuperación, posicionamiento en los motores de búsqueda y experiencia de usuario.

En tanto que la información sea un elemento constitutivo del mercado y la recuperabilidad facilite el acceso a los usuarios o la recomendación para usarla, habrá necesidad de crear nuevas soluciones para posicionar los documentos y contenidos, con la finalidad de ponerlos en contacto con las personas.

REFERENCIAS

- Arlitsch, Kenning, Patrick OBrien y Brian Rossmann. 2013. "Managing Search Engine Optimization: An Introduction for Library Administrators". *Journal of Library Administration* 53 (2-3): 177-188.
- Bakker, P. 2012. "Aggregation, content farms and huffinization". *Journalism Practice* 6 (5-6): 627-637. doi:10.1080/17512786.2012.667266
- Baskerville, Richard. L. y A. Trevor Wood-Harper. 1996. "A Critical Perspective on Action Research as a Method for Information Systems Research". *Journal of Information Technology* 11: 235-246.
- Checkland, Peter y Sue Holwell. 1998. *Information, Systems and Information Systems Making Sense of the Field*. England: John Wiley and Sons.
- Checkland, Peter. 1981. *Systems Thinking, Systems Practice*, Chichester. England: John Wiley and Sons.
- Codina, Lluís, Mar Iglesias-García, Rafael Pedraza y Lucía García-Carretero. 2016. *Visibilidad y posicionamiento Web de informaciones periodísticas: El Framework SEO-RCP*. Barcelona: Serie Editorial DigiDoc-UPF. eRepositorio Universitat Pompeu Fabra. <https://www.upf.edu/web/digidoc/editorial-digidoc>
- Comfort, Louise K. 1985. "Action research: a model for organizational learning". *Journal of Policy Analysis and Management* 5 (1): 100-118.
- Curran, K. 2004. "Tips for Achieving High Positioning in the Results Pages of the Major Search Engines". *Information Technology Journal* 3 (2): 202-205. <https://doi.org/10.1108/10650750610706998>
- Ellis, Justin. 2011. "Traffic Report: Why pageviews and engagement are up at Latimes.com". *Nieman Journalism Lab*, August 16. <http://www.niemanlab.org/2011/08/traffic-report-why-pageviews-and-engagement-are-up-at-latimes-com/>
- Engerer, Volkmar. 2017. "Exploring Interdisciplinary Relationship Between Linguistics and Information Retrieval From the 1960s to Today". *Journal of the Association for Information Science and Technology* 68 (3): 660-680. <https://doi.org/10.1002/asi>
- Giomelakis, Dimitrios y Andreas Veglis. 2015. "Employing Search Engine Optimization Techniques in Online News". *Studies in Media and Communication* 3 (1): 22-33.
- Giomelakis, D. y Andreas Veglis. 2016. "Investigating Search Engine Optimization Factors in Media Websites". *Digital Journalism* 4 (3): 379-400. doi: 10.1080/21670811.2015.1046992
- Gonzalo, C., L. Codina y C. Rovira. 2015. "Recuperación de información centrada en el usuario SEO: categorización y determinación de las intenciones de búsqueda en la Web". *index.comunicación* 5 (3): 19-27.
- Guy, Marieke y EmmaTonkin. 2006. "Folksonomies: Tidying up tags?". *D-Lib Magazine* 12 (1). <https://doi.org/10.1045/january2006-guy>

- Hjørland, Birger. 2017. "Reviews of Concepts. Knowledge Organization". *Knowledge Organization* 44 (1): 55-64.
- Kemmis, Stephen, Robin McTaggart y Rondha Nixon. 2014. *The action research planner: doing critical participatory action research*. Singapore: Springer. <https://doi.org/0.1007/978-4560-67-2>
- Lewin, Kurt. 1948. *Resolving Social Conflicts*. New York: Harper.
- Lu, Cai mei, Jung-Ran Park y Xiaohua Hu. 2010. "User tags versus expert-assigned subject terms: A comparison of LibraryThing tags and library of congress subject headings". *Journal of Information Science* 36 (6): 763-779. <https://doi.org/10.1177/0165551510386173>
- MacFarlane, Andrew. 2016. "Knowledge Organization and its Role in Multimedia Information Retrieval". *Knowledge Organization* 43 (3): 180-183.
- Malaga, Ross A. 2007. "The Value of Search Engine Optimization: An Action Research Project at a New E-Commerce Site". *Journal of Electronic Commerce in Organizations* 5 (3): 68-82. <https://doi.org/10.4018/jeco.2007070105>
- Matusiak, Krystyna K. 2006. "Towards user-centered indexing in digital image collections. OCLC". *Systems & Services* 22 (4): 283-298.
- Melgar Estrada, Liliana, Michiel Hildebrand, Victor de Boer y Jacco van Ossenbruggen. 2017. "Time-Based Tags for Fiction Movies: Comparing Experts to Novices Using a Video Labeling Game". *Journal of the Association for Information Science and Technology* 68 (2): 348-364.
- Mietzner, Dana y Martin Kamprath. 2013. "A Competence Portfolio for Professionals in the Creative Industries". *Creativity & Innovation Management* 22 (3): 280-294.
- Naumis-Peña, Catalina. 2007. *Los tesauros documentales y su aplicación en la información impresa, digital y multimedia*. México, Buenos Aires: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, Library Outsourcing Service, Alfagrama.
- Raisinghani, Mahesh S. 2005. Future Trends in Search Engines". *Journal of Electronic Commerce in Organizations* 3: i-vii.
- Sen, R. 2005. "Optimal Search Engine Marketing Strategy". *International Journal of Electronic Commerce* 10 (1): 9-25.
- Serrano-Cobos, Jorge. 2015. *SEO. Introducción a la disciplina del posicionamiento en buscadores*. Barcelona: Editorial UOC.
- Sobak, Veslemoy y Nils Pharo. 2017 "Decentralized Subject Indexing of Television. Programs: The Effects of Using a Semicontrolled Indexing Language". *Journal of the Association for Information Science and Technology* 68 (3): 739-749. <https://doi.org/10.1002/asi>
- Soto-Hernández, Silvano y Catalina Naumis-Peña. 2014. "Análisis bibliotecológico de los noticieros televisivos mexicanos en la web". *El Profesional de la Información* 23 (1): 80-86. <https://doi.org/10.3145/epi.2014.ene.10>
- Soto-Hernández, Silvano. 2009. "Tratamiento documental del video: Propuesta metodológica". Tesis de maestría, Universidad Nacional Autónoma de México, Facultad de Filosofía y Letras.
- Soto-Hernández, Silvano. 2017. "Tratamiento documental: Representación y recuperación de los noticieros en la web". Tesis de doctorado, Universidad Nacional Autónoma de México, Facultad de Filosofía y Letras.
- Springer, Michelle, Beth Dulabahn, Phil Michel, Barbara Natanson, David Reser, David Woodward y Helena Zinkham. 2008. *For the common good: The Library of*

- Congress Flickr pilot project*. DC: Government of the United States; Library of Congress. http://loc.gov/rr/print/flickr_report_final.pdf
- Spyridou, P. y A. Veglis. 2012. "Sustainable Online News Projects: Redefining Production Norms and Practices". *Proceedings of the 10th World Media Economics and Management Conference*, Thessaloniki, May.
- Stabiner, K. 2011. "CJR Column Mentions The Simpsons. A second look at SEO". *Columbia Journalism Review* (March/April).
- Yablo, Stephen. 2014. *Aboutness*. Princeton, NJ: Oxford: Princeton University Press.
- Zhang, Jin y Alexandra Dimitroff. 2005. "The Impact of Metadata Implementation on Webpage Visibility in Search Engine Results", Part II. *Information Processing and Management* 41 (3): 691-715.
- Zing, Bo y Zhangxi Lin. 2006. "The Impact of Search Engine Optimization on Online Advertising Market". *Proceedings of the 8th International Conference on Electronic Commerce: The New W-Commerce: Innovations for Conquering Current Barriers, Obstacles and Limitations To Conducting Successful Business on the Internet*. New Brunswick, Canadá. <https://doi.org/10.1145/1151454.1151531>

