

Revista española de documentación científica

vol. 44, n. 2 (2021)

Estudios

¿Quién es autor? Las autorías múltiples, criterios y lineamientos

Alberto Chamorro

Producción científica sobre ansiedad bibliotecaria: un análisis bibliométrico y cienciométrico desde Scopus

Pier Suclupe-Navarro, Cesar H. Limaymanta, Néstor Holmes Ramírez, Héctor Guillén

Estado de la web local portuguesa: evidencias empíricas acerca de la brecha digital entre regiones

Valeriano Piñeiro-Naval, Paulo Serra

La transformación digital en los métodos y temas de la investigación brasileña de Información y Documentación 2010-2019

Francisco Carlos Paletta, José-Antonio Moreiro-González

Tendencias de las búsquedas de información sobre las colecciones SciELO, Redalyc y Dialnet realizadas a través de Google

Cristina Bojo-Canales, María Sanz-Lorente, Javier Sanz-Valero

Evaluación de la investigación con encuestas en artículos publicados en revistas del área de Biblioteconomía y Documentación

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

Curación de contenidos en periodismo. Indicadores y buenas prácticas

Javier Guallar, Rafael Pedraza-Jiménez, Mario Pérez-Montoro, Laura Anton

Noticias

ENDORSE (European Data Conference on Reference Data and Semantics)

Teresa Abejón Peña

ESTUDIOS / RESEARCH STUDIES

¿Quién es autor? Las autorías múltiples, criterios y lineamientos

Alberto Chamorro*

*Corporación Universitaria Minuto de Dios, Colombia

Correo-e: alberto.chamorro@uniminuto.edu | Orcid ID: <https://orcid.org/0000-0002-6914-3160>

Recibido: 03-03-20; 2ª versión: 27-04-20; Aceptado: 08-05-20; Publicado: 04-05-21

Cómo citar este artículo/Citation: Chamorro, A. (2021). ¿Quién es autor? Las autorías múltiples, criterios y lineamientos. *Revista Española de Documentación Científica*, 44 (2), e290. <https://doi.org/10.3989/redc.2021.2.1758>

Resumen: Este artículo tiene por objeto: 1. Poner en evidencia las prácticas injustificadas más recurrentes en la asignación de autorías en las publicaciones científicas. 2. Rastrear los criterios frecuentemente aceptados dentro de la comunidad académica para establecer quién puede, en propiedad, llamarse autor, y 3. Proponer un conjunto de medidas que permitan valorar apropiadamente artículos con múltiples autores. Para ello se hizo una revisión de artículos publicados en Pubmed y Scopus usando palabras claves como "autoría" (authorship), "criterios para establecer autorías" (authorship guidelines), "artículos con múltiples autores" (multiple authors). A partir de los hallazgos, se indagaron las irregularidades más relevantes y los principios más aceptados. Como resultado, a través de los documentos consultados fueron identificadas las prácticas más censurables y las fuentes de criterios más reconocidas para establecer las autorías: International Committee of Medical Journal Editors (ICMJE), Council Science Editors (CSE), The World Association of Medical Editors (WAME), National Institutes of Health (NIH), The American Psychological Association (APA) y The Center for Values, Ethics and the Law in Medicine, Sidney University (SEH). Debido a que gran parte de los lineamientos dados por estas organizaciones son en su mayor parte impracticables, se propone una nueva forma y se concluye que es menester que tanto financiadores, instituciones, editoriales, editores en jefe e investigadores, asuman ciertas funciones de control y seguimiento, de tal forma que se preserve la integridad científica de las publicaciones, sin interferencia de las métricas.

Palabras clave: autor; autoría; múltiples autores; integridad científica; ética en publicaciones

Who is author? Multiple authorships, criteria and guidelines

Abstract: This paper aims are: 1. Highlighting the most common unethical practices regarding authorship in scientific publications, 2. Tracing authorship criteria from most respectful academic sources and 3. Making recommendations to deal with multiple authored papers. To achieve this, a review was made on those articles obtained through a search on Pubmed and Scopus databases with the terms "authorship", "authorship guidelines" and "multiple authors". From this, the more common unethical practices and the most acknowledged sources of guidelines were identified, most of them coming from relevant medical associations: International Committee of Medical Journal Editors (ICMJE), Council Science Editors (CSE), The World Association of Medical Editors (WAME), National Institutes of Health (NIH), The American Psychological Association (APA) y The Center for Values, Ethics and the Law in Medicine, Sidney University – VELIM (SEH). Given that most of those guidelines are inapplicable in practice, a more complete view is needed, where all the parts involved will take some responsibility. It is required that financial organizations, scientific institutions, publisher, editors and researchers take some action to assure research quality and true information, without metrics interference.

Key words: author; authorship; multiple authors; research integrity; ethics on publications.

Copyright: © 2021 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1. INTRODUCCIÓN

Ser reconocido como autor tiene muchas implicaciones dentro de la comunidad científica. La autoría es una moneda de intercambio que permite crear valor, reconocer, otorgar o intercambiar créditos (Bemmet y Taylor, 2003). Así, cuantos más manuscritos se publican, mayores beneficios tendrá un autor, dado que del número y calidad de las publicaciones dependerá su categorización, su puesto dentro del escalafón docente, el incremento de su salario, el acceso a futuros fondos para la subvención de proyectos y la posibilidad de postularse a mejores cargos. Esto, sin mencionar el efecto reputacional que las publicaciones puedan tener sobre las instituciones universitarias a las que los autores se encuentran afiliados.

Las exigencias y estímulos derivados de las publicaciones han presionado a los investigadores para aumentar su número y frecuencia, y no pocos se han visto tentados a recurrir a prácticas éticamente discutibles (Angell, 1994), tales como: la realización de estudios triviales, reportar los resultados de investigación de manera fragmentada y presentar varias versiones de una misma investigación (Bailar, 1986). Dentro de esas prácticas también se ha vuelto recurrente listar como autores a académicos que no han tenido un contacto significativo con los trabajos, particularmente cuando la tendencia de publicaciones en coautorías parece ser hoy la regla.

En efecto, estudios recientes muestran como las publicaciones en coautoría han venido en aumento en diversas disciplinas, entre ellas: Economía (Kuld y O'Hagan, 2017), Administración (Koseoglu, 2016), Ciencias Naturales (Mallapaty, S. 2018) y en Humanidades y Ciencias Sociales (Taylor and Francis Group, 2017). No se puede negar que las nuevas exigencias de las ciencias plantean nuevos retos a la manera de producir conocimiento. Las preguntas a las que se enfrentan las ciencias de hoy son múltiples, diversas en naturaleza y complejas, por lo que requieren la combinación de diversas disciplinas, tecnologías y variedad de muestras (Baethge, 2008). Por ejemplo, el proyecto acerca del genoma humano, iniciado en el año 1988 y completado en el 2013, requirió la conformación de un consorcio internacional que involucró a científicos de diferentes países (National Human Genome Research Institute, 2019), igual sucede con proyectos de investigación relativos al Cáncer (Big-TenCRC, 2019), Pediatría (PeRC, 2019), Educación (CPRE, 2020), y Física (UKPC, 2019), entre muchos otros. El abordaje de muchos problemas requiere hoy el concurso de estadísticos, diseñadores metodológicos, expertos en modelos matemáticos, expertos en recogida de datos, programadores y teóricos disciplinares. Indiscutiblemente, la elaboración de textos

en coautorías seguirá en incremento. La publicación más emblemática, que ilustra hasta dónde puede llegar el número de colaboradores, lo constituyen las 9 primeras páginas dedicadas a 5.154 autores, de las 33 que conforman el artículo, publicado en *Physical Review Letters* (Aad, 2015), sobre el proyecto europeo ATLAS.

2. JUSTIFICACIÓN Y METODOLOGÍA

Al tiempo que la forma de producir conocimiento cambia, se requiere también tener mucha claridad acerca de quiénes pueden, en propiedad, llamarse autores. El concepto de autor no solo está relacionado con créditos y beneficios, sino que también implica un asunto ético: es quién asume la responsabilidad por las publicaciones, la veracidad de la información y sus consecuencias. Así, por ejemplo, en 1983 John Darsee se hizo tristemente célebre cuando fueron revelados como falsos los resultados de sus estudios realizados en las universidades de Emory y Harvard (Broad, 1983). Este hecho tuvo aún mayor repercusión al darse a conocer que muchos directores de los departamentos con quienes trabajó compartían las coautorías de sus publicaciones (Smith, 1994). Incidentes como este obligaron a las comunidades académicas a plantearse el problema ético del autor, ya que este tipo de casos ponen en riesgo la integridad científica, la calidad de la investigación, y la confianza en la información y en los hallazgos reportados.

En consecuencia, el presente artículo gira sobre la pregunta: ¿A quién puede atribuirse realmente el título de autor? O, dicho de otra manera: ¿Qué criterios deben cumplirse para ser incorporado en la lista de autores de una publicación?

Para responder la pregunta se optó inicialmente por hacer una revisión documental en Scopus y en PubMed, ensayando varias opciones de búsqueda, tales como: *authorship*, *guidelines* y *multiple authors*. Se usó como margen de tiempo los últimos 10 años (2010-2020). En la Tabla I, se relacionan las entradas con sus posibles combinaciones y los resultados por base de datos:

Tabla I. Resultado de las búsquedas sobre autoría múltiple por bases de datos (Número de documentos 2010-2020)

	Pub Med	Scopus
authorship	32.273	21.911
(authorship) AND ("guidelines")	1.357	1.107
(authorship) AND ("multiple author")	7	624
authorship and "guidelines" and "multi author"	4	39

El número de publicaciones referentes a los temas de autorías ha ido creciendo exponencialmente en los últimos 10 años. Se revela, además, una variedad de subcategorías dentro de este apartado, siendo la más notoria las autorías injustificadas, seguidas por la cuestión del orden de las firmas, el tema del género, el debate sobre los tutores como autores, y la disyuntiva entre autor y colaborador. Algunos estudios han diferenciado hasta 10 subcategorías: atribución, violaciones, parcialidad, responsabilidad, orden de presentación, citaciones y referencias, definición, estrategia de publicación, originalidad y sanciones (Hosseini & Gordijin, 2020).

Dentro de la categoría más amplia (authorship) se contó el número de documentos que, de algún modo, trataban sobre las prácticas más censurables. El concepto "Ghost" registró el mayor número de documentos tanto en Scopus como en PubMed, mientras que el de menor relevancia en ambas fue "Gift". Los resultados se pueden apreciar en la Tabla II.

Tabla II. Número de documentos sobre autoría múltiple según práctica fraudulenta

Authorship	Scopus	PubMed
Ghost	107	71
Honorary	84	69
Guest	60	51
Gift	78	38

Las categorías de "pressurized authorship" y "author pooling", no aparecen suficientemente documentadas, pero han sido incluidas en este documento para llamar la atención sobre este tipo de prácticas.

Finalmente, se revisaron las fuentes más recurrentes combinándolas con la categoría de "authorship", y se encontró el siguiente número de coincidencias (Tabla III):

Tabla III: Coincidencias encontradas de organizaciones en bases de datos sobre autoría múltiple

Authorship	ICMJE	NIH	APA	WAME	CSE
PubMed	370	1654	17	6	3
Scopus	163	100	47	7	2

La primera tabla nos permitió decantar un número muy limitado de artículos concernientes a la cuestión ética de las autorías múltiples, que serán citados a lo largo de este trabajo. La segunda tabla

nos muestra la preponderancia de algunas prácticas consideradas como injustificadas. Y, la tercera, evidencia la relevancia que algunas organizaciones tienen cuando se trata de abordar el tema ético de la autoría.

Gran parte de la literatura revisada muestra que los lineamientos éticos que han servido como referentes para definir quién puede ser considerado como autor, se han concentrado en los planteados por un reducido número de asociaciones. Gran parte de ellas en el campo de la Medicina y las Ciencias de la Salud. Los aspectos sobresalientes de cada una de estas asociaciones serán presentados a continuación, pero antes se definirán las prácticas inapropiadas más recurrentes que, precisamente, los lineamientos éticos de las diferentes asociaciones pretenden combatir.

3. AUTORÍAS INJUSTIFICADAS

A partir de los textos consultados, se puede evidenciar que la comunidad científica coincide en identificar algunas malas prácticas que han venido deteriorando el concepto de las autorías:

3.1. Autorías regaladas, honoríficas o invitadas (Da Silva T. , 2015). Estas tres denominaciones tienen en común que la persona que ha realizado todo el trabajo de investigación y redacción decide, voluntariamente, incluir como autores a otras personas que no han tenido ninguna participación en la publicación. Dependiendo de la motivación, recibe una u otra denominación. Así, en la autoría regalada, *gift author*, se incluye como autor a un junior o senior investigador con la esperanza que éste devuelva el favor en un futuro. Por su parte, la autoría honorífica (Michaela, 2016), *honorary author*, se refiere a aquella participación de investigadores seniors o personas con capacidad en la toma de decisión (jefes de departamentos, directores de programas, decanos, directores de centros de investigación o rectores), quienes, a cambio de ser incluidos en los artículos, facilitan el desarrollo del proyecto (garantizan los recursos necesarios: personas, tiempos y financiación). Finalmente, la práctica de incluir a un autor invitado, *guest author*, tiene por objeto, comúnmente, que un investigador de mayor trayectoria, incorporado al listado de autores, es con la esperanza de mejorar las oportunidades de publicación e incrementar las citas.

3.2. El Autor fantasma, ghost autor (Fletcher, 2005). Pese a que el investigador participa en el proyecto, y es pieza importante en la redacción, su nombre es omitido en la lista de auto-

res. Esto puede obedecer a diversas razones: cuando el autor tiene conflictos de intereses y no desea que su nombre sea público, o cuando el centro de investigación o laboratorio, después de haber llegado a las conclusiones de un proyecto, pasa los datos a un académico para que sea este quién escriba y publique el documento con apariencia de objetividad (Sismondo, 2005). Este tipo de publicaciones han sido centro de preocupaciones por parte de la comunidad académica (The PLoS Medicine Editors, 2009).

3.3. Autorías bajo presión, *pressurized authorship* (Das y otros, 2016). Tiene lugar cuando personas en posición de autoridad o con alguna ventaja ocasional, presionan a los investigadores originales para incluir otros nombres, ajenos al proyecto, en la lista de autores.

3.4. Comunidad de autores, *author pooling*. Es una especie de autoría invitada (*guest authorship*), con la variante de que un grupo de académicos han pactado previamente escribir cada uno por su cuenta determinado número de artículos, con el compromiso de listar como autores a todos los firmantes del pacto. El impacto de la producción de todos se incrementa en cuanto aumentan los miembros del equipo y el número de artículos acordados (Bhaskar, 2010).

La proliferación de estas prácticas desvirtúa la calidad de la información que los lectores de las revistas científicas reciben, provee falsos datos para la valoración que las entidades financiadoras realizan sobre las instituciones e individuos, promueve erróneamente los créditos que un investigador podría recibir por su trabajo e induce a la toma de decisiones equivocadas en materia de promociones. Esto sin mencionar el vacío de responsabilidad sobre la calidad de los datos recopilados y sus posibles consecuencias. Todas estas conductas son tipificadas como malas prácticas investigadoras que ponen en riesgo la integridad científica de los autores implicados, las instituciones que representan y las revistas en las que publican (Harvey, 2018).

4. ORIENTACIONES ÉTICAS PARA EL CONTROL Y EL MANEJO DE LAS AUTORÍAS

La cuestión de fondo es definir a quién se le puede atribuir en propiedad el título de autor. El asunto se hace más complejo si se tiene en cuenta que la comprensión de autor ha sufrido diversas acepciones (Da Silva y Dobráski, 2016). No es la intención de este trabajo rastrear históricamente el concepto sino, de acuerdo con los resultados de la búsqueda realizada, concentrarse en los lineamientos que las organizaciones con mayor influencia han

venido construyendo y que luego se han convertido en normas de buenas prácticas de editoriales científicas. Pese a que existen acercamientos previos al respecto (Ruíz-Pérez y otros, 2014), un nuevo acercamiento de carácter general se hace necesario.

El punto de partida será la organización de más amplia cobertura (CSE), luego la exposición se concentrará en las organizaciones médicas y del campo de la salud más representativas (ICMJE, WAME, NIH y APA) y, finalmente, se comentará un documento elaborado por Sydney Health Ethics (SEH) que, si bien no aparece referenciado en PubMed ni Scopus, logra consolidar una serie de principios generales de gran relevancia.

4.1 Consejo de Editores Científicos (Council of Science Editors, CSE)

El Consejo de Editores Científicos (Council of Science Editors, CSE) se creó en 1957 para apoyar las Ciencias Biológicas, pero después de varios años, cambió su enfoque al conjunto de las ciencias, y redefinió su misión: educar, orientar las prácticas de los editores y asesorarlos en los aspectos éticos de su quehacer (CSE, 2020). Dado su campo de influencia, CSE presenta el asunto a mayor escala. Reconoce que diversas instituciones académicas y entes financiadores han desarrollado sus propios lineamientos para definir el rol de los autores. También es consciente de que existen muchos principios, costumbres y prácticas que pueden diferir entre disciplinas, organizaciones y países. Por esto, el esfuerzo de CSE está dirigido a integrar toda esta diversidad en unos principios generales (CSE, 2012):

- Identificar quién es autor es responsabilidad del equipo que ha realizado el trabajo, y no del editor ni de quien realiza la publicación.
- Individuos cuya contribución no es suficiente para calificarles como autores, deben ser nombrados en la sección de reconocimientos.
- Todo individuo que se califique como autor o colaborador debe ser identificado de acuerdo con su participación.
- Quienes aparecen listados como autores deben revisar y aprobar el manuscrito final.
- Todos los autores deben especificar su tipo de contribución.
- La razón principal para identificar a los autores es establecer su grado de responsabilidad.
- Es responsabilidad del autor principal decidir quiénes pueden ser listados como autores y a quiénes debe darse reconocimiento. Así mismo, deben resolver las reclamaciones y los conflictos que surjan.

A diferencia de otras organizaciones, CSE explicita qué tipo de actividades no se califican como autorías: adquisición de fondos, recolección de datos, proporcionar retroalimentación tanto del proceso como del manuscrito, participar como consultor o recibir pagos por servicios realizados en la investigación, supervisar el trabajo o ser directivo del departamento donde tiene lugar la investigación.

4.2 Comité Internacional de Editores de Revistas Médicas (International Committee of Medical Journal Editors, ICMJE)

La organización ICMJE, también conocida como el grupo de Vancouver, está conformado por un pequeño grupo de editores de revistas y representantes de organizaciones en el campo de la Medicina y las Ciencias de la Salud, cuyo objetivo es asegurar la calidad de la conducta, los reportes y las publicaciones de los académicos en las Ciencias de la Medicina (ICMJE, 2020). Entre sus integrantes se encuentran: *British Medical Journal* (Revista Médica Británica); *Bulletin of the World Health Organization* (Boletín de la Organización Mundial de Salud); *Deutsches Ärzteblatt* (Revista Alemana de Medicina); *Ethiopian Journal of Health Sciences* (Revista de Ciencias de la Salud de Ethiopia); *JAMA* (Revista de la Asociación Americana de Medicina); *Journal of Korean Medical Science* (Revista Coreana de Ciencias Médicas); *New England Journal of Medicine* (Revista Médica de Nueva Inglaterra); *New Zealand Medical Journal* (Revista Médica de Nueva Zelanda); *The Lancet*; *Revista Médica de Chile*; *Ugeskrift for Laeger* (Revista Médica Danesa); *The U.S. National Library of Medicine* (Biblioteca Nacional de Medicina de Estados Unidos); y *The World Association of Medical Editors* (Asociación Mundial de Editores Médicos). Pese al reducido número de sus miembros, las recomendaciones y lineamientos de ICMJE, publicados desde 1978, son seguidos por gran parte de reconocidas casas editoriales, tales como ELSEVIER (ELSEVIER, 2019) y PlosOne (PlosOne, 2019), entre otras.

Históricamente este comité ha venido ajustando su comprensión de lo que debe esperarse de un autor (Misra, 2018). Sin embargo, fue después del escándalo del caso Darsee (Culliton, 1983), cuando se formularon criterios más precisos y exigentes. Desde entonces, para ser autor hay que cumplir los siguientes criterios en su totalidad (ICMJE, 2019):

- Haber realizado contribuciones sustanciales en la concepción y diseño del proyecto, o contribuir sustancialmente en el análisis e interpretación de los datos.
- Haber escrito el documento final o revisado críticamente haciendo aportes al contenido intelectual del mismo.