Para citar este texto:

- Naumis Peña, Catalina y Silvano Soto Hernández. 2018. "Representación temática y recuperación de los noticieros de televisión en la web: propuesta y aplicación de un modelo metodológico". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 175-191. <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57868>

DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57868>

Mujeres investigadoras en las primeras estructuras de organización en ciencias exactas e ingenierías en México de 1900-2000: Estudio Bibliométrico

María Elena Luna Morales*
Evelia Luna Morales*

Artículo recibido:
4 de enero de 2018
Artículo aceptado:
19 de septiembre de 2018
Artículo de investigación

RESUMEN

Se analiza la producción científica de México registrada en Web of Science de 1900 a 2000 en las áreas de ciencias exactas e ingenierías para caracterizar la forma en que se incorporan las mujeres investigadoras al desarrollo de estos campos de estudio. El trabajo se apoya en el método bibliométrico cuantitativo a través de la aplicación de indicadores de género, producción e impacto científico y el análisis de redes bibliométricas de coautoría. En México, las ciencias exactas e ingenierías registran sus primeras aportaciones durante la primera mitad del siglo XX. Las mujeres lo hicieron a partir de los años 40 a través de laboratorios Syntex donde publicaron sus contribuciones iniciales. La

* Coordinación General de Servicios Bibliográficos, Cinvestav-IPN, México
meluna@cinvestav.mx
eluna@cinvestav.mx

participación de las investigadoras se incrementó entre 1960 y 1970 con el reconocimiento de las mujeres en la educación superior y la investigación, así como la creación de distintas instituciones de investigación en el país, además de los cambios de las mujeres con respecto a la elección de carrera. Finalmente, las contribuciones de este trabajo representan una ventaja para la propia comunidad científica, pero también para las organizaciones orientadas a promover la equidad de género y las instituciones que integran investigadoras en este campo de estudio.

Palabras clave: Ciencias Exactas; Ingenierías; México; Estudio Bibliométrico de Género; Investigación Científica

A Bibliometric study of women researchers in early organizational structures of exact sciences and engineering in Mexico from 1900 to 2000

María Elena Luna-Morales and Evelia Luna-Morales

ABSTRACT

The scientific production of Mexico registered in Web of Science from 1900-2000 in the areas of Exact Sciences and Engineering is analyzed to describe the way in which female researchers are integrated in the development of these fields of study. The paper employs a quantitative bibliometric method, applying indicators of gender, scientific output and impact, and analysis of bibliometric networks of co-authorship. In Mexico, researchers in exact sciences and engineering published their first papers in the first half of the 20th century, but female researchers working in Syntax laboratories did not publish until the 1940s. As more women joined faculties in higher education and research centers in the 60s and 70s, publication by female researchers began to increase accordingly. This growth also appears to be associated with changes in career choice by women. While the contributions of this work represent an advantage for the scientific community itself, they also are useful for organizations that promote gender equity and institutions that integrate female researchers in this field of study.

Keywords: Exact Sciences; Engineering; Mexico; Bibliometrics, Gender Study; Scientific Research

INTRODUCCIÓN

La participación de las mujeres en las ciencias sigue siendo un tema de debate en todos los ámbitos (Buquet-Corleto, 2013; Pollack, 2013). Esta situación ha llamado la atención de distintos autores y la participación de algunas instituciones públicas y privadas que se han interesado por conocer los problemas a los que se enfrentan las investigadoras, como la formación académica, el progreso en la actividad científica y la ocupación de puestos altos, entre otros. Esto se ha reflejado en estudios como el presentado por The Boston Consulting Group (2013) y los producidos por la Organización de las Naciones Unidas (ONU), a través de su organismo especializado en educación, ciencia y cultura (Unesco, 2010).

Estos estudios se complementan con otros trabajos publicados, en los que se da a conocer el antagonismo y la exclusión hacia las mujeres (Sugimoto, 2013), así como la escasa explotación del potencial del género femenino en la actividad científica y tecnológica y la forma en que esto repercute en el desarrollo económico de los países (Mendieta-Ramírez, 2015).

Asimismo, se pueden mencionar distintas propuestas que giran en torno a la implementación de programas de posgrado más atractivos –que llamen la atención de las mujeres durante la elección de carreras– y la modificación de los planes de estudio desde los niveles educativos más bajos, para ir preparando y motivando a las niñas desde pequeñas (Miroux, 2011; Guevara-Ruiseñor y García-López, 2010).

El tema se complica cuando se analiza la participación de las investigadoras en las ciencias exactas e ingenierías, áreas que por mucho tiempo se han considerado del dominio masculino (Guevara-Ruiseñor y García-López, 2010). En estos campos, las mujeres han prosperado de manera impresionante, pero no lo suficiente como para lograr una mayor presencia. A nivel mundial, en física, las mujeres tienen representación de 10 % del universo y en ingenierías, de 30 % (Unesco, 2017). Según el Sistema Nacional de Investigadores (SNI) de México, entre 2013 y 2015 las mujeres no fueron mayoría en ninguno de los campos que conforman las ciencias exactas e ingenierías.

Didou y Etienne (2011) y Meza-Montes (2005) señalan, en físico-matemáticas y ciencias de la Tierra, que las mujeres representan 18 % del total, en ingenierías, 19 % y en biología y química, 40 %. Esta escasa representación ha orientado a las propias mujeres a realizar estudios más reflexivos que permitan identificar la forma en que funcionan los regímenes de género existentes para identificar los factores que las limitan (Narváez, 2016).

Pollack (2013) infiere que la razón por la que aún hay muy pocas mujeres en la ciencia no tiene que ver con la falta de talento, sino por la desestimación

que las mismas se tienen por los constantes bloqueos de los que son víctima: escasas contrataciones, sueldos bajos, exiguos reconocimientos, poco liderazgo en proyectos y puestos de trabajo, entre otros aspectos (Rossiter, 1982). A lo anterior se suman factores culturales y psicológicos, medio por el que se caracteriza a la mujer como carente de iniciativa (Pollack, 2013). A estos agentes se unen otros de orden político, económico y legal que, en conjunto, determinan el rol de las mujeres en la ciencia y la investigación, sobre todo, cuando se trata de elegir carrera, donde la familia juega un papel fundamental (Guevara-Ruiseñor y García-López, 2010).

La bibliometría en la evaluación de la ciencia

Los estudios bibliométricos surgieron desde hace más de un siglo. Sin embargo, fue Pritchard (1969) quien dio a conocer el término *bibliometría* por primera vez en un trabajo publicado en *Journal of Documentation*. Los estudios de la ciencia con enfoque bibliométrico se incrementaron a partir de la década de los años 70 del siglo XX. El aumento en el número de publicaciones y el exceso de información dio lugar a nuevos estudios orientados a la identificación de patrones de publicación y comunicación científica.

Los estudios métricos se apoyan en distintas teorías y modelos matemáticos con los que se pueden concretar objetivos relacionados con el comportamiento de la información (Gorbea-Portal, 2013). El método más utilizado es el cuantitativo que, en combinación con la estadística, permite cuantificar los datos a través de la aplicación de indicadores bibliométricos (Velasco *et al.*, 2012).

Los avances computacionales y la necesidad de visualizar los resultados de la producción científica desde una perspectiva distinta han conducido a que la bibliometría se relacione con el análisis de redes sociales como una forma de interpretar los aspectos de relación y cooperación entre los integrantes de un conjunto (Pinto, Moreiro-González y Oliveira de Meira, 2009).

Pajek (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>), software gratuito para Windows, se ha convertido en una herramienta muy práctica para el análisis y visualización de redes sociales (Batagelj y Mrvar, 2011). El software se apoya en la teoría de grafos: considera los nodos como individuos y las aristas como vínculos y cuenta con un amplio número de métodos que ayudan a determinar patrones de comportamiento en la colaboración científica (Hernández-García, 2014).

El análisis estructural se basa en las subestructuras que conforman una red. Para comprender estas estructuras es importante identificar los componentes que participan en las relaciones: diadas y triadas cuya conexión puede

dar lugar a grandes estructuras a través de la cohesión que ejercen. El tamaño de los nodos lo determina quien construye la red, y por este medio puede representar distintas variables, entre otras, la producción que registran los autores, el impacto científico y el número de documentos en colaboración (Oscar-Lluch, 2010).