- Haber participado en la aprobación del manuscrito final.
- Ser responsable del contenido del documento, responder de su veracidad e integridad en cada una de sus partes, cuando sea requerido.

La declaración de ICMJE reacciona con firmeza ante los abusos presentados y es enfática en indicar que todo autor debe cumplir con todos los cuatro criterios enunciados anteriormente, y todo aquel que cumpla con esos criterios ha de ser reconocido como autor. Ser autor implicará, además, asumir la responsabilidad y dar cuenta de todos los aspectos y etapas de la investigación y su correspondiente publicación. Aquellos que participan del primer criterio han de tener la oportunidad de colaborar en los otros tres, pero aquellos que solo toman parte de algunos de los procesos merecen ser reconocidos como colaboradores.

Sin embargo, la experiencia ha demostrado que no siempre se cumple con los principios de ICMJE, el trabajo en equipo permite que cada quién colabore de acuerdo con su experiencia, y no todos los participantes pueden responder en su totalidad de todas las fases del proceso de investigación. Estos principios han sido extremadamente restrictivos e impracticables.

4.3 La Asociación Mundial Editores Médicos (The World Association of Medical Editors, WAME)

La Asociación Mundial de Editores Médicos (World Association of Medical Editors, WAME), tuvo su origen en 1995 y ha tenido como objeto la estandarización de los procesos editoriales que publican los resultados de las investigaciones en medicina (BMJ, 1995).

Si bien el pronunciamiento de ICMJE, al cual parece adherirse gran parte de la comunidad académica, deja muchas claridades, también ha suscitado diversas reacciones. WAME, por ejemplo, reconoce que no existe un acuerdo universal sobre la definición de autor o colaborador, y que, por ende, las revistas deberían publicar lineamientos al respecto. La declaración de la Asociación Mundial Médica de Editores (WAME, 2019) coincide con ICMJE en cuanto que autor es aquel que ha realizado una contribución sustancial, pero, WAME se distancia al reconocer que existen algunas funciones, tales como escribir y revisar el manuscrito, que pueden variar. No todos los que participan lo hacen en todos lo cometidos. Por consiguiente, es altamente recomendable definir las funciones del autor y su correspondiente ordenamiento desde los inicios del proyecto de investigación, a fin de evitar disputas y demoras en la publicación.

WAME recomienda que el autor principal, o quien tiene las funciones de contactar con la revista, asuma la responsabilidad de indicar el tipo de participación de los diferentes autores. Sin embargo, es tarea de cada autor elaborar la descripción de su papel. Si bien cada autor es responsable de la totalidad del manuscrito y de la integridad científica del proyecto, en caso de que se requiera más información o surjan reclamaciones, cada quién responderá de acuerdo con su cometido y experticia.

Los lineamientos de WAME fueron más allá del rigor de ICMJE, introducen un poco de flexibilidad, pero dejan sin resolver cuáles son los cometidos, el grado de participación requerido, y bajo qué criterios se puede considerar qué se trata de una "substantial contribución".

4.4 El Instituto Nacional de Salud (National Institute of Health, NIH)

El NIH (2020) es la agencia encargada de dirigir y promover la investigación en asuntos médicos de los Estados Unidos, y forma parte del Departamento de Salud y Servicios Humanos. Además, se encarga de proveer recursos para investigación, supervisar la calidad de las publicaciones y las consideraciones éticas que esto implica.

Con respecto al tema de las autorías, su enfoque pretende dar mayor objetividad. NIH identifica cinco grandes actividades que requieren ser realizadas para la producción de una publicación, resolviendo con eso el problema de los roles requeridos. También descompone cada actividad en otras más pequeñas y señala, según el grado de realización y compromiso, si tal acción conlleva a reclamar autoría o más bien ser reconocido como colaborador.

Así, las cinco grandes actividades que componen la actividad de producción de un artículo resultado de investigación serían: 1. Diseño de la investigación e interpretación de los resultados. La sola generación de la idea no amerita autoría, a menos que haya un total involucramiento en el proyecto. 2. Supervisión del proyecto. Supone un acompañamiento del desarrollo de la investigación en todas sus fases, excluye ser autor cuando esta supervisión corresponde a tareas de formación, entrenamiento, mentorías o tutorías. 3. Soporte técnico o administrativo. Se realizan actividades de manejo de recursos financieros, materiales, animales o personas objeto de estudio. Este tipo de contribución es merecedora de reconocimiento, mas no es constitutiva de autoría. 4. Recolección y manejo de datos. Tiene especial relevancia cuando se trata de un trabajo experimental y original. Cuando se trata de un trabajo técnico califica como autoría cuando se incorporan nuevos métodos, lecturas sofisticadas

de imágenes o análisis estadísticos. 5. Escritura del manuscrito. La redacción del manuscrito final es un requerimiento esencial para ser reconocido como autor, no así la sola lectura, comentarios y sugerencias, a menos que la retroalimentación suministrada requiera involucrarse con mayor profundidad con el proyecto. Las figuras jerárquicas como jefes de laboratorio, departamento o directores no deben ser listadas como autores por el solo hecho de la posición que ostentan.

La distinción de las actividades constituye un avance que permitirá plantearse el asunto de las funciones, desempeño, tiempos de dedicación y calidad de la contribución. Aspectos que serán integrados por propuestas que tendrán lugar en la Asociación Americana de Psicología (APA).

4.5 The American Psychological Association (APA)

La Asociación Americana de Psicología (APA), en contra de lo que sucede en algunos países e instituciones, considera que los tutores puedan aparecer como coautores de las tesis o trabajos de sus doctorandos. Y, con el ánimo de esclarecer el grado de autoría que los tutores podrían reclamar con relación a las disertaciones presentadas por sus estudiantes, estipuló algunos lineamientos que legitimaban tal participación (Fields, 1983): 1. El supervisor o tutor sólo puede ser segundo autor. 2. El tutor o supervisor de tesis está obligado a aparecer como segundo autor siempre y cuando éste haya sido determinante en definir las variables, contribuya notablemente en las interpretaciones o aporte bases de datos. 3. La posición de segundo autor puede ser extendida como cortesía si el tutor o supervisor ha estado demasiado involucrado en las diferentes fases de la investigación. 4. El tutor o supervisor no debe ser reconocido como autor si su papel se reduce a animar, hacer comentarios o críticas, ofrecer ayuda financiera o realizar asistencia editorial.

A diferencia de las organizaciones que establecen principios o lineamientos, Winston propone un ingenioso procedimiento para determinar, con mayor objetividad, quiénes pueden reclamar autorías (Winston, 1985). Para ello desglosó todo el proceso de investigación, incluyendo la publicación, en diversas tareas, que van desde la concepción de la idea de investigación, su diseño, revisión bibliográfica, recolección de datos, hasta terminar con la escritura y aprobación final del manuscrito. La propuesta de Winston fue adoptada por la Asociación Americana de Psicología (APA Science Student Council, 2006) y ha servido de inspiración para crear tablas más rigurosas que permiten resolver disputas cuando los resultados iniciales son

Tabla IV. Valores para determinar autoría

Tipo de actividad	Puntos totales	Puntaje de los autores			
		Iniciales de autores			
Conceptualización de la idea de investigación	90				
Definición o cristalización de la idea de investigación	60				
Revisión de literatura (resúmenes de artículos, libros, conferencias)	20				
Creación del diseño de la investigación	80				
Selección y diseño de instrumentos	30				
Selección de encuestas e instrumentos de análisis	40				
Realización de análisis de estadísticas (incluye trabajo computacional)	40				
Interpretación de estadísticas	80				
Manuscrito					
Escritura de la introducción	90				
Escritura de la sección del método	80				
Escritura de la sección de resultados	80				
Escritura de sección de discusión	100				
Escritura de conclusiones	60				
Escritura de limitaciones del estudio	60				
Escritura sobre futuras dirección del estudio	60				
Manejo del proceso de postulación y publicación					
Respuesta a las observaciones de los evaluadores	10				
Realizar cambios de acuerdo con los comentarios recibidos	60				
TOTAL					

Fuente: Belwalkar y Toaddy, 2014.

muy parecidos. Arriba se muestra la Tabla IV, ajustada en el 2014 por miembros de APA (Belwalkar y Toaddy, 2014).

La lista de chequeo ha sido diseñada para ayudar a los colaboradores a decidir si ellos merecen ser reconocidos como autores en una publicación y determinar su orden según el puntaje obtenido. Todos los investigadores deben completar la rejilla en la medida que el proyecto avanza, los puntos totales de los autores por cada actividad deben ser igual al número propuesto en cada una. Al final, el orden de los autores se distribuye de acuerdo con los puntajes obtenidos, de mayor a menor, y quienes acumulen menos de 50 puntos, serán listados en la sección de reconocimientos.

Inspirados en fuentes diversas, tales como el Comité de Editores de Revistas Médicas (ICMJE), el Comité Ético de Publicaciones (CPE), el Instituto Nacional de Salud (NIH) y la Conferencia Internacional sobre Harmonización (ICH), Phillippi y su equipo (Phillippi, y otros, 2018) plantean 3 propuestas, una para proyectos de corte cuantitativo, otra para

cualitativos y una más para síntesis de literatura. Las dos primeras comparten 8 ítems a considerar: 1. Responsabilidad. 2. Coordinación y Comunicación. 3. Protección de los participantes. 4. Diseño. 5. Datos. 6. Análisis. 7. Escritura. 8. Aprobación final del manuscrito y sometimiento a publicación. Por otra parte, los proyectos de revisión de literatura no consideran los aspectos experimentales y se centran en otras actividades: 1. Responsabilidades. 2. Coordinación y Comunicación. 3. Revisión del diseño y búsqueda de literatura. 4. Fuentes sometidas a revisión. 5. Evaluación de datos y síntesis. 6. Escritura del manuscrito. 7. Aprobación final. Para cada ítem o actividad se definen los roles de cada participante, poniendo especial atención en aquello que corresponde al autor principal, al segundo en la lista y al investigador senior. Téngase en cuenta que estas propuestas han sido elaboradas para resolver posibles disputas entre autores de tesis y tutores en el campo de la Psicología.

Los autores reconocen la conveniencia de su propuesta para resolver disputas menores, sin em-

bargo, cuando se trata de múltiples autores que exceden un equipo de 4, esta no parece ser funcional. Por otra parte, tampoco existen estudios que permitan validar que tales propuestas han sido eficaces.

Los intentos por llegar a criterios más objetivos, aunque han sido de gran ayuda, no resuelven totalmente el problema. La cuestión de quién puede ser reconocido como autor es, ante todo, un problema ético que tiene que ser resuelto por los participantes, para lo cual es necesario volver a los lineamientos o principios generales, como lo sugiere el grupo VELiM (SHE)

4.6 El Centro por los Valores, la Ética y la Ley en Medicina, Universidad de Sydney - VELiM(SHE)

Fundado en 1995, El Centro por los Valores, la Ética y la Ley en Medicina VELiM, cambió luego su nombre por Sydney Health Ethics (SHE), y se ha convertido en un referente internacional en políticas éticas para la investigación científica (Sydney University, 2020). Este centro realizó otro gran esfuerzo por incluir en el debate las dificultades experimentadas, y propuso una guía para establecer autorías mucho más comprensivas, con 9 principios reguladores:

- Integridad y transparencia son los principios de la actividad científica, lo que supone un compromiso de trato justo y respetuoso entre los académicos.
- La publicación es una responsabilidad inherente a la actividad científica. El conocimiento y sus beneficios deben ser de dominio público.
- La responsabilidad es compartida por los equipos de trabajo de acuerdo con grados (senior o junior), experiencia y cometidos.
- Las discusiones y acuerdos sobre las expectativas en autorías deben plantearse desde el comienzo de un proyecto y se deben dejar por escrito.
- Cada persona incorporada en la lista de autores debe tener suficiente justificación para estar allí.
- Toda publicación es liderada por un autor principal
- El papel de autor principal es distribuido entre los diferentes productos que un equipo de investigación puede generar.
- El autor principal lidera la redacción del manuscrito y asume las tareas propias del proceso de gestión editorial.

Siendo este un asunto ético, que valora el grado de contribución de los participantes de una publicación, es menester que desde el inicio del proyecto se establezcan las responsabilidades en el desarrollo de la investigación, los compromisos en

la publicación y el orden en que aparecerán los autores. Es dentro del equipo de trabajo donde debe ejercerse la autorregulación. El primer principio constituye la fuente de todos los demás. Si la equidad, el trato justo y el respeto están en la base del trabajo científico, ninguno debería ostentar lo que no merece ni reclamar aquello que no ha trabajado, ni mucho menos, tomar para sí el crédito que le corresponde a otro.

5. DISCUSIÓN Y CONCLUSIONES

La manera de dar crédito por autorías está sujeta a múltiples convenciones existentes entre culturas (Fetters y Elwym, 1997), instituciones, grupos de investigación, disciplinas, asociaciones académicas y regulaciones gubernamentales (Zuckerman, 1968). En algunos contextos los supervisores o tutores raramente aparecen en los artículos, pero en otros casos, es una norma que los supervisores o jefes de departamentos aparezcan encabezando la lista de autores. En algunas disciplinas o revistas los autores son presentados obligatoriamente por orden alfabético (Over y Smallman, 1970). Los criterios de orden son muy variados, investigadores seniors pueden reservarse el primer lugar pero, a veces, el último. La primera posición, a veces, designa un investigador junior, o quien ha dedicado más tiempo, o quien haya realizado aportes más significativos (Gaeta, 1999). La editorial Nature, por su parte, contempla la autoría de los consorcios de investigación (Nature Research, 2020), para tal efecto se asumen como autores a todos los miembros del consorcio, y en caso de que algunos miembros no estén dispuestos a compartir la responsabilidad de lo publicado, deben listarse separadamente quiénes sí y quiénes no. Tal tipo de autorías extremadamente múltiples requieren la previa consulta con el editor en jefe. En general, no existen criterios universales comúnmente aceptados.

Las autorías múltiples no son condenables y parecen ser el futuro de las ciencias, sin embargo, el manejo de ellas puede ser problemático. Recientes estudios han identificado autores excesivamente prolíficos, al punto de publicar más de 72 artículos por año, una tasa asombrosa de un documento cada 5 días. El científico de materiales, Akihisa Inoue, constituye un caso excepcional. Desde 1976 reporta 2.566 artículos en Scopus, un artículo por semana por 43 años ininterrumpidos (Loannidis, J., Klavans, R. y Boyack, K., 2018). El número de artículos publicados por persona en las Ciencias Médicas puede sobrepasar los 600, mientras que en Física y Astronomía esa cifra puede llegar a 1000 (Schmidt, 2017). Al parecer la condición de ser un autor hiperprolífico depende de la discipli-

na (Física y Ciencias Médicas), instituciones (Erasmus University Rotherdam, Harvard University) y consorcios. Así, por ejemplo, el proyecto europeo ATLAS, sobre el Bosón de Higgs, está conformado por un equipo de más de 1000 integrantes donde todos son listados como autores. La condición de autor constituye un atributo de su membresía. El incremento de publicaciones con más de mil autores, por artículo, en física (Mallapaty, 2018) es una fuerte tendencia que cuestiona las prácticas de ponderar a un autor o a sus publicaciones separadamente, fuera de contexto, e invita a considerar otras opciones de valoración (Pritychenko, 2015).

La definición sobre quién puede llamarse en propiedad autor seguirá siendo un asunto complejo y abierto a discusión. En esencia, éticamente la cuestión es dar reconocimiento a quien se lo merece. Pero, además, tener la claridad sobre quién ha de responder por la información recopilada, su análisis e implicaciones. Asunto aún más difícil cuando las publicaciones con múltiples autores seguirán en aumento.

Además de las organizaciones ya comentadas, vale la pena mencionar que a nivel europeo y dentro de España, han existido pronunciamientos que, comparados con lo ya expuesto, requieren mayor refinamiento y precisión. Así, de acuerdo con el Código de Conducta Europeo para la Integridad en la Investigación (All European Academies (ALLEA), 2017), no necesariamente todos los autores responden por la totalidad de la publicación, pero todos deben cumplir con la condición de realizar aportes substanciales en el diseño de la investigación, la recolección de los datos, el análisis y su interpretación. Como puede apreciarse, el enunciado es ambiguo e incompleto. Ambiguo porque induce a pensar que las dos primeras acciones son condicionantes necesarias y las dos últimas optativas. Por otra parte, funciones como la escritura del manuscrito o su aprobación final han sido omitidas en la formulación. También, el Consejo Superior de Investigaciones Científicas (CSIC), (2016) de España, ha establecido algunos lineamientos para determinar las autorías. Sin embargo, su acento ha estado más en el aspecto de la responsabilidad. Todos los firmantes asumen la responsabilidad de la publicación y queda a su discreción revelar el tipo de contribución de cada uno. Quien actúe como interlocutor ante el editor tiene el deber de listar a "quienes hayan contribuido significativamente en la concepción, planificación diseño, ejecución, obtención de datos, interpretación y discusión de los resultados del trabajo". No queda claro el significado de "contribución significativa", ni si basta participar en una fase del proceso o en todas. (Consejo Superior de Investigaciones Científicas, 2016).

En medio de tantos ensayos y enfoques, la exigencia por atribuir créditos donde es debido, ha conducido a propuestas innovadoras que otorgan mayor transparencia en las publicaciones con autorías múltiples. Sin duda, algunas líneas de trabajo previas han sido redefinidas y combinadas por el Consorcio para el Avance de la Administración de la Investigación (Consortia Advancing Standards in Research Administration, CASRAI). CASRAI ha distinguido 14 funciones posibles que puede desempeñar un autor: conceptualización, curaduría de datos, análisis, adquisición de fondos, investigación, metodología, administración del proyecto, recursos, software, supervisión, escritura del documento original, revisión de correcciones y edición. Además de distinguir los roles, CASRAI también ha desarrollado una herramienta que permite calificar el grado de participación de cada uno: Contributor Roles Taxonomy, CRediT. (CASRAI, 2019). Sopor-tada por un grupo reconocido de universidades y casas editoriales, CRediT está siendo difundida ampliamente y sería deseable que tal iniciativa fuese adoptada como una buena práctica editorial (McNutt y otros, 2018), más aún, cuando los datos de CRediT son compatibles con ORCID y Crossref. Aunque las ventajas de adhesión a CRediT son innegables, esto por sí mismo no es suficiente, se requiere la confluencia de los diferentes grupos de interés a fin de lograr la justa ponderación y reconocimiento a las contribuciones que cada autor hace, no sólo en un documento, sino a las ciencias y al conocimiento en general.

Con el ánimo de contribuir a la construcción del concepto de autoría y como conclusión de este trabajo se propone el siguiente enunciado:

Todo aquel que haya participado conjuntamente en el proceso de producción de una publicación resultado de investigación, puede reclamar ser reconocido como autor si cumple con las siguientes condiciones:

- *Desde el inicio del proyecto o desde el momento de su vinculación, se definió por escrito su cometido y sus resultados esperados,*
- *El equipo de trabajo considera que dentro de la ejecución del proyecto realizó contribuciones substanciales de calidad y tiempo en al menos tres de las siguientes actividades:*
- *En la concepción de la idea original, o en el refinamiento y reformulación de dicha idea, o en el diseño de la investigación, o en la revisión de la literatura, o en la construcción de herramientas, o en el manejo de los grupos de estudio o trabajos de laboratorio, o en la recolección y organización de los datos, o en el análisis e interpretación de los datos, o en la construcción*

de aplicativos o software usados dentro del proyecto, o en la redacción del manuscrito, o en las correcciones y aportes finales del documento a publicar.

- Y, además, leyó y aprobó la versión final del documento
- Y asume la responsabilidad, tanto por la parte que le corresponde como por el contenido total de la publicación.

Toda persona que participe en alguna de estas actividades ha de brindársele la oportunidad de participar en las otras. Aquel que, habiendo participado en el proyecto, pero que no califica como autor, debe ser mencionado en la sección de reconocimientos.

Gran parte de esta propuesta contempla la estructura de ICMJE, pero en su contenido está más cerca de WAME. La lista de actividades se encuentra inspirada en las propuestas de APA, al igual que los conceptos de calidad y tiempo. A diferencia de otros enunciados, se le da importancia al juicio del grupo de investigación sobre los aportes de cada investigador. La cifra de un mínimo de actividades asegura el mayor involucramiento de los participantes. Aunque esta versión se acerca a la presentada por McNutt y otros (2018) se han dejado resaltadas en *italicas* las diferencias.

Por otra parte, una apropiada valoración, créditos y responsabilidades, derivadas de publicaciones con autorías múltiples, depende de los diferentes grupos de interés que intervienen en la publicación de un artículo resultado de investigación. La primera tarea, por tanto, es diferenciar estos grupos, y después determinar su tipo de responsabilidad.

Entre los grupos de interés se proponen los siguientes: las agencias financiadoras, las instituciones promotoras de las publicaciones, las revistas, los editores en jefe y los autores.

- *Las agencias financiadoras* deben concentrar la valoración de las propuestas de investigación por la calidad y alcance de estas, y no por la fama ni las métricas de quiénes las presentan. De igual manera, deben establecer criterios diferenciados entre disciplinas, instituciones e investigadores, no todos los agentes tienen los mismos criterios para atribuir o reconocer autorías.
- *De las instituciones promotoras de publicaciones* (casas editoras, universidad y centros) ha de esperarse su compromiso ético con la promoción y defensa de la integridad científica, expresada en filiaciones a comités de ética internacional, como lo puede ser el Committee

on Publication Ethics (COPE, 2019), y la adhesión a principios centrados más en la calidad de los contenidos que en las figuras académicas o métricas. Es altamente recomendable que las instituciones promotoras de revistas científicas se suscriban a la Declaración de San Francisco (San Francisco Declaration on Research Assessment, 2019) y el manifiesto de Leiden (Leiden Manifest for Research Metrics, 2015). Estas declaraciones surgieron por iniciativa de editores e instituciones académicas que invitan a valorar los artículos más por su contenido que por el índice h de sus autores. Esta actitud, sin duda, desincentiva la proliferación de autorías fraudulentas.

- *Por otra parte, las editoriales, o las revistas mismas* deben asegurarse de expresar públicamente cuáles son sus políticas y lineamientos para establecer las autorías, así como los procedimientos para dirimir los conflictos que al respecto puedan suscitarse. Es también importante, en aras de la transparencia y calidad de la información, demandar a los autores su respectivo código ORCID y perfil en Google Scholar, registros que permiten tener la trazabilidad pública de su producción científica. Y como se ha señalado arriba, la adhesión CRediT que introduce la especificación de los roles o actividades realizadas por cada investigador.
- *A nivel de los editores* conviene aclarar que no es su responsabilidad determinar quién puede ser autor o no, en tanto que el editor actúa impulsado por el principio de buena fe. Sin embargo, sí es de su competencia proporcionar la información necesaria para distinguir entre funciones de autoría o de colaboración. Además, el editor, basado en su experiencia, debe reaccionar e indagar ante los primeros signos de alerta (COPE, 2018):
 - Aparecen listados autores de la compañía que financia el estudio más no de quien lo realiza.
 - Los nombres de los autores no son familiares a su área de experticia, desempeño o línea de investigación.
 - No se especifican claramente las contribuciones de los autores.
 - Correspondencia del artículo con el número de autores. Así, son sospechosos los proyectos de alta complejidad con pocos autores o, por el contrario, muy simples con una docena de titulares.
 - Artículos con títulos similares publicados con diferentes nombres o alianzas.
 - Cambio de autores, eliminación o adición dentro del proceso de publicación.
 - Existencia de autores prolíficos, excesivas e inverosímiles publicaciones por año.

- Redacción o revisión final del manuscrito por alguien ajeno a los firmantes.
- El autor de contacto no responde con claridad ante preguntas u observaciones sobre el manuscrito.
- El estilo y lenguaje de la carta de presentación difiere ampliamente del manuscrito.

Ante cualquier alerta se debe proceder según un plan de tratamiento de casos como lo sugiere el Committee on Publication Ethics (COPE, 2006): suplir información relevante sobre políticas de autoría, enviar formato de declaración de autor, sugerir las posibilidades de casos que se pueden presentar, contactar con autores listados y manifestar por escrito observaciones y tratamiento a seguir. Finalmente, tomar acciones correctivas para evitar que se repitan futuros casos.

Por parte de los autores su compromiso con la ética y la integridad científica es aún mayor. La sociedad, en general, tiene el derecho de ser informada de manera veraz sobre quiénes realmente están haciendo aportes sustanciales en el avance

de las ciencias y quienes están en capacidad de responder auténticamente por la recolección de los datos, los hallazgos y sus implicaciones. Es conveniente, por tanto, que desde el inicio de un proyecto de investigación se definan los aportes de cada uno de los participantes, el tipo de contribución en la redacción del texto final, quienes podrán reclamar autorías y en qué orden serán listados. Igualmente, dejar claro quiénes merecen aparecer en la sección de reconocimientos. Todos estos acuerdos deben constar por escrito en un acta firmada por los participantes.

Debido a que el ejercicio de producción de conocimiento difiere entre las diferentes disciplinas y que la incorporación de nuevas tecnologías hace el trabajo mucho más diverso, no puede ser valorado igual quien participa de un consorcio con miles de integrantes, produciendo un documento por semana, que un equipo de tres que logran encontrar la cura de una enfermedad o un aislado filósofo quien trata de comprender nuevas formas de contrato social. Sin duda, todos producen conocimiento científico.

6. REFERENCIAS

- Aad, G. et al. (2015). A measurement of the Higgs boson mass is presented based on the combined data samples of the ATLAS. *Physical Review Letters*, 114(9), 1-33. <https://doi.org/10.1103/PhysRevLett.114.191803>
- All European Academies (ALLEA). (2017). European Code of Conduct for Research Integrity. Disponible en: <https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf> [Fecha de Consulta: 12/4/2020].
- Angell, M. (1994). Publish or perish: a proposal. *Annals of Internal Medicine*, 104(2), 261-262. <https://doi.org/10.7326/0003-4819-104-2-261>
- APA Science Student Council. (2006). *A Graduate Student's Guide to Determining Authorship Credit*. Disponible en: <https://www.apa.org/science/leadership/students/authorship-paper.pdf> [Fecha de consulta: 17/7/2019].
- Baethge, C. (2008). Publish Together or Perish. *Deutsches Ärzteblatt International*, 105(20), 380-383. <https://doi.org/10.3238/arztebl.2008.0380>
- Bailar, J. (1986). Science, Statistics and Deception. *Annals of Internal Medicine*, 104(2), 259 - 260. <https://doi.org/10.7326/0003-4819-104-2-259>
- Belwalkar, B. y Toaddy, S. (2014). Authorship Determination Scorecard. Disponible en: <https://www.apa.org/science/leadership/students/authorship-determination-scorecard.pdf> [Fecha de consulta: 14/6/2019].
- Bemmet, D. y Taylor, D. (2003). Unethical practices in authorship of scientific papers. *Emergency Medicine*, 15(3), 263-270. <https://doi.org/10.1046/j.1442-2026.2003.00432.x>
- Bhaskar, E. (2010). Guest authorship: the practice of author pooling. *Open Medicina*, 4(1). Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3116671/> [Fecha de consulta: 16/05/2020]
- BigTenCRC – Big Cancer Research Consortium (2019). Clinical Research. Disponible en: <https://www.bigtenrc.org/clinical-research> [Fecha de consulta: 16/6/2019].
- BMJ. (1995). World Association of Medical Editors (WAME) launched. 310(761) <https://doi.org/10.1136/bmj.310.6982.761a>
- Broad, W. J. (1983). Notorious Darsee case shakes assumptions about science. *New York Times*. Disponible en: <https://www.nytimes.com/1983/06/14/science/notorious-darsee-case-shakes-assumptions-about-science.html> [Fecha de consulta: 16/6/2020].
- CASRAI. (2019). CRediT – Contributor Roles Taxonomy. Disponible en: <https://casrai.org/credit/> [Fecha de consulta: 10/4/2020].
- Consejo Superior de Investigaciones Científicas. (2016). Guía de buenas prácticas de las publicaciones periódicas y unitarias de la Agencia Estatal Consejo Superior de Investigaciones Científicas. Disponible en: http://revistas.csic.es/public/guia_buenas_practicas_CSIC.pdf [Fecha de consulta 15/4/2020].
- COPE - Committee on Publications Ethics - (2006). What to do if you suspect ghost, guest or gift authorship. Disponible en: <https://publicationethics.org/files/Ghost.pdf> [Fecha de consulta: 18/4/2020].
- COPE - Committee on Publications Ethics. (2018). How to recognize potential authorship problems. Disponible en: [Fecha de consulta: 18/4/2020]
- COPE - Committee on Publications Ethics (2019). Principles of Transparency and Best Practice on Scholarly Publishing. Disponible en: <https://publicationethics.org/node/19881> www.publicationethics.org [Fecha de consulta 19/4/2020].

- CPRE – Consortium for Policy Research in Education (2020). Overview. Disponible en: <https://www.cpre.org/>; <https://www.cpre.org/overview> [Fecha de consulta: 18/6/2020].
- CSE – Council Science Editors (2012). White Paper on Publication Ethics. Disponible en: <https://www.councilscienceeditors.org/resource-library/editorial-policies/white-paper-on-publication-ethics/> [Fecha de consulta: 13/6/2019].
- CSE – Council Science Editors. (2020). History of CSE. Disponible en: <https://www.councilscienceeditors.org/about/history-of-cse/> [Fecha de consulta: 8/4/2020].
- Culliton, B. (1983). Coping with fraud: the Darsee Case. *Science*, 220(4592), 31-35. <https://doi.org/10.1126/science.6828878>
- Da Silva, J. y Dobráski, J. (2016). How Authorship is DeUned by Multiple Publishing Organizations and STM Publishers. 23(2), 97 - 122. <https://doi.org/10.1080/0898621.2015.1047927>
- Da Silva, T. (2015). Multiple Authorship in Scientific Manuscripts: Ethical. *Science and Engineering Ethics*, 1457-1472. <https://doi.org/10.1007/s11948-015-9716-3>
- Das, K. V., Patil, S., Ray, J., Deepak, K. (2016). Knowledge on ethical authorship: A comparative study between medical and pharmacy faculty. *Journal of Youth Pharmacists*, 8(2), 136-140. <https://doi.org/10.5530/jyp.2016.2.15>
- ELSEVIER (2019). Factsheet: Authorship. Disponible en: https://www.elsevier.com/__data/assets/pdf_file/0006/653883/Authorship-factsheet-March-2019.pdf [Fecha de consulta 8/4/2020].
- Fetters, M., y Elwym, T. (1997). Authorsip. Assesment of authorship depend on culture. *BMJ*, 315(7110), 747-747. <https://doi.org/10.1136/bmj.315.7110.744>
- Fields, C. (1983). Professors' demands for credit as coauthors of students' research projects may be rising. *Chronicle of Higher Education*, págs. 7,10.
- Fletcher, R. (2005). Ghost writing initiated by commercial companies. *Journal of General Internal Medicine*, 20(6), 549-549. <https://doi.org/10.1111/j.1525-1497.2005.41015.x>
- Gaeta, T. (1999). Authorship: "Law and oder". *Academic Emergency Medicine*, 6(4), 297-301. <https://doi.org/10.1111/j.1553-2712.1999.tb00393.x>
- Harvey, L. (2018). Gift, honorary or guest authorship. *Spinal Cord*, 56(9). <https://doi.org/10.1038/s41393-017-0057-8>
- Hicks, D., Wouters, P., Waltman, L., Sarah, D., y Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 429-431. <https://doi.org/10.1038/520429a>
- Hosseini, M. y Gordijn, B. (2020). A review of the literature on ethical issues related to scientific authorship, *Accountability in Research*, <https://doi.org/10.1080/0898621.2020.1750957>
- ICMJE – International Committee of Medical Journal Editors (2019). Defining the Role of Authors and Contributors. Disponible en: <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html> [Fecha de consulta: 2/2/2019].
- ICMJE – International Committee of Medical Journal Editors (2019). Who is an Author? Disponible en: <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#two> [Fecha de consulta: 2/2/2019].
- ICMJE. (2020). *International Committe of Medical Journal Editors*. About ICMJE. Disponible en: <http://www.icmje.org/about-icmje/> [Fecha de consulta: 8/4/2020].
- Koseoglu, M. (2016). Growth and structure of authorship and co-authorship network in the strategic management realm: Evidence from the Strategic Management Journal. *Business Research Quarterly*, 19(3), 153-170. <https://doi.org/10.1016/j.brq.2016.02.001>
- Kuld, L. y O'Hagan, J. (2017). The trend of increasing co-authorship in economics: New evidence. *VOX*. Disponible en: <https://voxeu.org/article/growth-multi-authored-journal-articles-economics> [Fecha de consulta: 20/5/2020]
- Leiden Manifesto for Research Metrics. (2015). Disponible en: <http://www.leidenmanifesto.org/> [Fecha de consulta 20/5/2019].
- Loannidis, J., Klavans, R. y Boyack, K. (2018). Thousands of scientists publish a paper every five days. *Nature* (561), 167-169. <https://doi.org/10.1038/d41586-018-06185-8>
- Mallapaty, S. (2018). *Paper authorship goes hyper*. Nature Index, Disponible en: www.natureindex.com/news-blog/paper-authorship-goes-hyper [Fecha de consulta: 18/6/2019].
- McNutt, M., Bradford M., Drazan, J., Hadson, B., Howard, B. (2018). Trasnparency in authors' contribucions and responsibilities to promote integrity in scientificc publication. *National Academy of Sciences*, 115(11), 2557-2560. <https://doi.org/10.1073/pnas.1715374115>
- Michaela, P. (2016). Credit where credit is due: best practices for authorship attribution. Disponible en: https://www.aje.com/dist/docs/Authorship_Attribution_EN.pdf [Fecha de consulta: 15/4/2021].
- Misra, D. P. (2018). Integrity of Authorship and Peer Review Practices: Challenges and Opportunities for Improvement. *Journal of Korean Medical Science*, 33(46), e287. <https://doi.org/10.3346/jkms.2018.33.e287>
- National Human Genome Research Institute. (2019). The Human Research Project. Disponible en: <https://www.genome.gov/human-genome-project> [Fecha de consulta: 15/3/2019].
- Nature (2020). Authorship. Disponible en: <https://www.nature.com/nature-research/editorial-policies/authorship> [Fecha de consulta: 10/4/2020].
- NIH - National Institute of Health. (2020). Who we are. Disponible en: www.nih.gov; <https://www.nih.gov/about-nih/who-we-are> [Fecha de consulta: 16/4/2020].
- Over, R. y Smallman, S. (1970). Citation Idiosyncrasies. *Nature*, 228(5278), 1357-1357. <https://doi.org/10.1038/2281357a0>
- PerRC – Pediatric Research Consortium (2019). Current Projects. Disponible en: <https://perc.research.chop.edu/?q=listing> [fecha de consulta: 15/3/2019].
- Phillippi, J., Likis, F., & Tilden, E. (2018). Authorship Grids: Practical tools to facilitate coallaboration and ethical publication. *Research in Nursing & Health*, 41(2), 195-208. <https://doi.org/10.1002/nur.21856>

- PlosOne. (2019). Authorship. Disponible en: <https://journals.plos.org/plosone/s/authorship> [Fecha de consulta: 11/4/2020].
- Pritychenko, B. (2015). Fractional authorship in nuclear physics. *106*(1), 461-468. <https://doi.org/10.1007/s11192-015-1766-4>
- Rennie, D., Yank, V., Emmanuel, L. (1997). When authorship fail. A proposal to make contributors accountable. *JAMA*, 278(7), 579-585. <https://doi.org/10.1001/jama.1997.03550070071041>
- Ruiz-Pérez, R., Marcos-Cartanega, D., Delgado López-Cozar, E. (2014). La autoría científica en las áreas de ciencia y tecnología. Políticas internacionales y prácticas editoriales en las revistas científicas españolas. *Revista Española de Documentación Científica*, 37(2) <https://doi.org/10.3989/redc.2014.2.1113>
- San Francisco Declaration on Research Assessment. (2019). Disponible en: www.sfdora.org [Fecha de consulta: 4/5/2019]
- Schmidt, M., Fecher, B. y Kobsda, C. (2017). How many authors does it really need to write a paper? Elephant in the lab. Zenodo. <https://doi.org/10.5281/zenodo.818231>
- Sismondo, S. (2005). Ghost Management: How Much of the Medical Literature Is Shaped Behind the Scenes by the Pharmaceutical Industry? *PLoS Med*, 4(9). <https://doi.org/10.1371/journal.pmed.0040286>
- Smith, J. (1994). Gift authorship: a poisoned chalice? (J. Smith, Ed.) *The BJM*, 309, 1456-1457. <https://doi.org/10.1136/bmj.309.6967.1456>
- Sydney University. (2020). Research Publications and Outputs. Disponible en: <https://ses.library.usyd.edu.au/handle/2123/10565> [Fecha de consulta: 9/4/2020]
- Taylor and Francis Group. (2017). Co-Authorship in the humanities and Social Sciences. A Global View. Disponible en: <https://authorservices.taylorandfrancis.com/wp-content/uploads/2017/09/Coauthorship-white-paper.pdf> [Fecha de consulta: 15/4/2020].
- The Center for Values, Ethics and the Law in Medicine -VELIM. (2011). *Guidelines on Collaborative Research and Authorship*. Disponible en: <https://ses.library.usyd.edu.au/handle/2123/13253> [Fecha de consulta: 10/6/2019]
- The PLoS Medicine Editors. (2009). Ghostwriting: The Dirty Little Secret of Medical Publishing That Just Got Bigger. *Plos Med*, 6(9). e1000156. <https://doi.org/10.1371/journal.pmed.1000156>
- UKPC. (2019). What UKPC Does. Disponible en: <https://www.ukcp.ac.uk/pmwiki.php/UKCP/WhatUKCPDoes> [Fecha de consulta: 15/3/2019]
- WAME – The World Association of Medical Editors (2019). Recommendations on Publication Ethic Policies for Medical Journals. Disponible en: <http://wame.org/recommendations-on-publication-ethics-policies-for-medical-journals> [Fecha de consulta 16/6/2020]
- Winston, R. B. (1985). A suggested Procedure for Determining Order of Authorship Research in Publications, *Journal of Counseling and Development*, 63, 515 - 518. <https://doi.org/10.1002/j.1556-6676.1985.tb02749.x>
- Zuckerman, H. (1968). Patterns of Name Ordering Among Author of Scientific Papers: A Study of Social Symbolism and Its Ambiguity. *American Journal of Sociology*, 74(3), 276-291. <https://doi.org/10.1086/224641>

ESTUDIOS / RESEARCH STUDIES

Producción científica sobre ansiedad bibliotecaria: un análisis bibliométrico y cienciométrico desde Scopus

Pier Suclupe-Navarro*, Cesar H. Limaymanta*,**, Néstor Holmes Ramírez***, Héctor Guillén*

* Universidad Nacional Mayor de San Marcos. Perú

Correo-e: pier.suclupe@unmsm.edu.pe | ORCID iD: <https://orcid.org/0000-0001-7232-2103>

Correo-e: climaymanta@unmsm.edu.pe | ORCID-ID: <http://orcid.org/0000-0002-8797-4275>

Correo-e: hector.guillen1@unmsm.edu.pe | ORCID iD: <https://orcid.org/0000-0002-2644-8167>

**Universidad Peruana de Ciencias Aplicadas (UPC). Perú

Correo-e: pcmaclim@upc.edu.pe | ORCID-ID: <http://orcid.org/0000-0002-8797-4275>

***Biblioteca de la Universidad de Morelos. México

Correo-e: nestor@um.edu.mx | ORCID iD: <https://orcid.org/0000-0001-5597-4513>

Recibido: 29-01-20; 2ª versión: 23-04-20; Aceptado: 27-04-20; Publicado: 29-04-2021

Cómo citar este artículo/Citation: Suclupe-Navarro, P.; Limaymanta, C.H.; Ramírez, N.H.; Guillén, H. (2021) Producción científica sobre ansiedad bibliotecaria: un análisis bibliométrico y cienciométrico desde Scopus. *Revista Española de Documentación Científica*, 44 (2), e291. <https://doi.org/10.3989/redc.2021.2.1753>

Resumen: La ansiedad bibliotecaria (AB) es el estado de incertidumbre que experimentan algunos usuarios cuando acuden a la biblioteca, cuyo estudio comenzó en los años 80, pero que carece de un análisis métrico actualizado. El objetivo de este estudio fue hacer un análisis bibliométrico y cienciométrico de los artículos sobre AB indizados en la base de datos Scopus entre 1989 y 2018. Se analizaron 118 publicaciones según indicadores de producción, colaboración, impacto y mapas bibliométricos. Se encontró que las publicaciones sobre AB han aumentado progresivamente, que la colaboración de autores fue baja, que una revista influyente en el área fue *Library Review*, y que los autores con más impacto fueron Onwuegbuzie y Jiao. Asimismo, destacaron palabras clave sobre la formación del usuario y cuatro frentes de investigación: instrumentos de medición, resultados, marcos teóricos y tipos de unidades de análisis. Dichos resultados sugieren que la AB es un tema fructífero de investigación.