Es muy amplio el número de trabajos que abordan el tema de la participación de las mujeres en los distintos ámbitos y campos de investigación, incluidas las ciencias exactas e ingenierías (Contreras-Gómez *et al.*, 2015; Cepeda-Zetter, González-Brambila y Pérez-Angón, 2017; Buquet-Corleto, 2013; Pollack, 2013). El tema se aborda desde diferentes enfoques: político, social, educativo, investigación y económico; por hombres, mujeres y un variado número de instituciones y organismos públicos y privados (Valle-Díaz Muñoz y Garay-Sánchez, 2012; The Boston Consulting Group, 2013; Moya-Anegón *et al.*, 2007).

Sin embargo, en México y en el área de ciencias exactas e ingeniería, no hay estudios que aborden el tema del género desde el punto de vista bibliométrico, en complemento con el análisis de redes de coautoría para determinar el aprovechamiento de capital científico (Bourdieu, 2003). Es por ello que, en este estudio, se busca analizar la producción científica de México registrada en Web of Science (WoS) de 1900 a 2000 en las áreas de ciencias exactas e ingenierías, para caracterizar la forma en que se incorporan las investigadoras al desarrollo de estos campos de estudio y la manera en que contribuyen a la formación de las estructuras de organización. Cabe aclarar que se toma este periodo de estudio porque los años 90 están identificados como los años de crecimiento y consolidación de la ciencia en México (Luna-Morales, 2012), y en este caso se cubriría la etapa de mayor crecimiento para ver los incrementos de mujeres en las ciencias exactas e ingenierías.

WoS registró que, entre 1900 y 2000, hubo 68 415 autores de ciencias exactas e ingenierías. Entre ellos, 4 541 eran mujeres de cualquier parte del mundo y 2 957 están adscritas a instituciones de México.

METODOLOGÍA

A partir de búsquedas en las bases de datos que conforman la colección principal de WoS, se recuperó la producción científica de México de 1900-2000. Las búsquedas se realizaron por *Address* para recuperar lo publicado bajo "MEXICO", excluyendo "NUEVO MEXICO OR NM". Los resultados se limitaron utilizando los refinamientos que incluye la base de datos, por este medio se seleccionaron los años y las áreas de estudio (física, química, matemáticas y las distintas ingenierías).

Los registros se recuperaron de 500 en 500 debido a que WoS establece este límite. Se utilizó Microsoft Excel para conformar un archivo único donde se desagregaron uno a uno algunos campos como dirección de adscripción de los autores (C1), correspondencia de los autores (RP), categorías temáticas (SC) y autores (AU), lo anterior para normalizar y conformar catálogos únicos.

Para generar el catálogo de direcciones de autores se separaron, en un archivo independiente en Excel, los campos de afiliación y correspondencia para desagregar y seccionar cada uno de los apartados que los componen: institución principal, departamentos y unidades, subdependencias, laboratorios, secciones y programas, código postal y entidad federativa y país de origen de los autores. Cada campo se completó y normalizó para originar el índice de direcciones constituido por instituciones, dependencias, estados y países.

Para construir el índice de nombres de autores fue indispensable separarlos uno a uno sin perder la relación con el resto de los campos de la referencia bibliográfica. Esta separación permitió reunir en una sola lista todos los nombres de autores, la cual se organizó en forma alfabética, lo que ayudó a agrupar las variantes de nombres y, con ello, conseguir una normalización más rápida, facilitando la clasificación de los autores por género: mujer (M) y hombre (H).

La clasificación de los autores por género se realizó utilizando distintas fuentes de búsqueda: 1) bases de datos: WoS y Scopus; 2) páginas web de universidades públicas y privadas, centros e institutos de investigación nacionales e internacionales; 3) listas de autores en el SNI; 4) anuarios de instituciones; 5) Google Scholar, y 6) búsquedas libres en internet. Con este método, se identificó el género de 68 415 autores de un total de 72 701, lo que hace inferir que 4 286 no se identificaron.

Los índices contruidos se transfirieron a una base de datos en Access como tablas independientes, como una forma para generar consultas a partir de la relación entre tablas. Para el desarrollo de las redes bibliométricas de coautoría se utilizó Pajek.

Para determinar la participación de las mujeres en las estructuras de investigación en ciencias exactas e ingenierías, se crearon tres matrices que dieron lugar a tres redes de coautoría. La primera cubre de 1900-1970, la segunda de 1971-1975 y la tercera de 1976-1980. En estos casos se aplicó un algoritmo de visualización de datos por separación de componentes. De este modo es más fácil identificar las colaboraciones según el número de autores que conforman la estructura de cada red, así como las relaciones entre los autores que participan y los campos de investigación a la que pertenecen.

Cabe aclarar que, debido al incremento en el número de autores en años posteriores a 1980, se tomó la decisión de no integrar una red por quinquenio como se había planeado originalmente. Sin embargo, para no dejar duda con respecto al incremento de científicas sobre todo en el último periodo del estudio (1990-2000), se desarrolló la red que abarca de 1996 a 2000, donde se muestra únicamente la distribución de autores por género a través del algoritmo de visualización de optimización de filtro solo, que se encarga de encontrar el estado de equilibrio y mostrar la distribución de nodos sin necesidad de estructurar en redes, lo anterior como una forma de evitar la saturación en la visualización de la red. Es importante mencionar que para contribuir a lo anterior y reducir la cifra de autores, se excluyeron los del área de física de partículas y campos integrados a las colaboraciones *Big Science* con más de 200 coautores. También se descartó la participación de los autores que registran entre uno y cuatro trabajos.

PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

Producción científica

La *Figura 1* muestra, por serie anual, los incrementos de la producción. La línea con crecimientos más altos corresponde a la producción global, que incluye todas las áreas y todos los autores sin limitación de género. Las líneas referentes a trabajos-investigadoras y trabajos-investigadoras-México son muy parecidas, excepto las variantes que se muestran al final de los años 70 y segunda mitad de los 80.

Los crecimientos mostrados en los primeros 70 años son producidos por autores (hombres). Se trata de una contribución mínima, debido al bajo número de investigadores que había en esta época. A comienzos de los años 70 se ven cambios en las tres variables analizadas. Los crecimientos se consolidan principalmente durante las siguientes décadas (80 y 90). La producción global muestra un aspecto importante, pues dobla la de las investigadoras en general y la de adscritas a instituciones del país. Aunque hay años en los que las líneas se acercan mucho, en general es muy atípico el crecimiento de esta disciplina.

No obstante, es muy probable que los crecimientos mostrados tengan que ver con los siguientes aspectos: 1) implementación de posgrados de estudio; 2) consolidación de grupos de investigación, y 3) incrementos en la colaboración científica nacional e internacional como las *Big Science* (Luna-Morales y Collazo-Reyes, 2002).

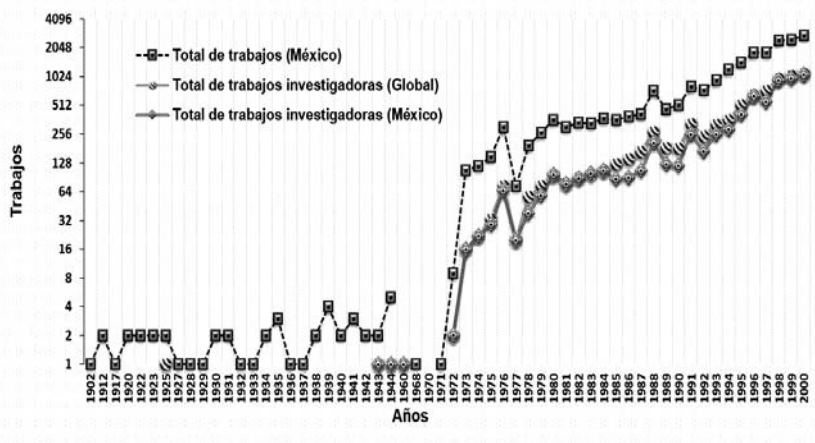


Figura 1. Producción científica anual en ciencias exactas e ingenierías: 1900-2000

La *Figura 2* muestra por año la producción científica que los investigadores (mujeres y hombres), integrados a las ciencias exactas e ingenierías, registran en México. Como se mencionó anteriormente, los años 40 marcan los orígenes en el desarrollo de estos campos de estudio, aunque es básicamente en los 70 cuando se observan crecimientos constantes.

De acuerdo con la *Figura 2*, los investigadores presentan crecimientos mayores a lo largo del periodo de estudio. Las investigadoras hacen lo mismo, pero su producción es dos veces menor que las de los investigadores. Las mujeres alcanzaron los 500 trabajos hasta 1988, rompieron con este tope en 1991 y, en adelante, subieron la producción, pero no lo suficiente para igualar la producción de los investigadores.

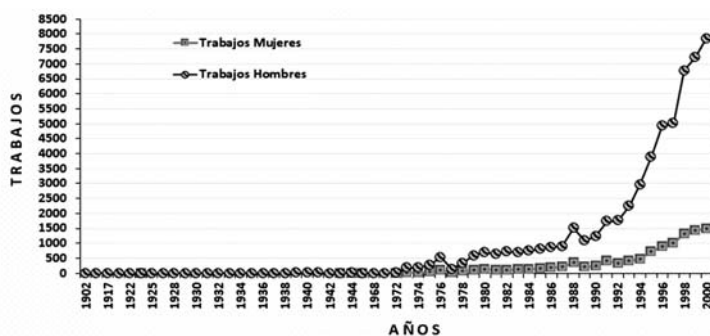


Figura 2. Distribución por serie anual de la producción científica generada por ambos géneros en ciencias exactas e ingenierías: 1900-2000

Mujeres investigadoras por institución

La *Figura 3* presenta la producción científica generada por investigadoras y por institución, así como la presencia que en años tienen las mujeres. Se distinguen cinco grupos de instituciones que son fundamentales para el progreso que logran las investigadoras en el país:

- 1) Representado por la UNAM, que muestra la producción más alta generada y tiene 29 años con presencia femenina;
- 2) Conformado por el Cinvestav y la UAM, ambas instituciones registran crecimientos importantes en producción científica y cuentan con 27 años integrando mujeres en su planta académica y de investigación;
- 3) Constituido por el Instituto Politécnico Nacional (IPN), el Instituto Nacional de Investigaciones Nucleares (ININ), la Benemérita Universidad Autónoma de Puebla (BUAP) y el Instituto Mexicano del Petróleo (IMP), cuya presencia de mujeres es de 24-27 años y la producción varía entre 250 y 400 trabajos;
- 4) Compuesto por instituciones que producen de 100 a 200 trabajos, como la Universidad de Sonora (UNISON) y el Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), entre otras, y reportan de 13 y 24 años con presencia de mujeres. En este grupo destaca la Universidad de Guanajuato (UGTO) por la producción que registra, pero no por los años que tiene integrando investigadoras, y
- 5) Formado por las instituciones que registran menos de 100 trabajos y con escasos años con presencia de mujeres, aunque hay excepciones como Syntex, que registra menor producción, pero tienen mayor presencia de mujeres con 57 años. De hecho, es en Syntex donde se identifican las primeras mujeres investigadoras en estas disciplinas.

El resto de las instituciones registran menor producción científica y no rebasan los 20 años con presencia de mujeres, porque son instituciones de reciente creación.

Para un mejor análisis de los datos presentados en la *Tabla 1* las instituciones se clasificaron en tres grandes grupos según el tipo de investigación que realizan: 1) *tradicionales*, por la trayectoria que tienen en la actividad científica y el número de mujeres que integran entre sus afiliados; 2) *emergentes*, comienzan a incorporar investigadoras a partir de los años 80, y 3) *intermitentes*, afilian investigadoras de manera esporádica.

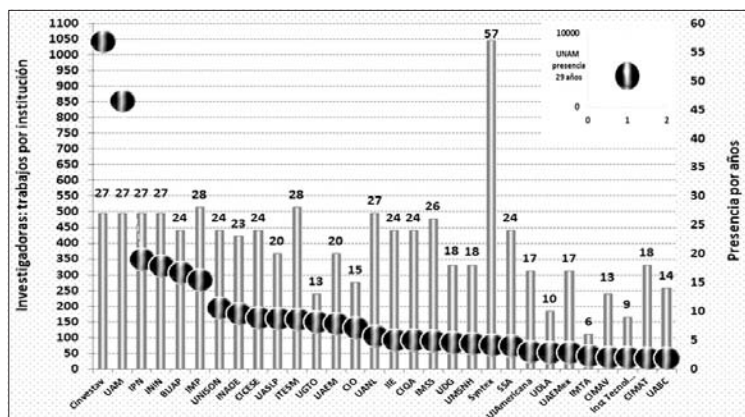


Figura 3. Producción científica por instituciones y presencia de investigadoras en años: 1900-2000

Instituciones tradicionales

La *Tabla 1* muestra por quinquenios y por instituciones la distribución de trabajos publicados por investigadoras, así como la presencia que en años registran las científicas. Cabe aclarar que el valor entre paréntesis corresponde al número de investigadoras y el dato externo a los paréntesis corresponde a los artículos. En este sentido, se pueden identificar las instituciones que más tiempo tienen en la actividad científica y las primeras que integraron mujeres entre sus grupos de investigación.

De 17 instituciones seleccionadas como tradicionales, Syntex y la UNAM son las que tiene más años incorporando investigadoras: 57 y 29 años, respectivamente. No obstante, la UNAM reporta más trabajos publicados por mujeres y mayor número de científicas, mientras que Syntex es de las que menos producción registra y menor número de investigadoras.

Dos casos que también son interesantes son los del Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM) y el IMP, ya que ambos tienen 28 años incorporando mujeres en sus equipos de investigación. Entre las dos instituciones apenas rebasan los 400 trabajos y reúnen menos de 200 investigadoras en el periodo analizado. Como se observa en la *Tabla 1*, destaca la presencia de universidades públicas y privadas, centros e institutos de investigación, hospitales e industrias, dando lugar a la clasificación de los siguientes sectores de producción: académica, investigación, salud e industria.

Desde sus orígenes estas instituciones están contribuyendo al crecimiento de la producción nacional y a la fecha continúan siendo base fundamental por la madurez que han desarrollado, con trayectorias de 23 a 27 años trabajando con mujeres investigadoras y con una producción que varía entre 72 y 1 042 trabajos.

| INSTITUCIÓN | 1900-1970 | 1971-1975 | 1976-1980 | 1981-1985 | 1986-1990 | 1991-1995 | 1996-2000 | TOTALES | PRESENCIA X AÑOS |
|----------------|-----------|------------|-------------|--------------|--------------|--------------|---------------|-------------|------------------|
| Syntex | 3 (3) | 13 (4) | 17 (7) | 10 (6) | 5 (2) | 16 (13) | 16 (2) | 80 | 57 |
| UNAM | | 42 (31) | 200 (97) | 307 (144) | 493 (239) | 876 (348) | 2301 (988) | 4219 | 29 |
| ITESM | | 8 (3) | 23 (4) | 8 (5) | 13 (10) | 29 (13) | 75 (42) | 156 | 28 |
| IMP | | 4 (4) | 22 (17) | 18 (9) | 20 (17) | 54 (34) | 166 (96) | 284 | 28 |
| Cinvestav | | 5 (2) | 43 (20) | 81 (42) | 97 (54) | 170 (102) | 646 (271) | 1042 | 27 |
| UAM | | 2 (2) | 23 (18) | 46 (22) | 68 (45) | 207 (146) | 508 (127) | 854 | 27 |
| ININ | | 2 (2) | 27 (12) | 21 (17) | 22 (19) | 76 (39) | 181 (136) | 329 | 27 |
| IPN | | 3 (3) | 12 (7) | 16 (8) | 22 (18) | 53 (47) | 242 (167) | 348 | 27 |
| UANL | | 2 (1) | 5 (2) | 5 (2) | 6 (4) | 13 (8) | 73 (37) | 104 | 27 |
| IMSS | | 8 (5) | 10 (8) | 20 (9) | 9 (11) | 18 (6) | 23 (26) | 88 | 26 |
| BUAP | | | 5 (3) | 16 (11) | 40 (23) | 56 (37) | 191 (111) | 308 | 24 |
| CICESE | | | 8 (2) | 6 (3) | 10 (3) | 23 (10) | 116 (51) | 163 | 24 |
| CIQA | | | 4 (4) | 9 (6) | 11 (5) | 9 (4) | 58 (31) | 91 | 24 |
| UNISON | | | 12 (1) | 18 (5) | 22 (18) | 37 (21) | 104 (55) | 193 | 24 |
| IIE | | | 7 (2) | 8 (4) | 16 (9) | 22 (17) | 39 (33) | 92 | 24 |
| SSA | | 2 (2) | 9 (2) | 12 (4) | 7 (5) | 13 (11) | 29 (52) | 72 | 24 |
| INAOE | | | 6 (5) | 8 (4) | 13 (4) | 24 (14) | 125 (59) | 176 | 23 |
| TOTALES | 3 | 91 | 433 | 609 | 874 | 1696 | 4893 | 8599 | 57 |