Palabras clave: ansiedad bibliotecaria; Bibliometría; Cienciométrica; usuarios de información; Scopus; VOSviewer; emparejamiento bibliográfico; cocitación.

Scientific production on library anxiety: a bibliometric and scientometric analysis from Scopus

Abstract: Library anxiety (LA) is the state of uncertainty experienced by some users when they come to the library, the study of which began in the 1980s but lacks updated metric analysis. The objective of this study was to make a bibliometric and scientometric analysis of the articles on LA indexed in the Scopus database between 1989 and 2018. 118 publications were analyzed according to indicators of production, collaboration, impact and bibliometric maps. It was found that publications on LA have progressively increased, that the collaboration of authors was low, that an influential journal in the area was *Library Review*, and that the authors with the greater impact were Onwuegbuzie and Jiao. Also, keywords on user training and four research fronts were highlighted: *measurement instruments, results, theoretical frameworks and types of analysis units*. These results suggest that LA is a fruitful research topic.

Keywords: library anxiety; Bibliometrics; Scientometrics; information user; Scopus; VOSviewer; bibliographic coupling; co-citation.

Copyright: © 2021 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1. INTRODUCCIÓN

1.1. Ansiedad y ansiedad bibliotecaria

De forma general, y desde una perspectiva psicológica, la ansiedad se define como un estado emocional de "agitación e inquietud desagradable, caracterizado por la anticipación del peligro [...] y la sensación de catástrofe inminente" (Sierra y otros, 2003). Según Browne (2018), la ansiedad constituye un mecanismo biológico común y normal de las personas para enfrentar situaciones amenazantes. De forma más específica, la ansiedad ha sido estudiada en dos dimensiones: "como respuesta emocional y como rasgo de la personalidad" (García y Cano, 2014), denominadas ansiedad de *estado* y ansiedad de *rasgo*, respectivamente (Leal y otros, 2017). La ansiedad de rasgo es la característica estable de la personalidad de reaccionar de forma ansiosa, con independencia de la situación o circunstancia. Por el contrario, la ansiedad de estado es una reacción emocional transitoria vinculada a una situación específica (Arcas y Cano, 1999).

De acuerdo con ello, la ansiedad bibliotecaria (AB en adelante) se puede entender como el estado emocional de incertidumbre e incomodidad que experimentan los usuarios al momento de tener contacto con la biblioteca para requerir información. Mellon (1986) usó por primera vez el término y lo definió como "el sentimiento de temor" que experimentan los usuarios y "que les impide permanecer en la biblioteca", sin sacarle el máximo provecho para su formación. En este sentido, como indica Ramírez (2015), la AB es una manifestación de la ansiedad de estado. Es decir, es una manifestación transitoria en el usuario que se presenta en el contexto particular de la biblioteca y sus servicios. Onwuegbuzie y otros (2004) agregan que la AB es un "estado negativo" que se da en un continuo, donde los usuarios "que experimentan los niveles más altos [de AB] tienen un mayor riesgo de exhibir conductas de evasión" (p. 55).

Considerando lo anterior, el estudio de la AB se inscribe dentro del paradigma cognitivo en bibliotecología y ciencias de la información, donde se considera "variables afectivas, emocionales y cognitivas" (Tamayo-Rueda y Meneses-Placeres, 2018), como la actitud, las valoraciones, el comportamiento, la experiencia cognitiva del usuario, sus motivaciones y expectativas (Tamayo-Rueda y Meneses-Placeres, 2018; Izquierdo, 1999).

1.2. Investigación sobre la ansiedad bibliotecaria

Diversos estudios han buscado medir la AB para determinar los factores que hacen que se manifieste. En general, dichos estudios siguen el enfoque

cuantitativo y se centran en estudiar las percepciones de estudiantes universitarios sobre la biblioteca. Un ejemplo representativo es el trabajo de Ramírez (2015), quien agrega que es posible estudiar la AB en un contexto latinoamericano. Para medir la AB se han elaborado escalas psicométricas, como la escala de Bostick (Onwuegbuzie y otros, 2004), la escala multidimensional MLAS (Van Kampen, 2004) y la escala AQAK (Anwar y otros, 2012). Carlile (2007) hace una revisión de los trabajos sobre el tema publicados entre 1986 y 2006 y los clasifica según el enfoque de estudio, el tipo de artículo, la metodología aplicada y el tipo de escala empleada para medir el fenómeno. La autora encuentra que si bien la escala de Bostick es aceptada como confiable, los estudios no muestran ser concluyentes. Asimismo, encuentra que los autores más prolíficos son Jiao y Onwuegbuzie. Respecto a esto, Cleveland (2004), en su revisión de los principales trabajos de Jiao y Onwuegbuzie, concluye que ambos autores son claves para comprender la importancia de la AB como un fenómeno real que debe ser estudiado en el campo bibliotecológico.

1.3. Bibliometría y cienciometría

La bibliometría es "la aplicación de los métodos matemáticos y estadísticos a la cuantificación de libros y otros medios de comunicación" (Pritchard, 1969). Es decir, es el conjunto de métodos cuantitativos útiles para describir y medir la literatura académica (Roemer y Borchardt, 2015); por ejemplo, para medir la cantidad de publicaciones y las citas que estas reciben (Cortés, 2007). Respecto a lo último, el análisis bibliométrico se basa en la idea de que existe un vínculo fuerte y directo entre las citas y el contenido de los artículos citados, "lo que significa que el análisis de citas puede ser el principal instrumento para evaluar el impacto de las publicaciones científicas" (Abramo y otros, 2019).

Por otro lado, la cienciometría es "el estudio de los aspectos cuantitativos de la ciencia y la tecnología vistos como un proceso de comunicación" (Mingers y Leydesdorff, 2015). Incluye dentro de sus principales temas la cuantificación de las investigaciones y su impacto, las citaciones, el mapeo de las áreas científicas y el uso de indicadores para las políticas y gestión de la investigación. Por esta razón, existe un solapamiento entre bibliometría y cienciometría, ya que ambas incluyen aspectos cuantitativos de la ciencia como proceso de comunicación (Björneborn y Ingwersen, 2004).

En bibliometría y cienciometría se usan indicadores para describir, evaluar y hacer seguimiento a las publicaciones (Riggio, 2017). Tradicionalmente se clasifican en indicadores de producción, de impacto

Tabla I. Tipos de indicadores bibliométricos y cienciométricos

Indicador	Descripción	Ejemplo
Producción	Cuantifica las publicaciones de autores, instituciones o países, para determinar el dinamismo del campo de investigación.	Año, idioma, número y tipo del documento, ciudad de afiliación, número de instituciones, temática, índice de productividad, etc.
Colaboración	Describe y diagnostica las relaciones de colaboración entre autores, instituciones o países.	Índice de colaboración, grado de colaboración, coeficiente de colaboración, etc.
Impacto	Evalúa la relevancia de las publicaciones en función del número de citas recibidas, para determinar su influencia a corto plazo.	Cuartiles y factor de impacto JCR (Journal Citation Reports), cuartiles e índice de impacto SJR (SCImago Journal Rank), número de citas, h-Index, etc.
Basado en mapas	Identifica especialidades e interdisciplinariedad en la investigación. Mapea los ejes de la ciencia.	Mapas de emparejamiento bibliográfico, cocitación, coautoría y coocurrencia de términos en palabras clave, títulos y resúmenes.

Fuente: Riggio (2017), Peralta y otros (2015), y Yang y Ding (2012).

y de colaboración (Peralta y otros, 2015). Además de estos, se usan indicadores bibliométricos basados en mapas (Yan y Ding, 2012) (Tabla I).

Respecto a los indicadores de producción, el índice de productividad (IP en adelante) permite categorizar a los autores y se calcula con el logaritmo del número de publicaciones: si el $IP \geq 1$ el autor es un gran productor, si $0 < IP < 1$ es un productor intermedio, si el $IP = 0$ es un productor transitorio (Mohanty y otros, 2018). En el caso de los indicadores de colaboración, el índice de colaboración mide el peso promedio del número de autores que firman un documento; su valor mínimo es 1, sin tener un valor máximo determinado (Lawani, 1981). El grado de colaboración señala la proporción de documentos escritos en coautoría múltiple; su valor oscila entre 0 y 1, donde los valores próximos a 1 indican que existe mayor frecuencia de documentos con, por lo menos, dos autores (Subramanyam, 1983; Salas y otros, 2018). Por último, el coeficiente de colaboración refleja al mismo tiempo el número y la proporción de autores por documento; su valor oscila entre 0 y 1, que tiende a 0 cuando predominan los trabajos de un solo autor (Ajiferuke y otros, 1988).

Sobre los indicadores de impacto, el número total de citas es un indicador básico que mide en enteros las veces que un ítem es citado (Waltman, 2016). Así, cuantas más citas tenga una publicación, mayor será su impacto. Al ser acumulativo, este indicador tiende a incrementarse con el tiempo, a diferencia de otros indicadores, como el promedio de citas recibidas, que no es acumulativo sino "independiente del número" (Waltman, 2016). En general, se considera que la citas y la citación "reflejan, con importantes limitaciones, aspectos relacionados al impacto y a la relevancia científica" (Aksnes y otros, 2019).

Finalmente, dentro de los indicadores basados en mapas, el análisis de emparejamiento bibliográfico (AEB en adelante) permite relacionar "documentos que referencian [o citan] un mismo grupo de documentos" (Boyack y Klavans, 2010). Es decir, el AEB vincula dos ítems A y B que hacen referencia en su bibliografía a un tercer ítem C. Cuando ello sucede, se dice que A y B están emparejados bibliográficamente por C. Así, cuanto mayor es el número de referencias comunes de dos publicaciones, "significa que ambas pertenecen a la misma área [de investigación]" (Habib y Afzal, 2017). Por otro lado, contrario al AEB, el análisis de cocitación (AC en adelante) permite relacionar dos publicaciones que son citadas juntas en las referencias de una tercera publicación (Marshakova-Shaikovich, 1973; Small, 1973). De acuerdo con MacCain (1991), el AC de revistas permite "destacar la relación entre temas, especializaciones de investigación, [así como] la organización de la literatura académica en varios niveles de generalidad" (p. 295). En el caso de dos o más autores, existe cocitación si un tercero los cita en sus referencias (van Eck y Waltman, 2014). En este sentido, cuando "dos autores son citados de manera frecuente en las mismas publicaciones, su relación es fuerte" (Perianes-Rodríguez y otros, 2016).

1.4. Objetivos de la investigación

Se hizo una búsqueda en Scopus, Web of Science y Google Académico y no se encontraron trabajos métricos que midan las publicaciones sobre AB. Considerando que el tema resulta de interés por estudiar la subjetividad del usuario y que tener un panorama métrico del mismo sería útil para los investigadores interesados, el objetivo principal de este trabajo es hacer un análisis bibliométrico y cienciométrico de los artículos sobre AB indizados

en la base de datos Scopus. Los objetivos específicos son los siguientes: (1) analizar la producción según publicación quinquenal, tipo de publicación, países, idioma, áreas de investigación e índice de productividad; (2) determinar la colaboración de autores mediante el índice de colaboración, el grado de colaboración y el coeficiente de colaboración; (3) determinar el impacto y las agrupaciones de fuentes y autores mediante el análisis de citas, de emparejamiento bibliográfico y de cocitación; y (4) describir las tendencias de investigación mediante redes de coocurrencia de palabras clave y de términos de los resúmenes.

2. MÉTODO

En general, se siguió una metodología de cinco fases: (1) recopilación de datos, (2) definición de unidades de análisis, (3) definición de unidades de medida, (4) reducción de dimensionalidad y distribución de visualizaciones y (5) análisis e interpretación de la visualización (Vargas-Quesada y de Moya-Anegón, 2007). A continuación se describe las cuatro primeras fases. La última se desarrolla a lo largo del apartado *Resultados*.

En primer lugar, para buscar y recuperar los datos objeto de análisis se usó Scopus, un recurso confiable que abarca en gran medida las áreas de Ciencias Sociales, Humanidades y Artes (Elsevier, 2019). La búsqueda se hizo en noviembre de 2019 y se usó el término "library anxiety" para encontrar coincidencias en los títulos, resúmenes y palabras clave de los documentos publicados entre 1960 y 2018. El primer resultado fue de 129 ítems publicados entre 1989 y 2018. Dicho resultado se delimitó según tipo de documento (solo artículos originales y artículos de revisión) y según tipo de fuentes (solo revistas). La cadena de búsqueda final fue: TITLE-ABS-KEY ("library anxiety") AND PUBYEAR < 2019 AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re")) AND (LIMIT-TO (SRCTYPE, "j")), con la que se obtuvo 119 artículos, de los cuales se descartó uno por no tratar sobre AB, sino sobre la ansiedad en general. Los 118 artículos restantes se exportaron en un archivo .CSV según las siguientes opciones en Scopus: información de citas, información bibliográfica, resumen y palabras clave, y otra información.

En segundo lugar, se definieron cinco unidades de análisis: documentos, fuentes, autores, palabras clave y términos en los resúmenes. En tercer lugar, se establecieron seis unidades de medida: producción, colaboración, citas, emparejamiento bibliográfico, cocitación y tendencia de investigación. En cuarto lugar, para analizar dichas unidades se usaron los programas Excel 2016, Publish or

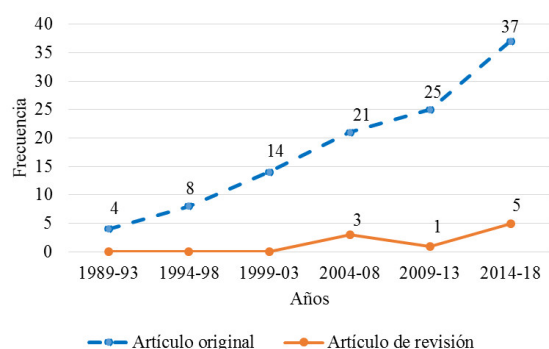
Perish y VOSviewer v1.6.13. Los dos primeros sirvieron para realizar las tablas y gráficos de los indicadores de producción (según año, tipo, idioma, país, área de investigación e índice de productividad de autores) y la colaboración de autores (según índice, grado y coeficiente de colaboración). VOSviewer (Van Eck y Waltman, 2019) se usó para analizar las citas y crear los mapas de emparejamiento bibliográfico y de cocitación (según fuentes y autores), además de los mapas de tendencia de investigación (según coocurrencia de palabras clave y de términos en los resúmenes). Para crear la mayoría de estos mapas se usó el método *fractional counting*, el cual asigna el mismo peso a cada acción y es recomendado para este tipo de análisis (Perianes-Rodríguez y otros, 2016; van Eck y Waltman, 2014; Martorell y otros, 2019).

3. RESULTADOS

3.1. Indicadores de producción

La producción quinquenal sobre AB en el periodo 1989-2018, según tipo de documentos (artículos originales y artículos de revisión), indicó la tendencia de crecimiento de publicaciones hasta el quinquenio 2014-2018 (Figura 1). Este quinquenio fue el más productivo con 42 artículos. Cabe destacar que desde 1996 se ha publicado al menos un documento por año. Respecto al tipo de documento, el artículo original fue el predominante con 109 ítems, frente al artículo de revisión con solo 9 ítems, el cual empezó a publicarse a mitad de los quinquenios indicados (2004-2008).

Figura 1. Artículos sobre ansiedad bibliotecaria indizados en Scopus según quinquenio



Respecto a la producción por país, la mayoría de documentos tuvo afiliación de Estados Unidos (57%), seguido por Malasia (8%) y Pakistán (6%). Menos representados estuvieron Australia, Canadá, Georgia, Jamaica, Jordania, Irán, Israel,

Kuwait, Sudán, Reino Unido (2% cada uno) y Bangladesh, China, Alemania, Indonesia, Japón, México, Polonia (1% cada uno). Un 3% no tuvo afiliación de país. Asimismo, la mayoría de documentos se escribió en inglés (98%), y en menor proporción en español y japonés (1% cada uno).

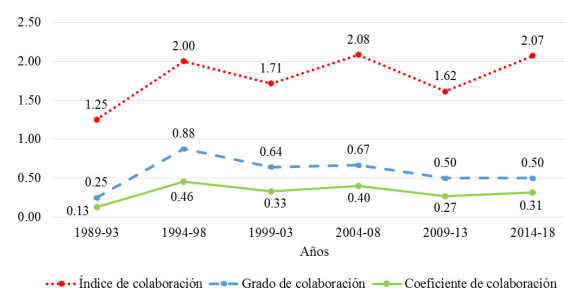
Respecto a las áreas temáticas, se encontró que los 118 artículos estudiados se distribuyen en 9 áreas: *Social Sciences* (77%), *Computer Science* (15%), *Arts and Humanities* (2%), *Decision Sciences* (2%), *Business, Management and Accounting* (1%), *Medicine* (1%), *Mathematics* (1%), y *Psychology* (1%). Cabe resaltar que algunos de estos artículos están descritos en Scopus en más de un área de temática. También que las dos áreas predominantes, *Social Sciences* y *Computer Science*, acumularon el 92% de publicaciones, y que las áreas *Arts and Humanities* y *Psychology* estuvieron poco representadas.

Por otro lado, se encontró un total de 150 autores. La gran mayoría (87%) tiene solo una publicación, mientras que el resto (13%) tiene dos o más publicaciones. Se encontró que Onwuegbuzie, A. J. (22 artículos) y Jiao, Q. G. (19 artículos) son grandes productores, con un IP de 1.3 cada uno. Karim, N. H. A. (8 artículos) y Anwar, M. A. (6 artículos) son productores intermedios con un IP de 0.9 y 0.8, respectivamente. Otros nueve autores intermedios tuvieron poca actividad (IP entre 0.6 y 0.3). Lo llamativo del campo de la AB es que el resto de autores (87%) son autores transitorios por tener una sola publicación. Es decir, la gran mayoría no suele seguir publicando sobre el tema después del primer trabajo.

3.2. Índice, grado y coeficiente de colaboración

Respecto a la colaboración de autores, el 57% de los trabajos se escribieron en colaboración, frente al 43% que tuvo un solo autor (Tabla II). No obstante, la mayoría de colaboración se dio solamente entre dos autores (36%). Cabe destacar que el quinquenio con mayor colaboración fue el periodo 2014-2018.

Figura 2. Índice, grado y coeficiente de colaboración de autores



La tendencia de colaboración según el índice, el grado y el coeficiente de colaboración (IC, GC y CC, respectivamente), para cada quinquenio, indicó un bajo IC de autores, que osciló entre uno y dos en los últimos años (Figura 2). Asimismo, el GC con mayor proporción de coautoría ocurrió en el quinquenio 1994-1998 (0.88), en los inicios de la investigación sobre AB. No obstante, la tendencia fue que en los últimos años dicha proporción bajó, siendo el periodo 2014-2018 el que presentó igualdad entre autores con una publicación y autores con al menos dos publicaciones. Por su parte, el CC tuvo un comportamiento similar al GC, con la diferencia que presentó valores menores. Es decir, en general la colaboración en el campo de la AB fue baja.

3.3. Análisis de citas, emparejamiento bibliográfico y cocitación de fuentes y autores

Fuentes

Se encontró un total de 42 revistas que publicaron sobre AB. Entre las 10 revistas más citadas destacaron *Library and Information Science Research* con 534 citas, seguida de *Library Review* con 222 citas y *College and Research Libraries* con 204 citas (Tabla III). Es decir, estas fueron las revistas con mayor impacto en términos de citas recibidas.