Tabla 1. Distribución de trabajos publicados por instituciones tradicionales y con mayor presencia en años de investigadoras: 1900-2000

Notas: El dato entre paréntesis corresponde al número de mujeres por institución y quinquenio. La tabla incluye a las instituciones más representativas según número de trabajos y mujeres investigadoras

Instituciones emergentes

La *Tabla 2* expone la producción de las investigadoras para las instituciones clasificadas como emergentes, debido al número de trabajos que presentan y el número de científicas que integran. En este grupo, las investigadoras empiezan a registrar trabajos de manera constante a partir del primer quinquenio de los años 80.

Como se observa, el número de investigadoras por quinquenio es escaso; hasta 1996-2000 se empiezan a mostrar los incrementos. La presencia en años de las investigadoras varía entre seis y 20 años. El grupo está constituido principalmente por universidades públicas y privadas, así como algunos centros de investigación e institutos tecnológicos, con temas de investigación procedentes de distintos sectores: académico, investigación e industria.

| INSTITUCIÓN | 1971-1975 | 1976-1980 | 1981-1985 | 1986-1990 | 1991-1995 | 1996-2000 | TOTALES | PRESENCIA X AÑOS |
|------------------|-----------|-----------|-----------|-----------|------------|-------------|---------|------------------|
| UASLP | | 1 (1) | 7 (2) | 10 (7) | 26 (19) | 116 (72) | 160 | 20 |
| UAEM | | | 5 (7) | 5 (4) | 16 (19) | 119 (74) | 145 | 20 |
| UDG | | | 3 (1) | 14 (3) | 15 (14) | 51 (44) | 83 | 18 |
| UMSNH | | | 3 (3) | 5 (1) | 8 (8) | 64 (45) | 80 | 18 |
| CIMAT | | | 3 (1) | 6 (2) | 7 (5) | 19 (16) | 35 | 18 |
| U Iberoamericana | 1 (1) | 1 (1) | 3 (1) | 12 (6) | 15 (8) | 23 (17) | 55 | 17 |
| UAEMex | | | 3 (2) | 5 (2) | 7 (5) | 36 (37) | 51 | 17 |
| CIO | | | 1 (1) | 5 (1) | 16 (8) | 108 (35) | 130 | 15 |
| UABC | | | | 4 (1) | 7 (6) | 24 (14) | 35 | 14 |
| UGTO | | | | 5 (5) | 24 (14) | 121 (77) | 150 | 13 |
| CIMAV | | | | 3 (2) | 6 (3) | 28 (17) | 37 | 13 |
| UDLA | | | | 1 (1) | 11 (11) | 41 (28) | 53 | 10 |

| | | | | | | | | |
|--------------------|----------|----------|-----------|-----------|------------|------------|-------------|-----------|
| Inst Tecnol Celaya | | | | | 8 (5) | 28 (14) | 36 | 9 |
| CENIDET | | | | | 7 (3) | 20 (10) | 27 | 8 |
| IMTA | | | | | 3 (4) | 39 (33) | 42 | 6 |
| UAQ | | | 1 (1) | | 1 (1) | 27 (24) | 29 | 6 |
| TOTALES | 1 | 2 | 29 | 75 | 177 | 864 | 1148 | 20 |

Tabla 2. Distribución de trabajos publicados por instituciones emergentes y con mayor presencia en años de investigadoras: 1971-2000

Notas: El dato entre paréntesis corresponde al número de mujeres por institución y quinquenio. La tabla incluye a las instituciones más representativas según número de trabajos y mujeres investigadoras.

Instituciones intermitentes

La *Tabla 3* presenta la producción de las investigadoras ligadas a las instituciones llamadas intermitentes, así como el número de científicas que hay por quinquenio y la presencia que las mismas tienen en años. Como se observa, estas instituciones registran la producción más baja y, en general, incluyen una o más investigadoras como parte de sus grupos de trabajo. Generalmente, las afilian por unos años, exceptuando el caso de la Universidad Autónoma de Guadalajara (UAG) y El Colegio Nacional (ECN), que tienen presencia de investigadoras por periodos de 19 y 14 años, respectivamente. Lo anterior es normal dado que las instituciones que conforman este grupo son industrias y centros de investigación que reciben apoyo de tipo gubernamental y filantrópico. En este caso es común que los investigadores cambien de institución constantemente, provocando inestabilidad y pocas posibilidades de llegar a contribuir de manera más constante al trabajo científico.

| INSTITUCIÓN | 1971-1975 | 1976-1980 | 1981-1985 | 1986-1990 | 1991-1995 | 1996-2000 | TOTALES | PRESENCIA X AÑOS |
|----------------------|-----------|-----------|-----------|-----------|-----------|-----------|---------|------------------|
| UAG | 1 (1) | 8 (5) | 1 (1) | 2 (4) | 2 (4) | 5 (10) | 19 | 19 |
| Colegio Naci | | | 2 (1) | 3 (1) | 8 (2) | 1 (1) | 14 | 14 |
| CIMMYT | | 1 (1) | 1 (1) | 2 (2) | 1 (1) | 2 (1) | 7 | 7 |
| Hosp Ctral Sur Pemex | | | | | 6 (4) | 1 (1) | 7 | 7 |
| Ind Negromex SA | | | | | 4 (5) | 3 (5) | 7 | 7 |

| | | | | | | | | |
|--------------------------|----------|-----------|----------|-----------|-----------|-----------|------------|-----------|
| INAH | 1 (1) | 1 (1) | | | 1 (1) | 3 (6) | 6 | 6 |
| UAChapingo | | | | | | 6 (6) | 6 | 6 |
| Inst Tecnol La Laguna | | | | | | 6 (7) | 6 | 6 |
| INIFAP | | | | 1 (1) | | 5 (6) | 6 | 6 |
| CONACYT | | | | 2 (1) | 1 (1) | 2 (3) | 5 | 5 |
| CIATEJ | | | | | 1 (1) | 3 (7) | 4 | 4 |
| Hylsa SA CV | | | | 1 (1) | | 3 (3) | 4 | 4 |
| Ap Green Refractories Co | | | 1 (1) | | 1 (1) | 2 (3) | 4 | 4 |
| Univ Mayab | | | | | | 4 (5) | 4 | 4 |
| Calipo Sa Cv | | | | | 4 (1) | | 4 | 4 |
| Ispat Mexicana SA CV | | | | | | 4 (3) | 4 | 4 |
| TOTALES | 2 | 10 | 5 | 11 | 29 | 50 | 107 | 19 |

Tabla 3. Distribución de trabajos publicados por instituciones iermitentes y con mayor presencia en años de investigadoras: 1971-2000

Notas: El dato entre paréntesis corresponde al número de mujeres por institución y quinquenio.

La tabla incluye a las instituciones más representativas según número de trabajos y mujeres investigadoras.

Estructuras de organización: 1900-1970

La *Figura 4* muestra la red de coautoría de los grupos de investigación que cubre el estudio: físicos, químicos, matemáticos e ingenieros en distintas especialidades. Se analizaron 70 años en los que no se observaron crecimientos extraordinarios. No obstante, permite ver la presencia de los primeros grupos de investigación conformando pequeñas redes, sobre todo en los campos de física y química. En el caso de la física, se trata de los primeros esfuerzos por hacer investigación de tipo académico desarrollada en la UNAM. Autores como Baños y Schrepf fueron los primeros en construir capital simbólico.

Por otro lado, están las redes del campo de química, donde se observa la presencia de las primeras investigadoras: Mercedes Velasco, María Luisa Franco y María E. Rivera Limón, vinculadas a Research Laboratories Syn-

tex S. A. Esto quiere decir que el desarrollo de la química en México estuvo muy ligado a la industria privada de la época, pero no sólo la química, sino también la energética y minera donde participaron empresas como American Smelting & Refining Co. y Moctezuma Copper Co.

Otros autores publicaron de manera individual. Entre ellos, los que registran mayor producción científica fueron Medina, Sánchez y Mechín. El primero estuvo ligado a la Dirección de Geografía Meteorología e Hidrología, Tacubaya. Sánchez se mantuvo bajo la afiliación de Pan American Institute of Geography & History, México, y Mechín, ligado a El Bordo en Pachuca, Hidalgo, México.

Sin duda, la industria desarrollada en México en la primera mitad del siglo XX fue generosa con las mujeres, sobre todo Syntex, que durante los años 40 encontró en territorio nacional una veta de recursos para comenzar el crecimiento de la industria farmacéutica más reconocida en el mundo (Hernández-García *et al.*, 2016).

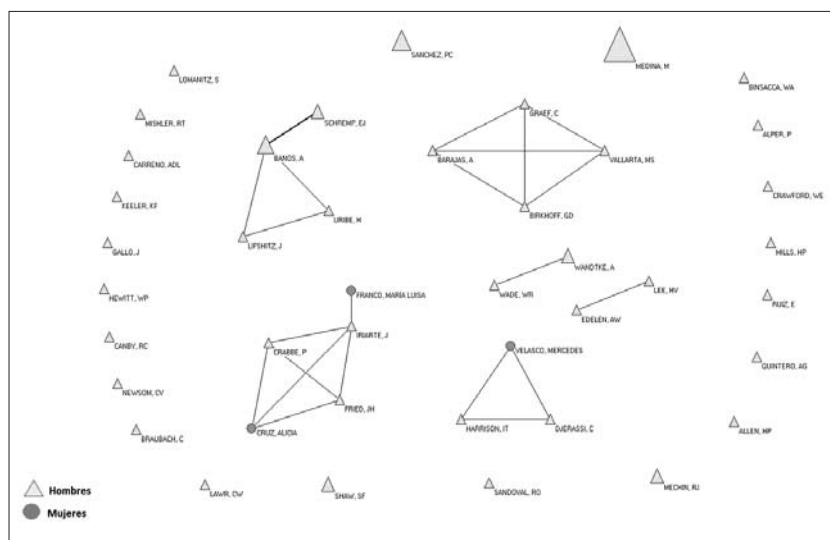


Figura 4. Red bibliométrica en ciencias exactas e ingenierías: 1900-1970

La Figura 5 muestra el desarrollo de la ciencia moderna en México en ciencias exactas e ingenierías, en las que se distinguen tres frentes de investigación: 1) académica, organizada y desarrollada en las universidades (la UNAM y el IPN que aparecen como únicos en este sector); 2) industrial, representado por la industria de la época (sobre todo farmacéutica, minería y metalúrgica), y 3) gubernamental, investigación generada en instituciones que dependían del

Estado como la Dirección de Estudios Geográficos y Climatología y la Dirección de Geología y Meteorología e Hidrología.

Esta figura hace referencia al quinquenio 1971-1975. En estos años es más evidente la presencia de investigadoras en los campos de estudio analizados, donde también se advierte la forma en que se van conformando las áreas de investigación, sobre todo la física, que registra un mayor número de redes independientes. Cada una de estas redes es liderada por autores que empiezan a sobresalir por el prestigio que van logrando, entre ellos investigadores. No obstante, la química registra la red más estructurada. Una situación similar se refleja en astrofísica, óptica y electrónica, matemáticas y el sector energético. El resto de las redes son de menor dimensión y corresponden a distintos campos de investigación donde predominan triadas, diadas y autorías únicas.

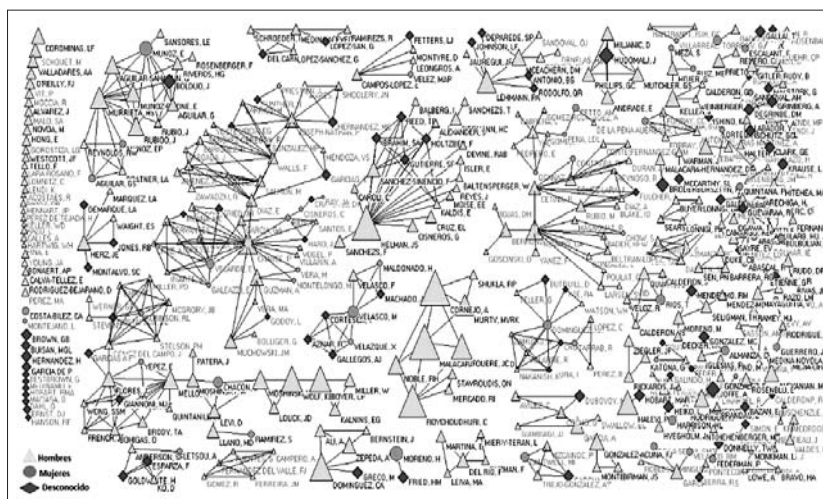


Figura 5. Red bibliométrica de coautoría en ciencias exactas e ingeniería: 1970-1975

La Figura 6 muestra redes más definidas y estructuradas representadas por autores que están logrando un alto prestigio en la investigación a través de las publicaciones dadas a conocer. Destaca como se observa un número cada vez mayor de investigadoras.

Entre los campos de investigación más constituidos destacan física y química, que en conjunto integran la red más amplia y estructurada. Por otro lado, física es el campo que más se diversifica, así lo muestran las pequeñas redes que se observan y donde se ve la presencia de autores con mayor prestigio por la producción que alcanzan.

Otro campo que también exhibe estructuras de organización bien definidas es astrofísica, óptica y electrónica. No obstante, podemos ver que predominan las autorías únicas, las diadas y triadas y en todos los casos se distinguen autores con alto capital científico (Bourdieu, 2003).

Cabe aclarar que conforme avanzan los años las investigadoras suman más y la colaboración científica es muy estrecha entre ambos géneros, lo que hace inferir que las investigadoras están contribuyendo en la construcción de los campos de investigación y en su consolidación, dando lugar a estructuras más organizadas que traen como consecuencia reconocimiento y prestigio a nivel individual, institucional, nacional e internacional.

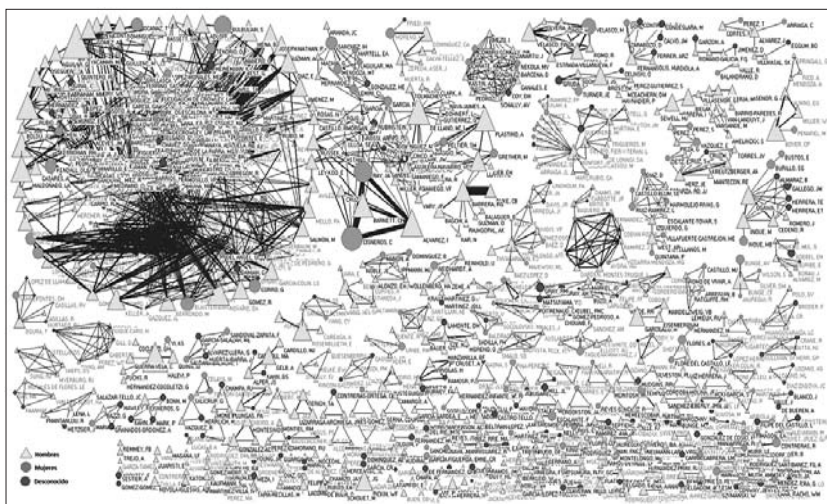


Figura 6. Red bibliométrica de coautoría en ciencias exactas e ingeniería: 1976-1980

Sin ir más allá de la década de los años 70 y por lo que muestra la red del quinquenio 1996-2000 (Figura 7), se puede deducir que las investigadoras en México hicieron su ingreso definitivo a las ciencias exactas e ingenierías en los 70 y se fortalecen en los periodos siguientes, sobre todo a partir de los 90, etapa considerada de consolidación de la ciencia en México (Luna-Morales, 2012).