Tabla II. Distribución de publicaciones según periodo y número de autores

	Periodo						Total
	89-93	94-98	99-03	04-08	09-13	14-18	
Número de autores							
1	3	1	5	8	13	21	51 (43%)
2	1	6	8	7	11	9	42 (36%)
≥ 3	0	1	1	9	2	12	26 (21%)
Total	4	8	14	24	26	42	118

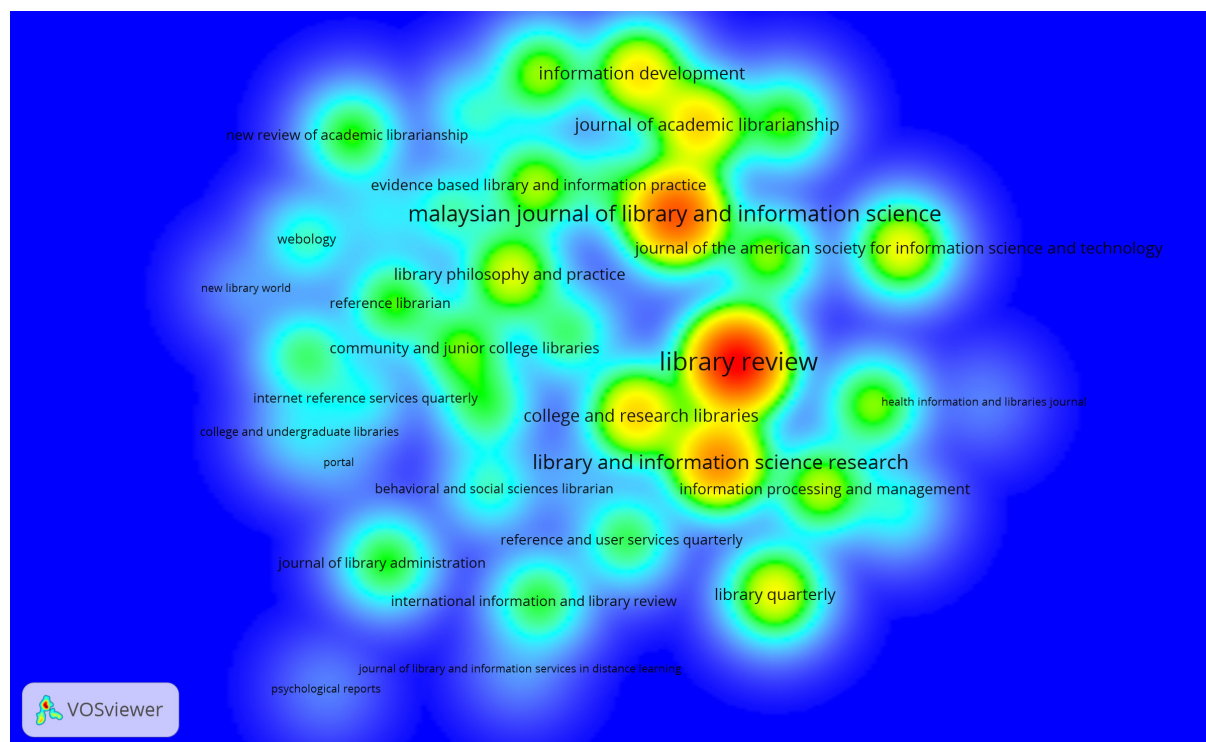
Tabla III. Diez revistas más citadas

Revista	Documentos	Total de citas
Library and Information Science Research	9	534
Library Review	12	222
College and Research Libraries	8	204
Library Quarterly	2	101
Reference and User Services Quarterly	6	96
Journal of Academic Librarianship	4	64
Malaysian Journal of Library and Information Science	7	55
Reference Services Review	4	53
Australian Academic and Research Libraries	2	47
Journal of American Society for Information Science and Technology	2	46

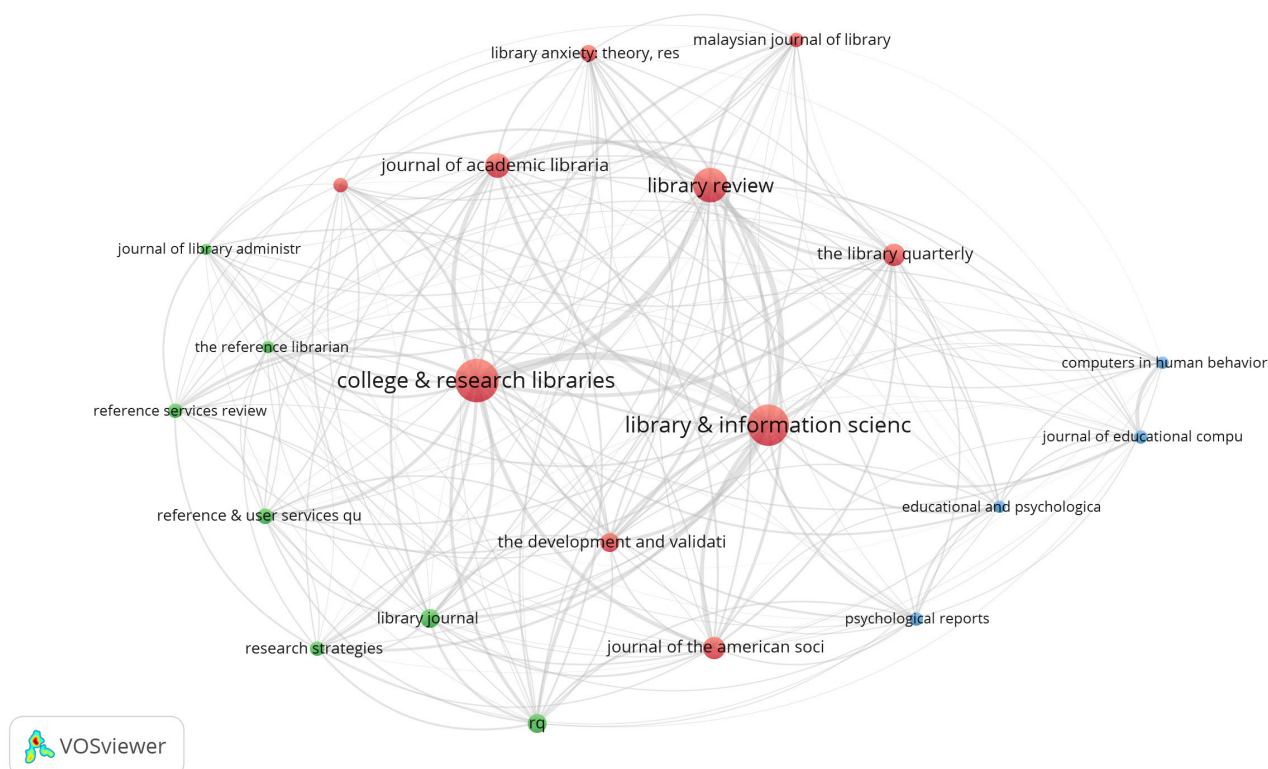
Para determinar las agrupaciones de las revistas, se hizo un análisis de emparejamiento biblio-

gráfico (AEB) (Figura 3). Se encontró, dentro de las revistas identificadas, dos agrupaciones principales: una cuyo foco fue *Library Review* (con una fortaleza de enlace de 207.8) y otra cuyo foco fue *Malaysian Journal of Library and Information Science* (con una fortaleza de enlace de 138.63). Es decir, el AEB permitió determinar que alrededor de estas revistas se emparejan las fuentes que tienen similares referencias en común.

Para complementar este resultado, se hizo un análisis de cocitación (AC) de fuentes (Figura 4). Se encontraron tres revistas predominantes según cocitación: *College & Research Libraries* (con una fortaleza de enlace de 172.2), *Library and Information Science Research* (con una fortaleza de enlace de 160.5) y *Library Review* (con una fortaleza de enlace de 119.2). Cabe destacar que dentro de estas fuentes, dos no son revistas: la tesis doctoral *The development and validation of the Library Anxiety Scale (1993)* y el libro *Library anxiety: theory, research and application (2004)*, con fortaleza de enlace de 47 y 36 respectivamente. En dicha tesis se propuso la escala de Bostick de AB y en el mencionado libro se presenta una síntesis de la investigación en AB desde 1986 hasta 2004.

Figura 3. Mapa de densidad de emparejamiento bibliográfico de revistas


Nota: Se usaron revistas con un documento como mínimo. De las 42 revistas, solo 40 estuvieron conectadas entre sí. Método de conteo: *fractional counting*. Método de normalización: *association strength*.

Figura 4. Mapa de cocitación de fuentes

Nota: Se hizo un tesaurus para estandarizar los nombres de las fuentes cocitadas en las referencias de los 118 documentos. Se usaron las fuentes que tuvieron 25 citas como mínimo. De las 1267 fuentes, solo 21 cumplieron dicho umbral. Método de conteo: *fractional counting*. Método de normalización: *LinLog/modularity*.

Autores

De los 150 autores identificados, Onwuegbuzie y Jiao, con 821 y 654 citas respectivamente, encabezaron la lista de los diez autores con mayor impacto en términos de citación (Tabla IV).

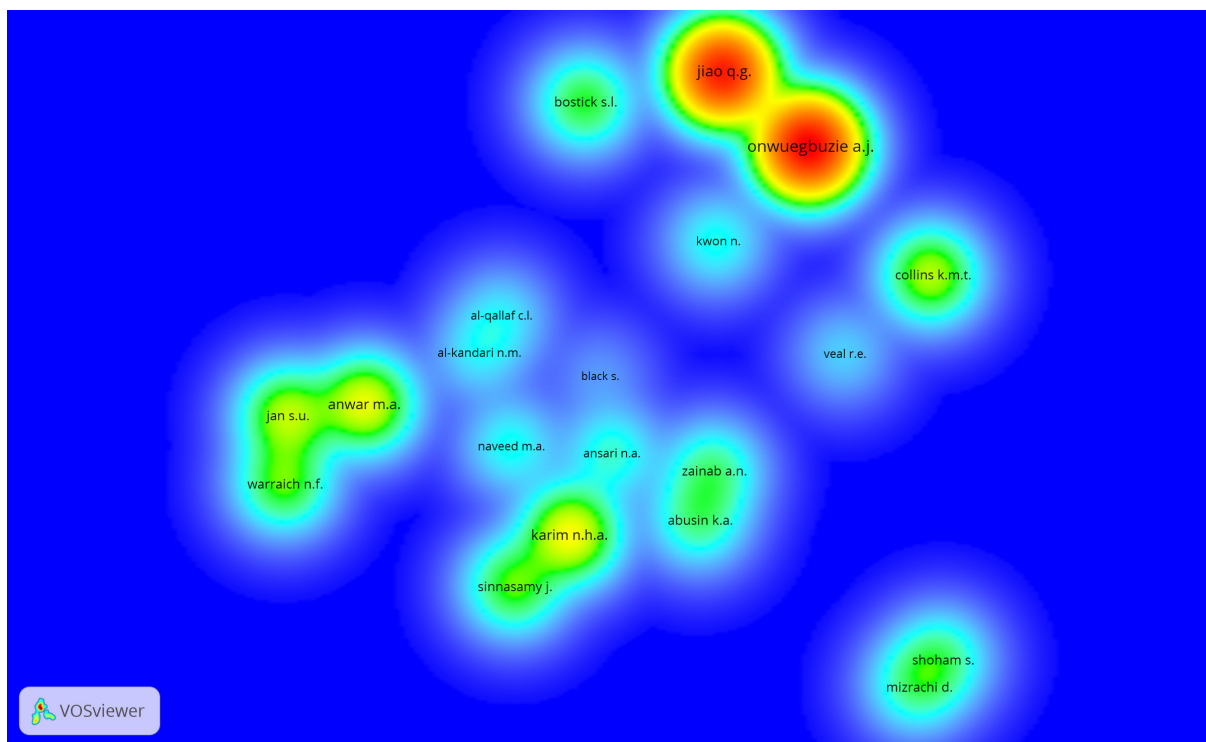
Tabla IV. Diez autores más citados

Autores	Documentos	Total de citas
Onwuegbuzie, A. J.	22	821
Jiao, Q. G.	19	654
Gross, M.	1	123
Latham, D.	1	123
Lichtenstein, A. A.	1	99
Kwon, N.	2	73
Anwar M. A.	6	71
Al-kandari N. M.	2	55
Al-qallaf C. L.	2	55
Karim N. H. A.	8	48

Por medio del AEB (Figura 5), se encontraron dos grandes agrupaciones de autores: una cuyo foco lo formaron Onwuegbuzie y Jiao, con fortaleza de enlace de 935.4 y 824.61, respectivamente; y otro cuyo foco lo formaron Karim y Anwar, con fortaleza de enlace de 212.77 y 197.75, respectivamente. Básicamente, el primer grupo corresponde a autores con publicaciones clásicas sobre AB, mientras que el segundo grupo corresponde a los autores con publicaciones más actuales. En general, alrededor de dichos autores se emparejaron otros autores que tienen referencias en común.

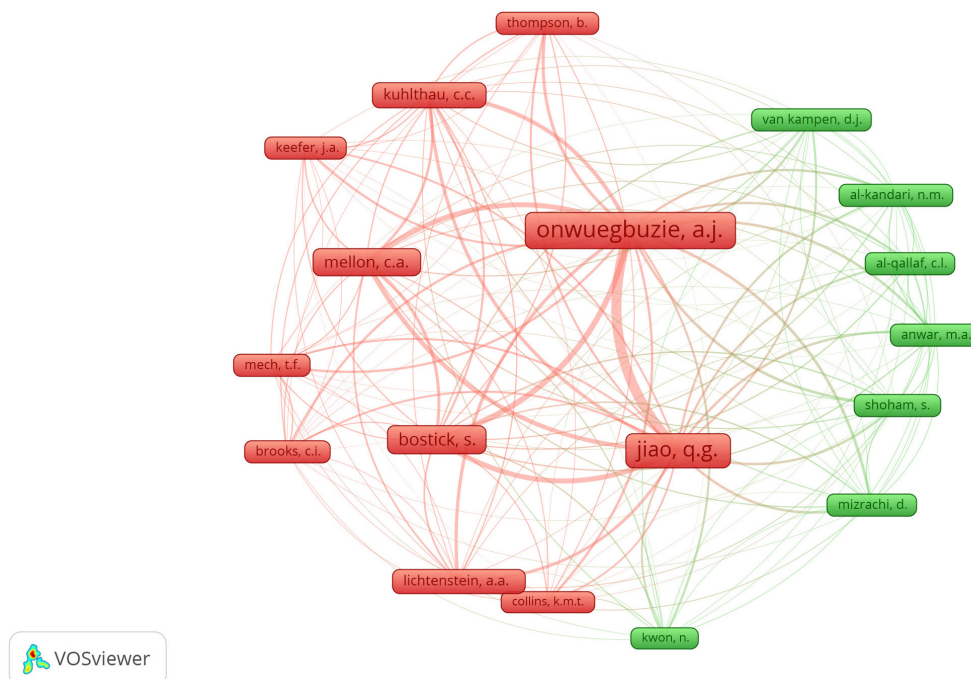
Para complementar el análisis anterior, se indagaron sobre los autores más cocitados (Figura 6). De forma similar a los análisis anteriores, Onwuegbuzie y Jiao fueron autores relevantes, pues alcanzaron el mayor número de citas (572 y 437, respectivamente) y la mayor fuerza de enlace en cocitación (390 y 334 respectivamente). Los siguientes 3 autores relevantes fueron Mellon (con 137 citas y 142 de fuerza de enlace), Bostick (con 151 citas y 117 de fuerza de enlace) y Kuhlthau

Figura 5. Mapa de densidad de emparejamiento bibliográfico de autores



Nota: Se usaron autores con dos publicaciones como mínimo. De los 150 autores, 20 cumplieron dicho umbral. Método de conteo: *fractional counting*. Método de normalización: *association strength*.

Figura 6. Mapa de cocitación de autores



Nota: Se hizo un tesoro para estandarizar los nombres de los autores cocitados en las referencias de los 118 artículos. Se usaron los autores que tuvieran como mínimo 25 citas. De los 150 autores, solo 18 cumplieron dicho umbral. Método de conteo: *fractional counting*. Método de normalización: *LinLog/modularity*.

(con 83 citas y 76 de fuerza de enlace), quienes hicieron aportes fundamentales a nivel teórico y metodológico con sus investigaciones. En ese sentido, estos autores fueron los más influyentes en términos de cocitación.

3.4 Tendencias de investigación en el campo de ansiedad bibliotecaria

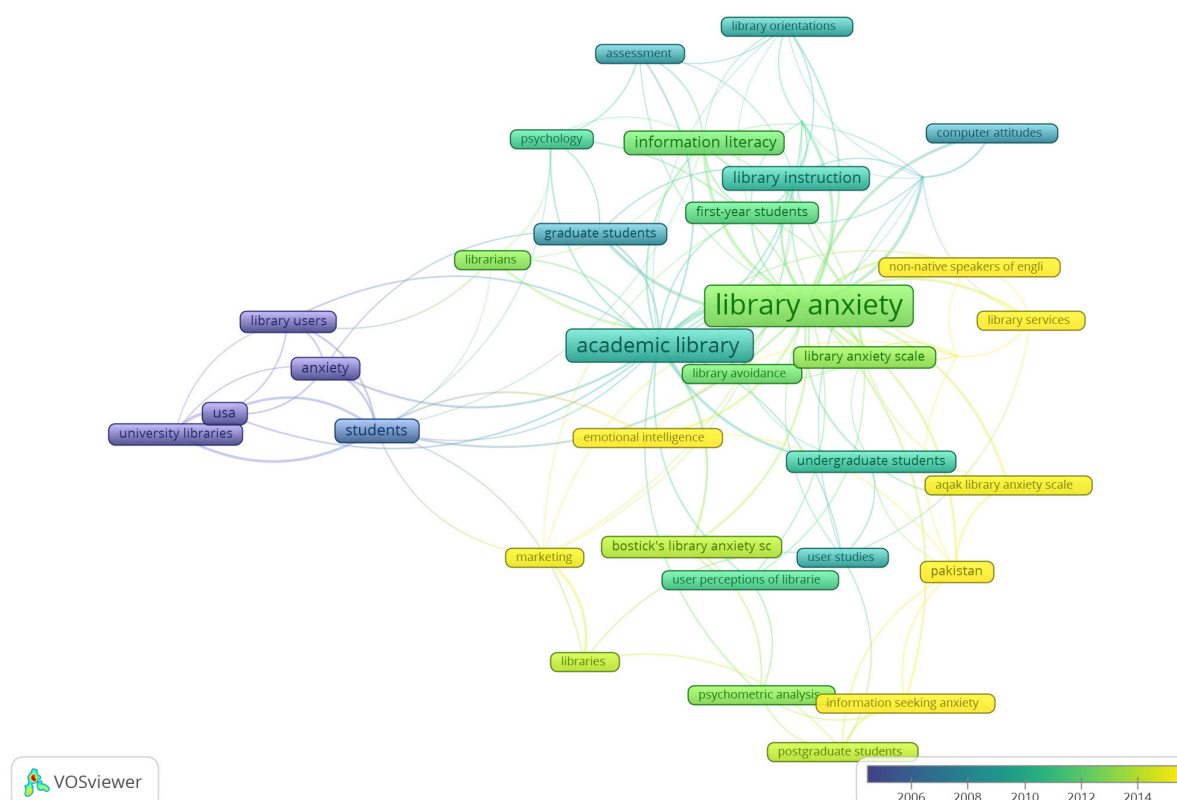
De acuerdo al mapa de superposición de palabras claves de autores (Figura 7), algunos términos destacados en los 118 artículos fueron *library anxiety*, *academic library*, *students*, *information literacy* y *library instruction* (con 39, 23, 10, 8 y 8 de fuerza de enlace respectivamente). Respecto a la línea temporal, el término *library anxiety* sigue siendo vigente y su estudio ha sido relacionado tradicionalmente con los términos *university libraries*, *students*, *USA* y *anxiety*. En años posteriores, el estudio de la AB se asoció con términos sobre la formación del usuario (*information literacy* y *library instruction*), y en años recientes con los términos *emotional intelligence*, *information seeking anxiety*, *marketing* y *Pakistan*.

Por último, el análisis de los términos extraídos de los resúmenes permitió visualizar 89 términos, los cuales se agruparon en cuatro frentes de investigación (Figura 8): (1) azul: términos relacionados a instrumentos de medición, como *scale*, *instrument*, *sub scale* y *LAS* (library anxiety scale). (2) Rojo: términos enfocados en marcos teóricos, como *information*, *service*, *librarian*, *process*. (3) Amarillo: términos asociados a los resultados, como *implication*, *graduate student*, *role*. (4) Verde: términos relacionados a tipos de unidades de análisis y sus características, como *university student*, *library use*, *age*, *sex*, *native language*.