Como se puede observar en la Figura 7, la presencia por género es contundente: los investigadores tienen ventaja, no obstante, las científicas suman cada vez más. Hay que aclarar que la red no muestra la presencia de mujeres que registran de uno a cuatro trabajos publicados y tampoco están presentes los autores de partículas y campos. Pese a lo anterior, las investigadoras han sido persistentes, lo que se ve reflejado en las redes y en el incremento de su participación por quinquenios.

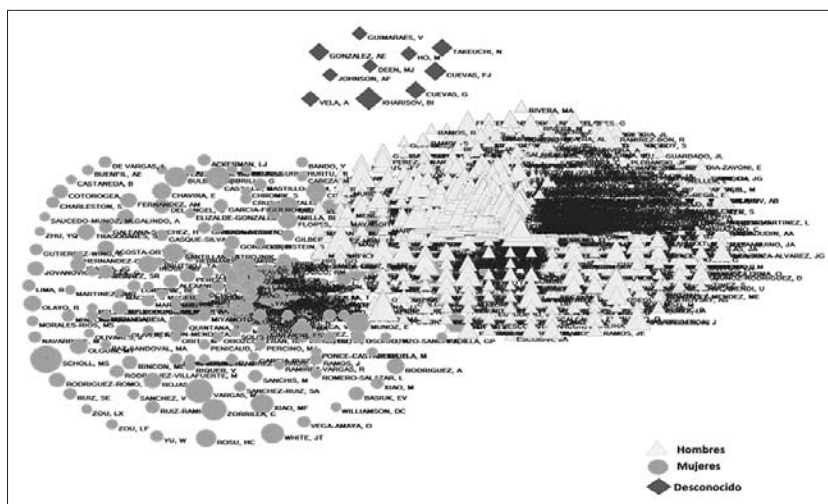


Figura 7. Red bibliométrica de coautoría en ciencias exactas e ingeniería: 1996-2000

DISCUSIÓN

Las mujeres investigadoras adscritas a las áreas de ciencias exactas e ingenierías han logrado progresos en México. Esto se refleja desde distintos aspectos: el incremento de la matrícula educativa, sobre todo en ingenierías (Valle Díaz-Muñoz y Garay-Sánchez, 2012; Razo-Godínez, 2008), y la producción científica registrada en la literatura de corriente principal. Lo anterior se corresponde con estudios internacionales como el de la Unesco (2017), que señala crecimientos de 47 % en la matrícula de educación superior en campos como las ciencias naturales y 25 % en las ingenierías.

Desde el punto de vista bibliométrico en México, la incorporación de investigadoras a las ciencias exactas e ingenierías ocurrió durante las décadas de 1940 a 1960 y fue en Syntex donde se afiliaron por primera vez (Hernández-García *et al.*, 2016: 10). Durante los años 70 se identificaron más investigadoras en estos campos, el número se incrementó en los 80 y, finalmente, en los 90 se consolidó su participación en la investigación científica. Estos impactos tienen que ver con la apertura y consolidación en los años 60 y 70 de diversas instituciones en el país: Cinvestav, IMP y la UAM (Torres-Cruz, 2009), así como la fundación del Consejo Nacional de Ciencia y Tecnología (Conacyt), organismo rector en el desarrollo de políticas científicas nacionales (Rubio-Castillo, 2009).

Estos eventos se complementan con el esfuerzo que realizan las instituciones de mayor trayectoria en México, así como la integración de programas de posgrados en ciencias exactas e ingenierías en distintas instituciones de educación superior y centros de investigación del país. Asimismo, a estos esfuerzos se suman los de instituciones que se fundaron durante los años 80 y 90 como los centros públicos Conacyt y los institutos tecnológicos (Rubio-Castillo, 2009). En conjunto, dichas iniciativas crearon las condiciones para la incorporación de investigadores (hombres y mujeres) al desarrollo científico. Sin embargo, las diferencias de género afectan a las investigadoras, sobre todo en el campo de las ciencias exactas e ingenierías, donde siguen teniendo escasa presencia. Esta situación es reconocida por las propias mujeres.

Por ello, a nivel internacional y nacional, se ha considerado necesario tomar en cuenta factores como la integración de programas educativos que consideren a las mujeres desde la niñez en estos campos de estudio, y que las instituciones de educación superior trabajen con una perspectiva de género que tome en cuenta la incursión de las mujeres en las ciencias exactas e ingenierías, a fin de incrementar la incursión en estos campos de estudio (Valle Díaz-Muñoz y Garay-Sánchez, 2012), además de eliminar las barreras que todavía se oponen a su desarrollo científico.

Las redes bibliométricas permiten ver que se trata de campos bien estructurados, sobre todo en la física, que, a partir de la primera mitad de los años 90, se integró a las colaboraciones de *Big Science*, provocando crecimientos en los trabajos publicados y en el número de investigadores (Luna-Morales y Collazo-Reyes, 2002), dando lugar a redes más complejas. A lo anterior se suman los cambios en las políticas de publicación que adquirieron los ingenieros en computación, quienes, durante los años 90, orientaron sus resultados de investigación a la publicación en revistas científicas (Luna-Morales, 2009), generando incrementos en la producción científica.

CONCLUSIONES

Los crecimientos más altos en ciencias exactas e ingenierías ocurren a partir de los años 70 y se consolidan durante los 90. Destacan por sus aportaciones y número de investigadoras la UNAM, el Cinvestav, la UAM-I y laboratorios Syntex, entre otras instituciones.

Las redes bibliométricas muestran que desde los años 40 y 60 del siglo XX se publicaron los primeros trabajos de investigadoras y se intensificó su participación durante el periodo de los 70.

La incorporación de mujeres a la investigación en ciencias exactas e ingenierías en México es un suceso que está ligado a tres eventos: 1) la incorporación de la mujer a la educación superior y la investigación científica, suceso ocurrido en los años 70; 2) la creación de nuevas instituciones en el país durante las décadas de los años 60 y 70, y 3) que las mujeres en nuestro país están rompiendo el cerco que, por años, ha impuesto la familia y la sociedad al prohibir su incorporación a las ciencias consideradas del dominio masculino (García-Guevara, 2013).

La complementación de estos aspectos ha permitido que la mujer investigadora consiga espacios a través de las distintas instituciones de investigación establecidas en el país –así como en las industrias instaladas en México–, logrando con ello visibilidad y reconocimiento mediante la difusión de trabajos publicados en revistas de corriente principal incluidas en los índices de citas WoS.

Sin duda, las contribuciones de este trabajo generan ventajas orientadas a diversos grupos, uno de ellos es la misma comunidad científica al contar con un estudio que cubre desde sus orígenes, abarcando el periodo de los años 90 considerado de consolidación de la ciencia en México (Luna-Morales, 2012). Asimismo, se trata de un tema que no ha sido abordado con anterioridad, en este sentido, los resultados serán de interés para otros sectores como las instancias que buscan la equidad de género, las instituciones de educación superior y de investigación que integran estos campos de estudio; de igual manera servirán en la definición de políticas científicas tanto institucionales como nacionales.

Debido a lo antes expuesto y por los crecimientos que muestran las ciencias exactas e ingenierías en el periodo analizado, se pretende a futuro abordar los campos de estudio de manera independiente por considerar que cada uno de ellos integra aspectos que merecen ser analizados por separado, con la idea de caracterizar más a detalle las formas de evolución y la incorporación de las científicas en periodos específicos de tiempo.