4. DISCUSIÓN Y CONCLUSIONES

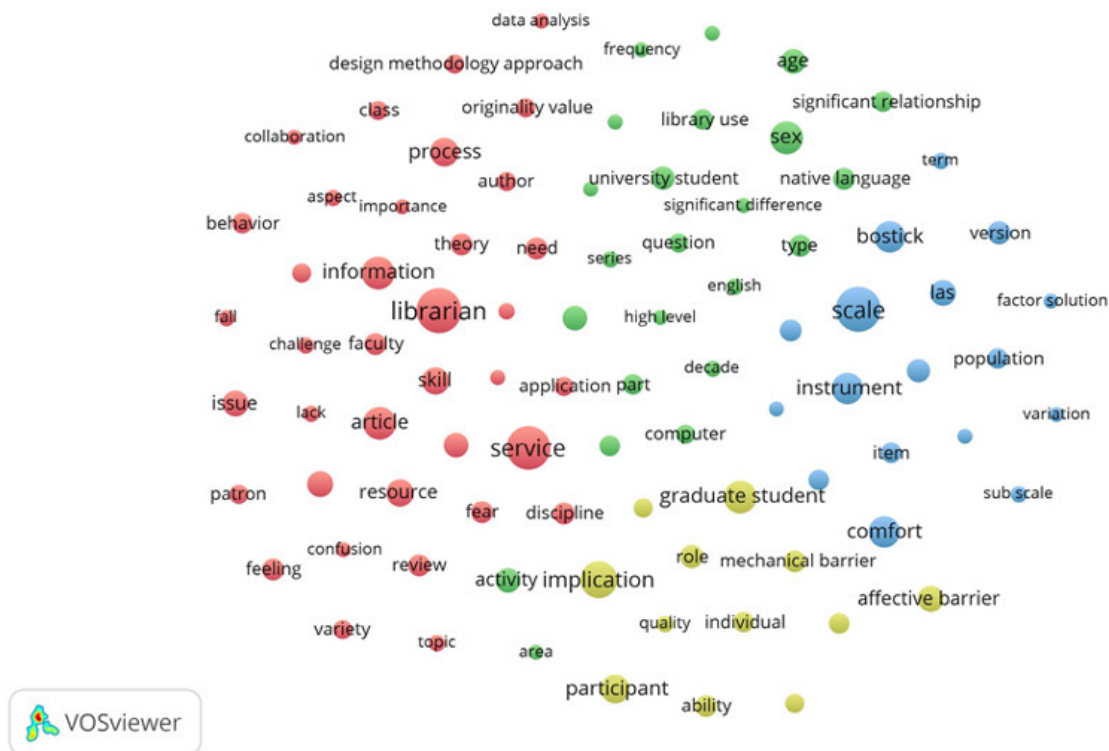
En primer lugar, la investigación en AB ha tenido un crecimiento progresivo, aunque su presencia no es comparable al de otros tópicos de investigación en bibliotecología y ciencias de la información (Togia y Malliari, 2017). Sin embargo, esta es una oportunidad para estudiar el tema en regiones como Latinoamérica, que solo cuenta con el estudio de Ramírez (2015) (el 1% de las publicaciones analizadas). Asi-

Figura 7. Mapa de superposición de palabras clave de autores



Nota: Se hizo un tesaurus para estandarizar las palabras clave de autores. Se usaron solo las que tuvieran tres ocurrencias como mínimo. De las 205 palabras clave, solo 35 cumplieron dicho umbral. Método de conteo: *fractional counting*. Método de normalización: *fractionalization*.

Figura 8. Mapa de red de términos en los resúmenes



Nota: Se hizo un tesauro para estandarizar los términos de los resúmenes. Se usaron solo los que tuvieran cuatro ocurrencias como mínimo. Método de normalización: *association strength*.

mismo, se reafirmó lo señalado por Cleveland (2004) sobre la importancia de los autores Onwuegbuzie y Jiao, quienes además tienen trabajos en conjunto, aunque se evidenció que, en general, la tendencia de trabajos en colaboración fue baja. Esto resulta llamativo en una época de crecimiento de trabajos interdisciplinarios y de colaboración en diferentes áreas (Aldrich y Al-Turk, 2018; Hall y otros, 2018; Bu y otros, 2018). Quizá por ello las publicaciones estudiadas destacan en el área *Social Science* y no se observe mayor interdisciplinariedad. Solo se pueden destacar los trabajos desde el área *Computer Science*, donde se estudia la AB asociada a la actitud y ansiedad hacia la tecnología (Jerabek y otros, 2001).

En segundo lugar, las revistas *Library and Information Science Research* y *Library Review*, fueron las de mayor impacto. Como complemento, las revistas *Library Review* y *Malaysian Journal of Library and Information Science*, así como los autores Onwuegbuzie, Jiao, Anwar y Karim generaron grandes agrupaciones de revistas y autores, respectivamente. Al respecto, según el portal Scimago, la revista *Library and Information Science Research* es la mejor posicionada en el cuartil 1 (Q1), seguida de *Malaysian Journal of Library and Information Science* en el

cuartil 2 (Q2) y de *Library Review* ubicada en el cuartil 3 (Q3). De forma similar, según cocitaciones, otra revista relevante fue *College & Research Libraries*, además de los autores Mellon, Bostick y Kuhlthau, quienes alcanzaron un alto número de citas y una alta fortaleza de cocitación. En resumen, el análisis descrito permitió tener un panorama de fuentes y autores relevantes, tanto con publicaciones clásicas como con investigaciones recientes.

En tercer lugar, la tendencia ha sido pasar de estudios sobre la naturaleza de la AB hacia su relación con la formación del usuario (*library instruction*, *information literacy*). Otras relaciones interesantes se dan con los tópicos *emotional intelligence*, *information seeking anxiety* y *marketing*. Por otra parte, según el análisis de resúmenes, los estudios sobre AB se han enfocado, básicamente, en estudiantes universitarios y no en usuarios de otras unidades de información. Asimismo, destacaron términos sobre elementos metodológicos de tipo cuantitativo, lo que sugiere una oportunidad para investigar la AB desde enfoques cualitativos. Todo lo mencionado indica lo fructífero de investigar el tema, dependiendo de la perspectiva que se siga (Wildemuth, 2017).

Finalmente, el estudio se basó solamente en artículos recuperados de Scopus y no en otro tipo de documento (capítulo de libro, editorial, conferencia y nota, que solo acumularon el 6% de las publicaciones sobre AB), por lo que se sugiere para futuros estudios incluir otros tipos de documentos y otras bases de datos. Del mismo modo, los mapas bibliométricos representaron solo parte de la actividad científica en el campo de la AB. No obstante, se buscó evidenciar la importancia de la bibliometría y cienciometría para describir temas de investigación en desarrollo como la AB, y suscitar investigaciones futuras al respecto dentro de la comunidad de bibliotecología y ciencias de la información, especialmente en Latinoamérica.

6. REFERENCIAS

- Abramo, G., D'Angelo, C.; Reale, E. (2019). Peer review vs bibliometrics: which method better predicts the scholarly impact of publications? *Scientometrics*, 121(1), 537–554. <https://doi.org/10.1007/s11192-019-03184-y>
- Ajiferuke, I.; Burell, Q.; Tague, J. (1988). Collaborative coefficient: A single measure of the degree of collaboration in research. *Scientometrics*, 14(5–6), 421–433. <https://doi.org/10.1007/BF02017100>
- Aksnes, D. W.; Langfeldt, L.; Wouters, P. (2019). Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories. *SAGE Open*, 9(1), 1–17. <https://doi.org/10.1177/2158244019829575>
- Aldrich, H.; Al-Turk, A. (2018). Crouching Authors, Hidden Pitfalls: Collaboration in Research. *Studi Di Sociologia*, 56(4), 341–328. https://doi.org/10.26350/000309_000041
- Anwar, M.; Al-Qallaf, C.; Al-Kandari, N.; Al-Ansari, H. (2012). AQAK: A library anxiety scale for undergraduate students. *Journal of Librarianship and Information Science*, 44(1), 36–46. <https://doi.org/10.1177/0961000611425568>
- Arcas, A.; Cano, A. (1999). Procesos cognitivos en el trastorno de ansiedad generalizada, según el paradigma del procesamiento de la información. *Psicología.com: Revista Electrónica de Psicología*, 3(1). Disponible en: <http://psiqui.com/1-6380> [Fecha de consulta: 04/09/2019].
- Björneborn, L.; Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology*, 55(14), 1216–1227. <https://doi.org/10.1002/asi.20077>
- Boyack, K.; Klavans, R. (2010). Co-Citation Analysis, Bibliographic Coupling, and Direct Citation: Which Citation Approach Represents the Research Front Most Accurately? *Journal of the American Society for Information Science and Technology*, 61(12), 2389–2404. <https://doi.org/10.1002/asi.21419>
- Browne, D. (2018). Anxiety: Overview, symptoms, causes, and treatments. Disponible en: <https://www.medicalnewstoday.com/articles/323454.php> [Fecha de consulta: 04/09/2019].

5. AGRADECIMIENTOS

Se agradece al grupo de investigación CIGETMEN de la Facultad de Letras y Ciencias Humanas de la Universidad Nacional Mayor de San Marcos, y a la Dirección de Investigación de la Universidad Peruana de Ciencias Aplicadas.

ACKNOWLEDGEMENTS

Thanks to the research group CIGETMEN of the Facultad de Letras y Ciencias Humanas, Universidad Nacional Mayor de San Marcos, and to the Dirección de Investigación de la Universidad Peruana de Ciencias Aplicadas.

- Bu, Y.; Ding, Y.; Liang, X.; Murray, D. S. (2018). Understanding persistent scientific collaboration. *Journal of the Association for Information Science and Technology*, 69(3), 438–448. <https://doi.org/10.1002/asi.23966>
- Carlile, H. (2007). The implications of library anxiety for academic reference services: A review of literature. *Australian Academic and Research Libraries*, 38(2), 129–147. <https://doi.org/10.1080/00048623.2007.10721282>
- Cleveland, A. (2004). Library anxiety: A decade of empirical research. *Library Review*, 53(3), 177–185. <https://doi.org/10.1108/00242530410526583>
- Cortés, D. (2007). Medir la producción científica de los investigadores universitarios: la bibliometría y sus límites. *Revista de Educación Superior*, 36(142), 43–65. Disponible en: <https://www.redalyc.org/articulo.oa?id=60414203> [Fecha de consulta: 11/09/2019].
- Elsevier. (2019). ¿Qué es Scopus? [En línea]. Disponible en: https://service.elsevier.com/app/answers/detail/a_id/15534/supporthub/scopus/#tips [Fecha de consulta: 02/10/2019].
- García, Z.; Cano, A. (2014). Estandarización y validación del Inventario de Situaciones y Respuestas de Ansiedad (ISRA) en la población de República Dominicana. *Summa Psicológica UST*, 11(2), 81–99. Disponible en <https://dialnet.unirioja.es/servlet/articulo?codigo=4954000> [Fecha de consulta: 21/09/2012]. <https://doi.org/10.18774/448x.2014.11.130>
- Habib, R.; Afzal, M. (2017). Paper recommendation using citation proximity in bibliographic coupling. *Turkish Journal of Electrical Engineering & Computer Sciences*, 25, 2708–2718. <https://doi.org/10.3906/elk-1608-180>
- Hall, K. L.; Vogel, A. L.; Huang, G. C.; Serrano, K. J.; Rice, E. L.; Tsakraklides, S. P.; Fiore, S. M. (2018). The science of team science: A review of the empirical evidence and research gaps on collaboration in science. *American Psychologist*, 73(4), 532–548. <https://doi.org/10.1037/amp0000319>
- Izquierdo, M. (1999). Una aproximación interdisciplinar al estudio del usuario de información: bases conceptuales y metodológicas. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*, 13(26), 112–134. <https://doi.org/10.22201/iibi.0187358xp.1999.26.3902>

- Jerabek, J. A.; Meyer, L. S.; Kordinak, S. T. (2001). "Library anxiety" and "computer anxiety:" Measures, validity, and research implications. *Library & Information Science Research*, 23(3), 277-289. [https://doi.org/10.1016/s0740-8188\(01\)00083-4](https://doi.org/10.1016/s0740-8188(01)00083-4)
- Lawani, S. M. (1981). Bibliometrics: Its Theoretical Foundations, Methods and Applications. *Libri*, 31(4), 294-315. <https://doi.org/10.1515/libr.1981.31.1.294>
- Leal, P.; Goes, T.; da Silva, L.; Teixeira-Silva, F. (2017). Trait vs. state anxiety in different threatening situations. *Trends Psychiatry Psychother*, 39(3), 147-157. <https://doi.org/10.1590/2237-6089-2016-0044>
- Marshakova-Shaivevich, I. (1973). System of documentation connections based on references (SCI). *Nauchno Tekhnicheskaya Informatsiya Seriya*, 2(6), 3-8. Disponible en: <http://garfield.library.upenn.edu/marshakova/marshakovanaughtech1973.pdf> [Fecha de consulta: 01/10/2019].
- Martorell, O.; Socias, A.; Otero, L.; Mulet-Forteza, C. (2019). Thirty-fifth anniversary of the International Journal of Hospitality Management: A bibliometric overview. *International Journal of Hospitality Management*, 78, 89-101. <https://doi.org/10.1016/j.ijhm.2018.10.013>
- McCain, K. W. (1991). Mapping economics through the journal literature: An experiment in journal cocitation analysis. *Journal of the American Society for Information Science*, 42(4), 290-296. [https://doi.org/10.1002/\(SICI\)1097-4571\(199105\)42:4<290::AI-D-AS15>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(199105)42:4<290::AI-D-AS15>3.0.CO;2-9)
- Mellon, C. A. (1986). Library Anxiety: A Grounded Theory and Its Development. *College & Research Libraries*, 47(2), 160-165. <https://doi.org/10.5860/crl.76.3.276>
- Mingers, J.; Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, 246, 1-19. <https://doi.org/10.1016/j.ejor.2015.04.002>
- Mohanty, B.; Sahoo, J.; Dash, N. (2018). Bibliometric Indicators for Assessing the Quality of Scholarly Communications: A Case Study on International Journal of Cooperative Information Systems. *Library Philosophy and Practice (e-Journal)*. Disponible en: <https://digitalcommons.unl.edu/libphilprac/2158> [Fecha de consulta: 01/10/2019].
- Onwuegbuzie, J.; Jiao, Q.; Bostick, S. (2004). *Library Anxiety: Theory, Research, and Applications*. EE. UU: Scarecrow Press Inc. Disponible en: <https://cutt.ly/Wrezi5D> [Fecha de consulta: 11/09/2019].
- Peralta, M.; Frías, M.; Chaviano, O. (2015). Criterios, clasificaciones y tendencias de los indicadores bibliométricos en la evaluación de la ciencia. *Revista Cubana de Información En Ciencias de La Salud*, 26(3), 290-309. Disponible en: <http://dspace.uclv.edu.cu/handle/123456789/5101> [Fecha de consulta: 18/09/2019].
- Perianes-Rodríguez, A.; Waltman, L.; van Eck, N. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics*, 10(4), 1178-1195. <https://doi.org/10.1016/j.joi.2016.10.006>
- Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25(4), 348-349. Disponible en: https://www.researchgate.net/publication/236031787_Statistical_Bibliography_or_Bibliometrics [Fecha de consulta: 02/09/2019].
- Ramírez, N. (2015). Ansiedad bibliotecaria en estudiantes universitarios. *Revista Interamericana de Bibliotecología*, 38(3), 227-236. <https://doi.org/10.17533/udea.rib.v38n3a07>
- Riggio, G. (2017). *Indicadores bibliométricos de la actividad científica de la República Dominicana*. [Tesis doctoral]. Madrid: Universidad Carlos III. Disponible en: <http://eprints.rclis.org/31698/> [Fecha de consulta: 01/10/2019].
- Roemer, R.; Borchardt, R. (2015). *Meaningful Metrics: A 21st-Century Librarian's Guide to Bibliometrics, Altimetrics, and Research Impact*. Chicago, Illinois: Association of College and Research Libraries. Disponible en: http://www.ala.org/acrl/sites/ala.org/acrl/files/content/publications/booksanddigitalresources/digital/9780838987568_metrics_OA.pdf [Fecha de consulta: 11/09/2019].
- Salas, G.; Ravelo-Contreras, E. L.; Mejía S., Andrades R., Acuña E., Espinoza F.; Núñez, M.; Barboza-Palomino, M.; Ventura-León, J.; Caycho-Rodríguez, T.; Pérez-Acosta, A. M. (2018). Dos décadas de Acta Colombiana de Psicología: un análisis bibliométrico. *Acta Colombiana de Psicología*, 21(2), 13-38. <https://doi.org/10.14718/acp.2018.21.2.2>
- Sierra, J.; Ortega, V.; Zubeidat, I. (2003). Ansiedad, angustia y estrés: tres conceptos a diferenciar. *Revista Mal-Estar e Subjetividade*, 3(1), 10-59. Disponible en: <https://biblat.unam.mx/es/revista/revista-mal-estar-e-subjetividade/articulo/ansiedad-angustia-y-estres-tres-conceptos-a-diferenciar> [Fecha de consulta: 04/09/2019].
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265-269. <https://doi.org/10.1002/asi.4630240406>
- Subramanyam, K. (1983). Bibliometric studies of research collaboration: A review. *Journal of Information Science*, 6(1), 33-38. <https://doi.org/10.1177/016555158300600105>
- Tamayo-Rueda, D.; Meneses-Placeres, G. (2018). Comportamiento informacional: revisión de teorías posibles para su estudio. *E-Ciencias de La Información*, 8(2), 2-19. <https://doi.org/10.15517/ECI.V8I2.32441>
- Togia, A.; Malliari, A. (2017). Research Methods in Library and Information Science. En: Oflazoglu, S. (ed.). *Qualitative Versus Quantitative Research*, pp. 43-64. IntechOpen. doi:10.5772/intechopen.68749. <https://doi.org/10.5772/intechopen.68749>
- Van Eck, N.; Waltman, L. (2014). Visualizing bibliometric networks. En Ding, Y.; Rousseau, R.; Wolfram, D. (Eds.), *Measuring scholarly impact: Methods and practice*. pp. 285-320. Springer. https://doi.org/10.1007/978-3-319-10377-8_13
- Van Eck, N.; Waltman, L. (2019). *VOSviewer Manual*. Disponible en: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.10.pdf [Fecha de consulta: 25/09/2019].
- Van Kampen, D. (2004). Development and Validation of the Multidimensional Library Anxiety Scale. *College and Research Libraries*, 65(1), 28-34. <https://doi.org/10.5860/crl.65.1.28>

- Vargas-Quesada, B.; de Moya-Anegón, F. (2007). *Visualizing the structure of science*. New York, NY: Springer. Disponible en: <https://www.springer.com/gp/book/9783540697275>.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391. <https://doi.org/10.1016/j.joi.2016.02.007>
- Wildemuth, B. (2017). Library Anxiety Impedes College Students' Library Use, but May Be Alleviated Through Improved Bibliographic Instruction. *Evidence Based Library and Information Practice*, 12(4), 275. <https://doi.org/10.18438/b8k082>
- Yang, E.; Ding, Y. (2012). Scholarly Network Similarities: How Bibliographic Coupling Networks, Citation Networks, Cocitation Networks, Topical Networks, Coauthorship Networks, and Coword Networks Relate to Each Other. *Journal of the American Society for Information Science and Technology*, 63(7), 1313–1326. <https://doi.org/10.1002/asi.22680>

ESTUDIOS / RESEARCH STUDIES

Estado de la web local portuguesa: evidencias empíricas acerca de la brecha digital entre regiones

Valeriano Piñeiro-Naval*, Paulo Serra**

* Observatorio de los Contenidos Audiovisuales, Universidad de Salamanca (España)

** LabCom – Comunicação e Artes, Universidade da Beira Interior, Covilhã (Portugal).

Correo-e: vale.naval@usal.es | ORCID iD: <http://orcid.org/0000-0001-9521-3364>

Correo-e: pserra@ubi.pt | ORCID iD: <http://orcid.org/0000-0001-7821-3880>

Recibido: 12-01-20; 2ª versión: 18-03-20; Aceptado: 02-04-20; Publicado: 27-04-2021

Cómo citar este artículo/Citation: Piñeiro-Naval, V.; Serra, P. (2021). Estado de la web local portuguesa: evidencias empíricas acerca de la brecha digital entre regiones. *Revista Española de Documentación Científica*, 44 (2), e292. <https://doi.org/10.3989/redc.2021.2.1761>

Resumen: En aras de corresponder a las demandas de sus ciudadanos, los municipios han impulsado el gobierno electrónico mediante el uso de las TIC. Sin embargo, esta adopción tecnológica no siempre se efectúa de forma homogénea entre unos territorios y otros, originando las conocidas como «brechas digitales». Así pues, los objetivos del estudio consistieron en medir, a través del análisis de contenido de los 308 sitios web municipales de Portugal, la calidad de estos servicios electrónicos, para luego detectar diferencias interregionales y apuntar algunos factores predictivos del estado de la red local portuguesa. Desde el punto de vista instrumental, se diseñó un «Índice General de Calidad Web» que sirvió para concretar el grado de desempeño nacional, cercano a un nivel correcto, así como para concluir que el litoral continental del país, la zona más poblada, rica y desarrollada, es también el área donde el gobierno electrónico está mejor implementado.