REFERENCIAS

- Batagelj, Vladimir y Andrej Mrvar. 2011. *Pajek Program for Analysis and Visualization of Large Networks: Reference Manual List of commands with version 2.05short explanation*. Slovenia: University of Ljubljana. Fecha de consulta: 14 de agosto de 2017. <http://vlado.fmf.uni-lj.si/pub/networks/pajek/doc/pajekman.pdf>
- Boston Consulting Group, The. 2013. *Overview women status in science*. Boston Massachussets: BBCG.

- Bourdieu, Pierre. 2003. *El oficio del científico: ciencia de la ciencia y reflexividad*. Barcelona: Anagrama.
- Buquet-Corleto, Ana Gabriela. 2013. "Sesgos de género en las trayectorias académicas universitarias: orden cultural y estructura social en la división sexual del trabajo". Tesis de doctorado, UNAM, Facultad de Ciencias Políticas y Sociales.
- Cepeda-Zetter, B., C. González-Brambila y M. A. Pérez-Angón. 2017. "Gender segregated analysis of mexican inventors in patent applications under the Patent Cooperation Treaty (PCT)". *Interciencia* 42 (4): 204-211.
- Contreras-Gómez L. E., R. Baquero-Parra, E. Robles-Belmont y M. A. Pérez-Angón. 2015. "Patrones de movilidad de los físicos mexicanos en el Sistema Nacional de Investigadores". *Interciencia* 40: 525-532.
- Didou, S. y G. Etienne. 2011. "El Sistema Nacional de Investigadores en 2009: ¿Un vector para la internacionalización de las élites científicas?" *Perfiles Educativos* 33 (132): 29-47.
- García-Guevara, P. 2013. "El Género y las ingenierías en computación". *Tarbiya* 34: 81-89.
- Gorbea-Portal, S. 2013. "Tendencias transdisciplinarias en los estudios métricos de la información y su relación con la gestión de la información y del conocimiento". *Perspectivas em Gestão & Conhecimento* 3 (1): 1-15.
- Guevara-Ruiseñor, E. L. y Alba E. García-López. 2010. "Orden de género y trayectoria escolar en mujeres estudiantes de ciencias exactas y naturales". *Investigación y Ciencia* 46: 10-17.
- Hernández-García, A. 2014. "Redes de colaboración científica y su efecto en la productividad. Un análisis bibliométrico". *Investigación Bibliotecológica* 27 (59): 159-175.
- Hernández-García, Y., J. A. Chamizo, M. Kleiche-Dray y J. M. Russell. 2016. "The Scientific Impact of Mexican Steroid Research 1935-1965: A Bibliometric and Historiographic Analysis". *Journal of the Association for Information Science and Technology* 67 (5): 1245-1256.
- L'Oreal-Unesco a The Boston Consulting Group. 2014. *Women & Science: Less than 1 researcher in 3 is a woman, according to a report commissioned by the L'Oreal Foundation*. Boston: Foundation L'Oreal. Fecha de consulta: 11 de octubre de 2017. <http://www.loreal.com/media/press-releases/2014/mar/women--science--less-than-1-researcher-in-3-is-a-woman--according-to-a-report-commissioned-by-the-l%E2%80%99or%C3%A9al-foundation>
- Luna-Morales, M. E. 2012. "Determinants of the maturing process of the mexican research output: 1980-2009". *Interciencia* 37 (10): 736-742.
- Luna-Morales, M. E. 2009. "La maduración de la ciencia mexicana: un análisis histórico bibliométrico de su desarrollo de 1980-2004". Tesis de doctorado, UNAM, Facultad de Filosofía y Letras, Centro Universitario de Investigaciones Bibliotecológicas.
- Luna-Morales, M. E. y F. Collazo-Reyes. 2002. "El Síndrome Big Science y su influencia en el proceso de maduración de la física mexicana de partículas elementales". *Revista Española de Documentación Científica* 25 (4): 409-420.
- Mendieta-Ramírez, A. 2015. "Desarrollo de las mujeres en la ciencia y la investigación en México: un campo por cultivar". *Agricultura, Sociedad y Desarrollo (ASyD)* 12: 107-115.

- Meza-Montes, L. 2005. "El grupo de trabajo mujeres en la física de la Sociedad Mexicana de Física", en *II Encuentro Participación de la Mujer en la Ciencia*, León, Guanajuato, México, mayo 19-20.
- Miroux, Anne. 2011. "Mainstreaming a gender perspective in science, technology and innovation policy", en *United Nations Commission on the Status of Women*, Fifty-fifth session, New York, 4 de febrero-22 de marzo. Fecha de consulta: 18 de julio de 2016. <http://www.un.org/womenwatch/daw/csw/csw55/panels/Pannel1-Miroux-Anne.pdf>
- Moya-Anegón, F., Z. Chinchilla-Rodriguez, B. Vargas-Quesada, E. Corera-Alvarez, A. Munoz-Molina, F. J. Munoz-Fernandez y R. Gomez-Crisostomo. 2007. "Scientific output by gender in Spain (Web of Science, 2004)". *Proceedings of ISSI 2007: 11th International Conference of the International Society for Scientometrics and Informetrics*, 568-582. Madrid.
- Narváez, Mónica. 2016. "Perspectiva sobre las mujeres en las ciencias exactas". *Radio Concencia*, Mérida, Yucatán, 4 de abril.
- Oscar-Lluch, J. 2010. "Aplicación del análisis de redes al estudio de la investigación española de historia de la ciencia". *REDES* 19 (6): 122-143.
- Pinto, A. L., J. A. Moreira-González y G. A. Oliveira de Meira. 2009. "Análisis de redes sociales a partir de recursos web y de bases de datos especializadas en literatura científica". *Anales de Documentación* 12: 139-158.
- Pollack, Martha E. 2013. "Why Are There Still So Few Women in Science?" *The New York Times*, 3 de octubre.
- Pritchard, A. 1969. "Statistical bibliography on bibliometrics". *Journal of Documentation* 25 (4): 348-349.
- Razo-Godínez, M. L. 2008. "La inserción de las mujeres en las carreras de ingeniería y tecnología". *Perfiles Educativos* (junio): 63-96.
- Retana-Guiascón, O. G. 2009. "La institucionalización de la investigación científica en México: breve cronología". *Ciencia* 94 (abril-junio): 47-51.
- Rosser, Margaret. 1982. *Women scientists in America: struggles and strategies to 1940*. Baltimore: London, John Hopkins University Press.
- Rubio-Castillo, F.A. 2009. "La Estructura Organizacional en Centros de Investigación, Desarrollo e Innovación; una Aproximación a la Experiencia Internacional". Tesis de doctorado, Universidad Autónoma de Querétaro, Facultad de Contaduría y Administración.
- Sugimoto, C. R. 2013. "Global gender disparities in science". *Nature* 12 (December): 211-213.
- Torres-Cruz, I. 2009. "A sus 35 años la UAM concreta su cuarto campus y piensa que no hay quinto malo; atiende a 45 mil estudiantes". *La Crónica*, junio. Fecha de consulta: 12 de abril de 2015. <http://www.cronica.com.mx/notas/2009/437832.html>
- Unesco. 2017. *Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM)*. París: Unesco.
- Unesco. 2010. *Gender, science and technology: report of the expert group meeting*. París: Unesco.
- Velasco, B., J. M. Eiros-Bouza, J. M. Pinilla y J. A. San Román. 2012. "La utilización de los indicadores bibliométricos para evaluar la actividad investigadora". *Aula Abierta* 40 (2): 75-84.
- Valle Díaz-Muñoz del, G. y A. de Garay-Sánchez. 2012. "La falta de inclusión de mujeres en ciencias exactas e ingenierías". *Ciencia* 4: 34-43.

Para citar este texto:

Luna Morales, María Elena y Evelia Luna Morales. 2018. "Mujeres investigadoras en las primeras estructuras de organización en ciencias exactas e ingenierías en México de 1900-2000: Estudio Bibliométrico". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 193-215.
<http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57899>

DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57899>