Palabras clave: e-gobierno; brecha digital; sitios web municipales; Portugal; análisis de contenido; Índice General de Calidad Web.

State of Portuguese local web: empirical evidence about digital divide between regions

Abstract: In order to correspond to their citizens' demands, municipalities have promoted electronic government through the use of ICT. However, this technological adoption is not always carried out homogeneously when comparing some territories with others, causing what are known as "digital divides". Thus, the objectives of the study consisted on measuring, through the content analysis of the 308 municipal websites of Portugal, the quality of these electronic services, to then detect interregional differences and point out some predictive factors of the state of Portuguese local network. From the instrumental point of view, a "General Web Quality Index" was designed, that served to specify the national performance level, which was close to correct, as well as to conclude that the country's continental coastline, the most populated, rich and developed area, is precisely where e-government is best implemented.

Keywords: e-government; digital divide; municipal websites; Portugal; content analysis; General Web Quality Index.

Copyright: © 2021 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1. INTRODUCCIÓN

Por muy superficial que esta sea, toda reflexión acerca de la sociedad actual nos tendría que remitir, de forma ineludible, a las «Tecnologías de la Información y la Comunicación» (TIC), aquellos elementos que facilitan tanto la conexión entre los distintos agentes sociales como su acceso al conocimiento generado. Constituyen, por tanto, un avance equiparable al que representó la imprenta; una revolución que también motivó, en aquel entonces, un cambio de paradigma social. En un ecosistema tan complejo, «la gestión de la información y su documentación resulta determinante para la eficacia de los procedimientos a los que se ven sometidos, de manera recurrente, los ayuntamientos» (Campillo-Alhama, 2013); los organismos más próximos al ciudadano y cuyas obligaciones radican en atender, en primera instancia, sus demandas y solicitudes. Para cumplir con este imperativo, las instituciones públicas han impulsado lo que se ha dado en denominar «e-gobierno» que, a grandes rasgos, consiste en el uso de las TIC para promover una reforma administrativa que devenga en la prestación eficiente de servicios (Helbig y otros, 2009; López-López y otros, 2018). Entre sus principales ventajas destacarían: una mayor accesibilidad a los medios públicos y la participación de los ciudadanos de la e-democracia, una mejor capacidad de respuesta a sus necesidades y, por último, una reducción de los tiempos y los costes burocráticos (Detlor y otros, 2010). El concepto y la práctica del e-gobierno son etiquetados mediante la utilización de términos como «digital», «en línea» y, últimamente, adjetivos más novedosos como «móvil», «ubicuo» o «inteligente» (Manoharan e Ingrams, 2018).

No obstante, la difusión del e-gobierno ha estado condicionada, desde sus inicios, por otro factor íntimamente asociado: la «brecha digital» (*digital divide*), una idea que ahonda en la relación que se establece entre los distintos grupos de individuos y las TIC. En este sentido, y según Helbig y otros (2009), cabe matizar que los estudiosos del e-gobierno se han centrado, habitualmente, en las iniciativas públicas que propician los servicios electrónicos y, en otros casos, en las oportunidades de participación de los ciudadanos, las empresas y demás partes interesadas; esto es, en todo lo relativo a la oferta de estos mecanismos. Por otro lado, y siguiendo a los mismos autores, los investigadores de la brecha digital analizan cómo diferentes grupos sociales intentan aprovechar estos dispositivos y otros recursos de las TIC; es decir, el modo mediante el que se comporta la demanda. Pero ¿qué es la brecha digital? Sucintamente, «se considera una distinción entre aquellos con y sin acceso

a Internet, o entre los usuarios y no usuarios de Internet» (Choudrie y otros, 2013). En los países desarrollados, los efectos de la brecha de acceso a la administración electrónica están disminuyendo como resultado de la expansión generalizada de la red (Ebbers y otros, 2016), si bien es cierto que los ingresos económicos y las infraestructuras continúan siendo factores clave a este respecto.

En suma, la presente investigación nace con un claro propósito: evaluar la calidad de los portales web oficiales de los 308 municipios portugueses, así como observar las eventuales diferencias en su implementación a nivel regional. Se sabe que la calidad de un producto o servicio es un atributo complejo y multidimensional (Carvalho y otros, 2016), que arroja luz a su grado de adecuación o ajuste a las necesidades de los usuarios (Karabasevic y otros, 2019). Así pues, propondremos una estrategia para medir el desempeño de estos portales basada en sus contenidos culturales y turísticos, su diseño web y su empleo de las principales herramientas sociales o 2.0.

1.1. Estado de la cuestión

Serán enumerados, de inmediato, algunos trabajos empíricos que han demostrado la existencia de brechas digitales interterritoriales no solo en Portugal, sino en otros países de su entorno; y que, además, han identificado los principales factores predictores del estado de la e-administración local.

Para comenzar este repaso por el ámbito internacional, el estudio de Ingrams y otros (2018) analiza una muestra de 100 *websites* de las principales ciudades del mundo, empleando una *checklist* compuesta por 104 ítems englobados en 5 parámetros, como son: 1) contenido, 2) usabilidad, 3) privacidad y seguridad, 4) servicios y, en último lugar, 5) implicación social y ciudadana. Entre sus hallazgos, los autores determinan que el tamaño poblacional y las características socioeconómicas de las localidades están positivamente asociadas con un mejor e-gobierno. En esta línea, Manoharan y otros (2017) ya habían testado con anterioridad la misma estrategia metodológica, validando así un modelo correlacional donde la adopción del e-gobierno dependía de cuestiones económicas, democráticas, educativas y de uso de Internet. En un contexto más próximo, Philip y otros (2017) ilustran la brecha digital existente en el seno de Gran Bretaña entre los ámbitos rural y urbano; unas desigualdades basadas, sobre todo, en las infraestructuras tecnológicas y los datos financieros.

En el caso español, García-Santamaría y Martín-Matallana (2017) se sirven de los datos proporcionados en el «Índice de Transparencia de los

Ayuntamientos» (Transparencia Internacional España, 2017); un informe donde son evaluados los portales de los 110 concejos españoles de mayor tamaño, y cuyos niveles de transparencia dependen de un total de 80 indicadores agrupados en 6 áreas. Tras el posterior tratamiento de estos datos secundarios, García-Santamaría y Martín-Matallana (2017) concluyen que el signo ideológico del partido político al frente del consistorio no afecta al grado de transparencia de su sede en línea. Por su parte, Piñeiro-Naval y otros (2017) sostienen que la calidad de los sitios web de los municipios españoles depende, principalmente, del número de habitantes y de los indicadores de desarrollo tecnológico; para el caso, del «Índice de Convergencia con la Sociedad de la Información», o ICSI (Gimeno, 2014), y del porcentaje de hogares con acceso a Internet.

Según Dias (2016), el número de trabajos sobre estas temáticas indexados en *Scopus* y publicados, desde 2005 hasta 2014, por investigadores afiliados a instituciones portuguesas ha aumentado de manera constante, así como las citas recibidas por dichos trabajos. Concretamente, entre 2003 y 2017, el volumen total de documentos relacionados con e-gobierno en Portugal ascendió a 163 (Dias, 2019), una cifra nada desdeñable. Precisamente, uno de esos estudios es el llevado a cabo por Dias y Costa (2013), donde se señala que, en términos geográficos, la e-administración se encuentra en una fase de desarrollo más avanzada en los municipios del litoral, que son los más poblados y, en paralelo, los que presentan mejores indicadores socioeconómicos. En cuanto a la investigación de Nevado-Gil y Gallardo-Vázquez (2016), su objetivo es analizar el grado de información en materia de «Responsabilidad Social» (a saber: comportamiento económico, social y medioambiental) contenida en los portales de los 58 municipios que conforman la región del Alentejo, así como identificar sus posibles factores explicativos. Para ello, generan un «Índice de Divulgación sobre Responsabilidad Social» (IDRS), que alcanza un promedio bajo y que correlaciona de manera significativa con el tamaño de las entidades y su capacidad económica e institucional, mientras que nada tiene que ver con el nivel de competencia política.

En Portugal también se recaban, desde 2013, datos acerca del «Índice de Transparencia Municipal» por parte de *Transparência e Integridade – Associação Cívica* (2017). A diferencia de su homólogo español (el ITA), el ITM se configura en función de 76 indicadores agrupados en 7 dimensiones: 1) información sobre la organización, composición social y funcionamiento del municipio; 2) planes e informes; 3) impuestos, tasas, tarifas, precios y

reglamentos; 4) relación con la sociedad; 5) contratación pública; 6) transparencia económica y financiera; y 7) transparencia en el área de urbanismo. Más adelante, en los apartados de metodología y resultados, incidiremos de nuevo en el ITM.

Por último, otra de las iniciativas de mayor trayectoria en Portugal referente a la evaluación de los sitios web locales es la llevada a cabo, periódicamente, por Soares y otros (2017), que se concreta en el «Índice de Presença na Internet das Câmaras» (IPIC). Este indicador está creado a partir de 4 criterios: 1) contenido; 2) accesibilidad, navegabilidad y facilidad de uso; 3) servicios online; y, para terminar, 4) participación. En la explotación de sus datos, Soares y otros (2017) emplean una estrategia comparativa interregional que les induce a concluir que el litoral continental aventaja al interior y a las islas; que el Área Metropolitana de Lisboa es la NUTS II más destacada, seguida del Algarve y del Norte; y que, en lo referente a los distritos, Bragança, Lisboa, Vila Real y Setúbal son los mejor posicionados.¹

En síntesis, parece que los factores determinantes que motivan las brechas digitales son los que siguen (Manoharan e Ingrams, 2018; Nevado-Gil y Gallardo-Vázquez, 2019; Pirni y otros, 2019; Várallyai y otros, 2015): 1) geográficos (ubicación de los municipios en los ámbitos rural o urbano), 2) demográficos (densidad poblacional), 3) económicos (ingresos, presupuesto y PIB de los territorios), 4) tecnológicos (infraestructuras de acceso a Internet), y 5) digitales (la propia calidad de los contenidos y los servicios web). Por ende, y tras la revisión de la literatura científica previa, nos planteamos las siguientes hipótesis:

- H₁: existirán diferencias a nivel geográfico en cuanto a la calidad de los *websites* municipales de Portugal, siendo los del litoral continental los más destacados.
- H₂: cuanto mayor sea la población de las localidades, mejor serán sus servicios *on-line*.
- H₃: cuanto mayor capacidad económica tenga el concejo, mayor será la calidad de su portal web.

A efectos estadísticos, la H₃ se subdivide en varias hipótesis de segundo orden:

- H_{3a}: a mayor presupuesto municipal, mejor sitio web.
- H_{3b}: a mayor gasto cultural, mejor sitio web.
- H_{3c}: a mayores ingresos turísticos, mejor sitio web.

Surgen, en términos políticos y digitales, las restantes hipótesis del estudio:

- H₄: el signo político del ayuntamiento no afectará a la calidad del *website*.
- H₅: a medida que los indicadores de desarrollo electrónico-administrativo y digital aumenten, la calidad del portal web municipal también mejorará.

Esta última hipótesis se desdobra, asimismo, en otras dos:

- H_{5a}: a mayor transparencia, mejor portal web.
- H_{5b}: a mayor presencia en Internet, mejor portal web.

A continuación, pasamos a detallar los objetivos y la metodología empleada en esta investigación.

2. OBJETIVOS

El cometido principal del presente estudio consistió en diseñar un procedimiento que permitiera medir la calidad de los sitios web municipales de Portugal. Supeditados a este propósito inicial y estrechamente relacionados con las hipótesis descritas con anterioridad, surgen dos objetivos específicos:

- OE₁: detectar la existencia de brechas digitales interterritoriales en cuanto a la implementación de los portales web.
- OE₂: señalar los principales factores que condicionan la puesta en marcha de estas herramientas tecnológicas al servicio de la e-administración local portuguesa.

Si establecemos una correspondencia entre los objetivos y las hipótesis, el OE₁ se instrumentaliza gracias a la H₁, mientras que el OE₂ se desdobra en el resto de las hipótesis.

3. METODOLOGÍA

En aras del acometimiento de estos propósitos, se practicó un análisis de contenido sistemático, objetivo y cuantitativo (Lacy y otros, 2015; Neuen-dorf, 2017) de los $N = 308$ portales web oficiales; es decir, del universo poblacional de municipios, un hecho que afirma la representatividad de los datos recabados. El listado de localidades del país se consultó en el repositorio de la *Associação Nacional de Municípios Portugueses*.²

La principal herramienta de todo análisis de contenido es el libro de códigos, "un manual de instrucciones a la medida de la metodología de investigación concreta" (Piñuel, 2002: 19) en el que se reflejan, con detalle, las variables de análisis y sus respectivas categorías. Así pues, el *codebook*, inspirado en estudios previos similares (Campillo-Alhama, 2013; Cardoso-de-Miranda y Muñoz-Cañavate, 2015; Carvalho y otros, 2016; Dias y Costa,

2013; Fernández-Cavia y otros, 2014; Ingrams y otros, 2018; Manoharan y otros, 2017; Mota y Losada, 2018; Nevado-Gil y Gallardo-Vázquez, 2016; Rocha, 2012), estaba conformado por un total de 80 ítems, en su mayoría nominales dicotómicos (conocidos también como «dummy») o, en su defecto, multicategoriales. Estos se referían tanto a la información cultural y turística, como al diseño web (estética o «*look and feel*», arquitectura de la información, usabilidad y accesibilidad) y, finalmente, a los parámetros de la web 2.0 (presencia de las localidades en redes sociales, microblogs, plataformas de vídeo, etc.). Este manual de instrucciones fue empleado por dos evaluadores, que se ocuparon de la codificación de las 308 unidades de análisis desde el 16 de marzo hasta el 8 de junio de 2017.

Para el cálculo de la fiabilidad, uno de los pasos más críticos de todo el proceso, fue seleccionada una muestra aleatoria de ~12% de los casos; esto es, $n = 37$ sedes analizadas, en simultáneo, por ambos miembros del equipo. El parámetro utilizado para el chequeo de la confiabilidad fue el «*Alpha* de Krippendorff» que, de entre las más de 20 medidas existentes (Zhao y otros, 2013), es la más robusta a nivel estadístico y la que mejor se ajusta tanto a las distintas escalas como a un número indeterminado de codificadores (Hayes y Krippendorff, 2007; Krippendorff, 2011). A este respecto, la fiabilidad promedio arrojada por las 80 variables fue adecuada: $M(a_k) = 0.80$.

3.1. Medidas

Antes de proceder con el reporte de datos estadísticos, es preciso aludir al «Índice General de Calidad Web» (IGCW), concebido para calibrar el estado de los portales de los 308 municipios portugueses, y que vertebró el apartado de resultados. Para su creación, fueron tenidos en cuenta los siguientes parámetros:

1. «Indicador de Información Turística» (Piñeiro-Naval y otros, 2017), compuesto por 31 ítems que aluden a contenidos web sobre ocio y entretenimiento de los municipios ($M_{I2T} = 0.493$, $DT = 0.117$; $a_c = 0.63$).
2. «Indicador de Información sobre Patrimonio Cultural» (Piñeiro-Naval y Serra, 2019a), configurado con arreglo a 20 variables que versan sobre la herencia cultural de las localidades ($M_{I2PC} = 0.468$, $DT = 0.136$; $a_c = 0.60$).
3. «Índice de Calidad Formal» (Piñeiro-Naval y otros, 2018), organizado a partir de 25 ítems que miden el desempeño de las sedes web en materia de estética, arquitectura de la información y usabilidad ($M_{ICF} = 0.506$, $DT = 0.136$; $a_c = 0.73$).

4. «Indicador de Web Social» (Piñeiro-Naval y Serra, 2019_b), diseñado en función de 4 variables relacionadas con la tenencia de perfiles en medios sociales digitales por parte de los concejos ($M_{IWS} = 0.527$, $DT = 0.306$; $\alpha_c = 0.64$).

Así pues, la expresión alfanumérica que explica el IGCW es la siguiente:

5. $IGCW = (I2T + I2PC + ICF + IWS) / 4$; donde los indicadores previos tienen, como puede observarse, el mismo peso relativo (25%).

El IGCW ($\alpha_c = 0.58$) presenta, a su vez, los siguientes valores descriptivos para el conjunto de la muestra: $M_{IGCW} = 0.499$ ($DT = 0.125$), con un rango que oscila entre «0.162» y «0.798». Por otro lado, cabe referir que la consistencia interna promedio de los indicadores ($\alpha_c = 0.64$) alcanza un valor aceptable si tenemos en cuenta el carácter exploratorio del trabajo (Hair y otros, 2010).

Finalmente, en la matriz SPSS (versión 24) donde se reunieron los datos pertenecientes a los 308 sitios web (evaluados en función de las 80 variables del estudio), fueron también incluidas determinadas variables independientes de tipo geográfico, demográfico, económico, político, electrónico-administrativo y digital. Solo así hemos sido capaces de triangularlas (Denzin, 2012) con el IGCW. Las fuentes secundarias de información a partir de las cuales se extrajeron estos datos (todos ellos actualizados en 2017, año en el que se examinaron los *websites*) son:

- *Instituto Nacional de Estatística*, gracias al que hemos podido ubicar a cada municipio en su correspondiente distrito y NUTS II (variables geográficas), así como conocer su población (variable demográfica).³
- *Direção Geral das Autarquias Locais*, mediante la que sabemos el presupuesto municipal (variable económica).⁴
- *Pordata – Base de Dados Portugal Contemporâneo*, que nos permitió conocer el gasto en actividades culturales y los ingresos en materia turística de cada localidad (variables económicas).⁵
- *Secretaria Geral do Ministério da Administração Interna*, de donde obtuvimos el partido político al frente de cada ayuntamiento.⁶
- *Transparência e Integridade – Associação Cívica* (2017), portal del que recogimos los datos relativos al Índice de Transparencia Municipal (ITM) de los 308 sitios portugueses (una variable electrónico-administrativa).
- *IPIC* (Soares y otros, 2017), una variable de desarrollo digital.

En el siguiente apartado se detallan los principales resultados del trabajo.

4. RESULTADOS

El primer bloque de resultados se corresponde con las distintas comparaciones, ya sean de índole geográfica o política, efectuadas con arreglo al IGCW. En la Tabla I figuran los datos que han arrojado las regiones en que se ha dividido el territorio portugués en función de este indicador agregado:

Tabla 1. Comparaciones interregionales a tenor del IGCW (ANOVA)

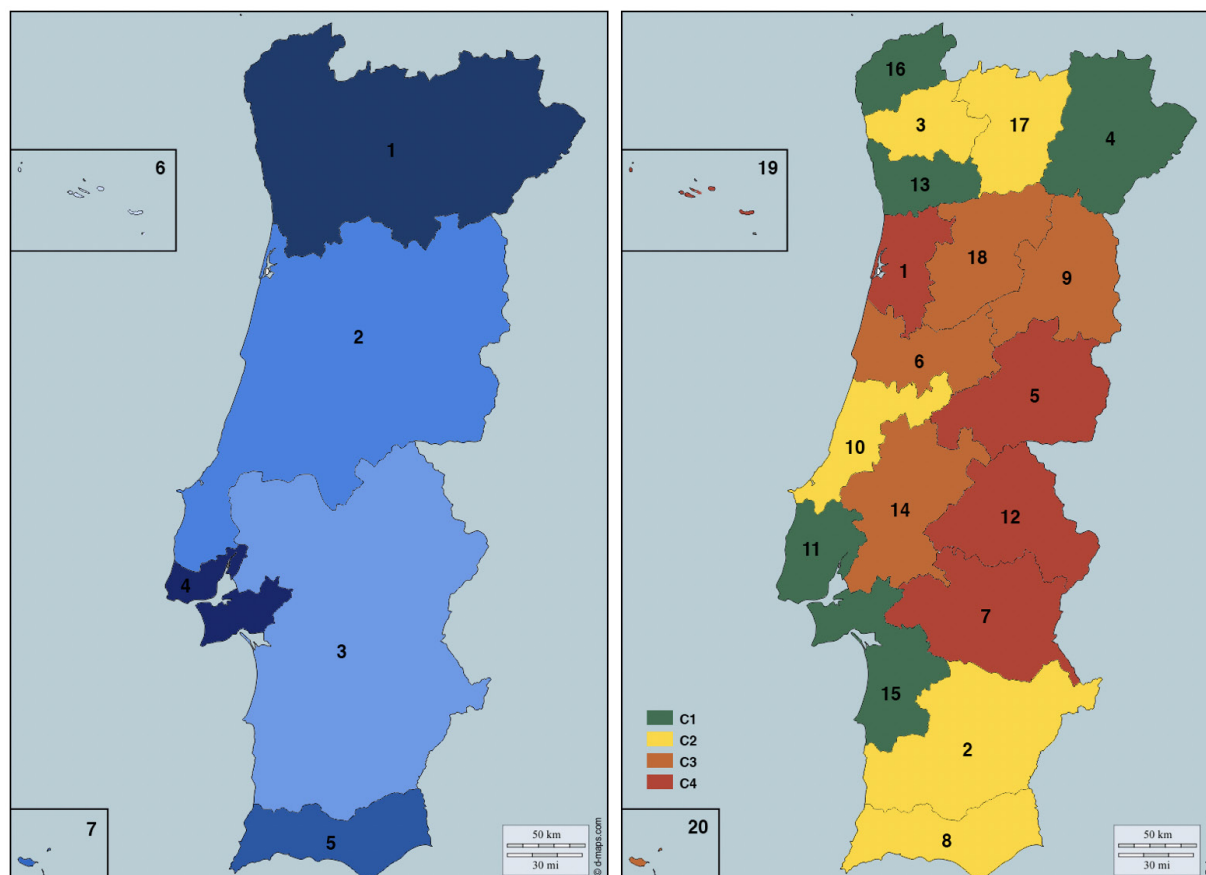
REGIONES	M_{IGCW}	DT	N
<i>Bloques</i>			
1. Litoral Continental	0.530	0.123	153
2. Interior	0.480	0.116	125
3. Islas	0.415	0.124	30
<i>NUTS II</i>			
1. Norte	0.541	0.128	86
2. Centro	0.478	0.109	100
3. Alentejo	0.474	0.107	58
4. Área de Lisboa	0.611	0.128	18
5. Algarve	0.516	0.107	16
6. Azores	0.378	0.121	19
7. Madeira	0.479	0.105	11
<i>Distritos–Regiones Autónomas</i>			
1. Aveiro	0.442	0.093	19
2. Beja	0.521	0.102	14
3. Braga	0.535	0.142	14
4. Bragança	0.569	0.136	12
5. Castelo Branco	0.462	0.090	11
6. Coimbra	0.495	0.105	17
7. Évora	0.441	0.107	14
8. Faro	0.516	0.107	16
9. Guarda	0.472	0.107	14
10. Leiria	0.507	0.126	16
11. Lisboa	0.553	0.134	16
12. Portalegre	0.455	0.096	15
13. Porto	0.565	0.123	18
14. Santarém	0.470	0.101	21
15. Setúbal	0.618	0.130	13
16. Viana do Castelo	0.603	0.089	10
17. Vila Real	0.504	0.121	14
18. Viseu	0.485	0.134	24
19. Açores	0.378	0.121	19
20. Madeira	0.479	0.105	11
Total	0.499	0.125	308

A partir de los datos reflejados en la Tabla I, podemos afirmar que se producen diferencias significativas en las tres comparaciones interregionales efectuadas. En primer lugar, y de los tres grandes bloques geográficos [$F_{IGCW \times \text{Bloques}} (2, 305) = 13.76$; $p < 0.001$; $\eta^2 = 0.083$], el más destacado es el litoral continental (H_1) que, asimismo, se diferencia de manera significativa tanto de su inmediato perseguidor, el interior [$t (276) = 3.39$; $p < 0.001$; $d = 0.434$] como, lógicamente, de las islas [$t (181) = 4.65$; $p < 0.001$; $d = 1.001$]; cuyos tamaños del efecto serían etiquetados, respectivamente, como «medio» y «elevado» (Cohen, 1988; Johnson y otros, 2008). Por su parte, el interior también aventajaría a las islas en su particular comparación [$t (153) = 2.75$; $p < 0.01$; $d = 0.608$]. Del mismo modo, podemos sostener, al nivel de las regiones NUTS II, que se dan importantes desajustes entre ellas [$F_{IGCW \times \text{NUTS II}} (6, 301) = 9.11$; $p < 0.001$; $\eta^2 = 0.154$], erigiéndose como la más destacada el Área Metropolitana de Lisboa, significativamente superior no solo a las Azores [$t (35) = 5.68$; $p < 0.001$; $d = 2$], la región más deprimida, sino también al

Norte [$t (102) = 2.11$; $p = 0.038$; $d = 0.583$], la segunda en importancia.

Los distritos también exhiben contrastes [$F_{IGCW \times \text{Distritos}} (19, 288) = 3.84$; $p < 0.001$; $\eta^2 = 0.203$], especialmente pronunciados entre Setúbal y las Azores [$t (30) = 5.33$; $p < 0.001$; $d = 1.918$]; es decir, los distritos con mayor y menor IGCW en sus sitios municipales. A este respecto, si estableciésemos como puntos de corte los cuartiles del IGCW de los distritos, el primero de los cuartiles (C1) tendría un rango de valores desde «0.618» hasta «0.553» (donde quedarían englobados, ordenadamente, Setúbal, Viana do Castelo, Bragança, Porto y Lisboa), el segundo (C2) desde «0.535» hasta «0.504» (con Braga, Beja, Faro, Leiria y Vila Real), el tercero (C3) desde «0.495» hasta «0.471» (con Coimbra, Viseu, Madeira, Guarda y Santarém), y el cuarto (C4) desde «0.462» hasta «0.378» (con Castelo Branco, Portalegre, Aveiro, Évora y Azores). De forma gráfica, pueden visualizarse en la Figura 1 tanto las regiones NUTS II del país como los 20 distritos, numerados atendiendo al orden de la Tabla I y agrupados según los cuartiles de su IGCW:

Figura 1. NUTS II y distritos portugueses (agrupados en los cuartiles del IGCW)



Finalmente, si descendemos a la unidad territorial de carácter más inmediato de todas cuantas hemos tenido en cuenta en este estudio, el listado «Top 5» de los municipios con mejor IGCW estaría formado por: Ponte de Lima (0.798), Alcochete (0.767), Montijo (0.758), Miranda do Douro (0.757) y Pombal (0.756).

Desde el punto de vista ideológico, y habiéndose identificado previamente el partido al frente de cada municipio⁷, se procedió a comparar aquellas localidades regidas por gobiernos de izquierdas ($M_{IGCW} = 0.494$; $DT = 0.128$) frente a derechas ($M_{IGCW} = 0.506$; $DT = 0.121$), llegando a la conclusión de que no se producen diferencias entre ambos [$t(306) = 0.78$; $p = 0.436$] (H_4).

En lo que respecta a las demás variables independientes tenidas en cuenta en este trabajo (de cariz demográfico, económico, electrónico-administrativo y digital), pueden verse en la Figura 2 los valores relativos a su correlación con el IGCW. Como se puede apreciar en la Figura II, todas las correlaciones son estadísticamente significativas, destacando las que se establecen entre el IGCW y la calidad percibida⁸ [$r(306) = 0.477$; $p < 0.001$], el IPIC [$r(306) = 0.381$; $p < 0.001$] (H_{5b}), el presu-

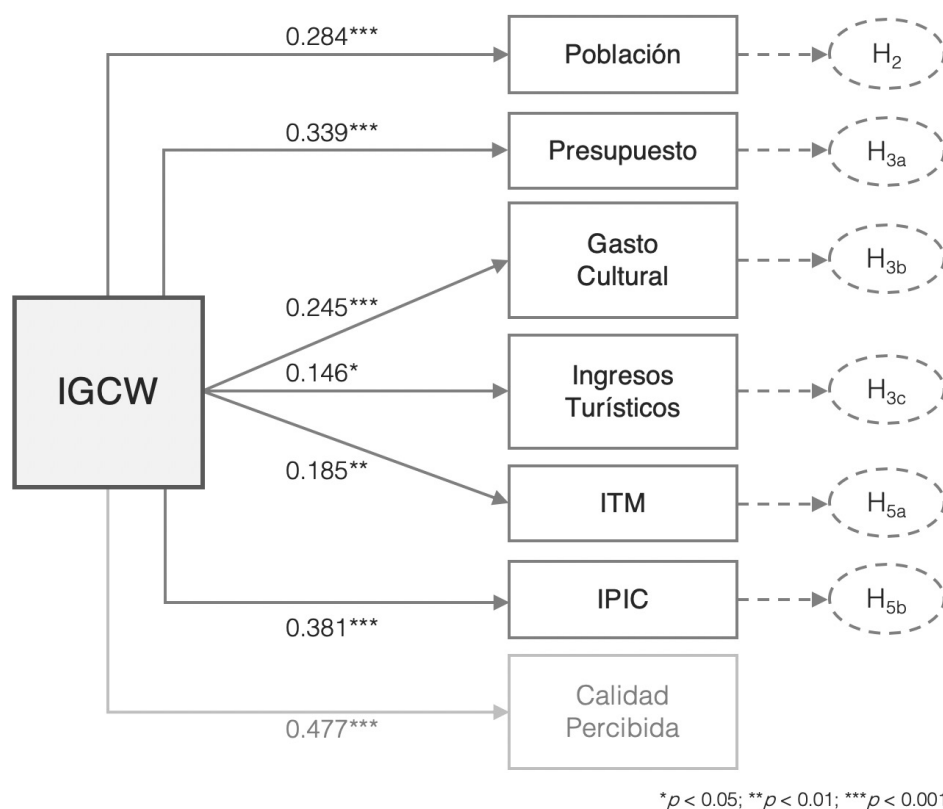
puesto municipal [$r(306) = 0.339$; $p < 0.001$] (H_{3a}), la población de la localidad [$r(306) = 0.284$; $p < 0.001$] (H_2) y, por último, el gasto en materia cultural [$r(306) = 0.245$; $p < 0.001$] (H_{3b}). Como complemento de estas correlaciones bivariadas, repárese en el siguiente análisis de regresión lineal múltiple, en el que se mide la dependencia entre el IGCW y las distintas variables predictoras en su conjunto:

Tabla II. Factores predictores del IGCW (Análisis de Regresión Múltiple)

FACTORES PREDICTORES	TOLERANCIA	FIV	β
Presupuesto	0.213	4.704	0.235*
Población	0.126	7.947	-0.051
Gasto Cultural	0.173	5.770	0.012
Ingresos Turísticos	0.321	3.118	0.041
ITM	0.895	1.117	0.034
IPIC	0.855	1.170	0.257***
Calidad Percibida	0.933	1.071	0.391***

FIV: Factor de Incremento de la Varianza. * $p < 0.05$, *** $p < 0.001$.

Figura 2. Asociación entre el IGCW y los factores predictores (r de Pearson)



El modelo ejecutado en la Tabla II resultó estadísticamente significativo [$F(7, 300) = 24.66$; $p < 0.001$], explicando las 7 variables predictoras, para el conjunto de la muestra ($N = 308$), el 35.1% de la varianza ($R^2_{ajustada} = 0.351$). Se observa que el factor predictor que más explica el IGCW es la variable inherente al estudio «calidad percibida» ($\beta = 0.391$; $p < 0.001$), seguida del IPIC ($\beta = 0.257$; $p < 0.001$) y del presupuesto municipal ($\beta = 0.235$; $p = 0.019$), justamente las que correlacionan en mayor medida con el indicador creado (véase la Figura 2).

5. DISCUSIÓN Y CONCLUSIONES

A tenor de los datos empíricos obtenidos, podemos deducir que el estado de los sitios web municipales de Portugal se aproxima a un nivel «correcto», ya que el IGCW creado arroja un promedio muy cercano al aprobado. No obstante, las organizaciones municipales, “formadas por recursos humanos, materiales, técnicos y sistemas de información que conectan las áreas, departamentos, servicios y negociados integrados en tales estructuras políticas y administrativas” (Campillo-Alhama, 2013), deben seguir insistiendo en una adopción más explícita y efectiva de las TIC y, en concreto, poner a disposición de sus ciudadanos unos portales web que deberían ser sustancialmente mejorados.

Esta tarea es más necesaria, si cabe, en el seno de aquellas localidades ubicadas tanto en el interior del país como en las islas, pues los resultados evidencian una clara persistencia de las brechas digitales entre estos bloques geográficos y el litoral continental. ¿A qué se pueden deber estos desajustes? En efecto, la literatura científica previa ya había señalado algunos factores predictivos de muy diversa naturaleza, como por caso: 1) los apriorísticos, es decir, la ubicación geográfica de las localidades; 2) los estructurales, como el tamaño poblacional y la capacidad económica; y 3) los internos, como el desarrollo tecnológico y administrativo. Por lo tanto, Portugal se alinea con otros países de su entorno, donde estos elementos desempeñan un rol fundamental en la correcta implementación de los servicios públicos electrónicos. Salemin y otros (2017) ya habían identificado un agregado de 157 manuscritos que ilustraban las diferencias entre el ámbito urbano y rural en cuanto al aprovechamiento de las TIC. En sintonía con las evidencias empíricas previas, la investigación que nos ocupa vendría a engrosar esta nómina de artículos.

Véase, a modo de resumen, la Tabla III con la descripción de las hipótesis del estudio y su consiguiente resolución:

Tabla III. Resumen de las hipótesis del estudio

HIPÓTESIS	DESCRIPCIÓN	RESOLUCIÓN
H ₁	IGCW Litoral Continental > Interior > Islas	✓
H ₂	+ Población $\Rightarrow$ + IGCW	✓
H _{3a}	+ Presupuesto $\Rightarrow$ + IGCW	✓
H _{3b}	+ Gasto Cultural $\Rightarrow$ + IGCW	✓
H _{3c}	+ Ingresos Turísticos $\Rightarrow$ + IGCW	✓
H ₄	IGCW Partidos Izquierdas $\approx$ Derechas	✓
H _{5a}	+ Transparencia $\Rightarrow$ + IGCW	✓
H _{5b}	+ Presencia en Internet $\Rightarrow$ + IGCW	✓

Con relación a la brecha digital, la consciencia acerca de los beneficios que reportan las TIC sigue siendo, por norma general, bastante débil entre los responsables políticos de las regiones menos desarrolladas (Kuk, 2002), independientemente de su tendencia ideológica. En paralelo, los altos niveles de desempleo y la precariedad laboral en dichas zonas conllevan que no todos los ciudadanos gocen de los medios suficientes y necesarios para adquirir dispositivos tecnológicos (ordenadores, tabletas, *smartphones*, etc.) y usarlos, con asiduidad, en su acceso a Internet.

Por otro lado, si nos remitimos al significativo concepto de «sostenibilidad», llegamos a la conclusión de que un e-gobierno sostenible debe incorporar la tecnología de última generación para brindar un servicio eficaz, robusto, que ahorre costes y que fomente la participación activa y la satisfacción de los usuarios (Joshi e Islam, 2018). En consecuencia, los sitios web municipales, en tanto que dispositivos articuladores de los procesos y las estrategias e-administrativas, deben adecuarse a unos estándares de calidad lo suficientemente elevados como para corresponder a las demandas no solo de sus ciudadanos, sino también de potenciales usuarios foráneos como, por ejemplo, los turistas.

En el plano de las limitaciones del estudio (y al margen de que muchos sitios web ya habrán cambiado desde la recogida de los datos), cabe reconocer que el IGCW se apoya, a nivel de contenido, en la comunicación cultural y turística, dejando de lado cuestiones de peso como la política o la economía. Del mismo modo, habrá, con certeza, otros factores que expliquen el estado de los portales web, como el grado de desempeño de los funcionarios (si los hubiera) encargados de su gestión y mantenimiento, o el profesio-

nalismo de los diseñadores del sistema informático (ambos sujetos sometidos al músculo financiero de la institución). Así pues, en futuros trabajos será preciso tener en cuenta más elementos para conformar nuevos indicadores agregados que se ciñan, con mayor precisión, al constructo multidimensional de «calidad web», así como incluir más variables independientes para completar un modelo que explique mejor el estado de la administración local portuguesa, cuyos esfuerzos en materia tecnológica y digital deberán aumentar a corto y medio plazo.

6. AGRADECIMIENTOS

Este trabajo ha sido financiado por la *Fundação para a Ciência e a Tecnologia* de Portugal (Ref.: SFRH/BPD/122402/2016). Queremos también agradecer a Rafael Mangana su colaboración en la recogida de datos.

ACKNOWLEDGEMENTS

This work was funded by the Portuguese Foundation for Science and Technology (Ref.: SFRH/BPD/122402/2016). We also want to thank Rafael Mangana for his collaboration in collecting data.

8. REFERENCIAS

- Campillo-Alhama, C. (2013). Gestión de la información y su evaluación en los gabinetes de comunicación municipal. *El profesional de la información*, 22 (6), 515-521. <https://doi.org/10.3145/epi.2013.nov.03>
- Cardoso-de-Miranda, E. A.; Muñoz-Cañavate, A. (2015). Los sitios web como servicios de información al ciudadano: un estudio sobre los 308 ayuntamientos de Portugal. *Anales de Documentación*, 18 (1). <https://doi.org/10.6018/analesdoc.18.1.212681>
- Carvalho, R. M. D.; Lopes, P. F.; Alexandre, I.; Alturas, B. (2016). Qualidade dos sítios Web da Administração Pública Portuguesa. *RISTI*, 20, 78-98. <https://doi.org/10.17013/risti.20.78-98>
- Choudrie, J.; Ghinea, G.; Songonuga, V. N. (2013). Silver Surfers, E-government and the Digital Divide: An Exploratory Study of UK Local Authority Websites and Older Citizens. *Interacting with Computers*, 25 (6), 417-442. <https://doi.org/10.1093/iwc/iws020>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd Ed.)*. Hillsdale, New Jersey: Lawrence Earlbaum Associates.
- Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research*, 6 (2), 80-88. <https://doi.org/10.1177/1558689812437186>
- Detlor, B.; Hupfer, M. E.; Ruhi, U. (2010). Internal factors affecting the adoption and use of government websites. *Electronic Government, An International Journal*, 7 (2), 120-136. <https://doi.org/10.1504/EG.2010.030923>
- Dias, G. P.; Costa, M. (2013). Significant socio-economic factors for local e-government development in Portugal. *Electronic Government, An International Journal*, 10 (3-4), 284-309. <https://doi.org/10.1504/EG.2013.058785>
- Dias, G. P. (2016). A decade of Portuguese research in e-government: evolution, current standing, and ways forward. *Electronic Government, An International Journal*, 12 (3), 201-222. <https://doi.org/10.1504/EG.2016.078415>
- Dias, G. P. (2019). Fifteen years of e-government research in Ibero-America: A bibliometric analysis. *Government Information Quarterly*, 36 (3), 400-411. <https://doi.org/10.1016/j.giq.2019.05.008>
- Ebberts, W. E.; Jansen, M. G. M.; Deursen, A. J. A. M. van (2016). Impact of the digital divide on e-government: Expanding from channel choice to channel usage. *Government Information Quarterly*, 33 (4), 685-692. <https://doi.org/10.1016/j.giq.2016.08.007>
- Fernández-Cavia, J.; Rovira, C.; Díaz-Luque, P.; Cavaller, V. (2014). Web Quality Index (WQI) for official tourist destination websites. Proposal for an assessment system. *Tourism Management Perspectives*, 9, 5-13. <https://doi.org/10.1016/j.tmp.2013.10.003>
- García-Santamaría, J. V.; Martín-Matallana, J. (2017). La transparencia municipal en España: análisis de los factores que más influyen en el grado de transparencia. *Revista Latina de Comunicación Social*, 72, 1148-1164. <https://doi.org/10.4185/RLCS-2017-1212>
- Gimeno, M. (2014). *e-España. Informe anual sobre el desarrollo de la sociedad de la información en España*. Madrid: Fundación Orange.
- Hair, J. F.; Black, W. C.; Babin, B. J.; Anderson, R. E. (2010). *Multivariate data analysis (7th Ed.)*. Upper Saddle River, New Jersey: Pearson Prentice Hall.

7. NOTAS

- ¹ NUTS es el acrónimo de «Nomenclatura de las Unidades Territoriales Estadísticas».
- ² Información disponible en: <https://www.anmp.pt> [Fecha de consulta: 20/02/2020].
- ³ Información disponible en: <https://ine.pt> [Fecha de consulta: 20/02/2020].
- ⁴ Información disponible en: <http://www.portalautarquico.dgal.gov.pt> [Fecha de consulta: 20/02/2020].
- ⁵ Información disponible en: <https://www.pordata.pt> [Fecha de consulta: 20/02/2020].
- ⁶ Información disponible en: <https://www.sg.mai.gov.pt> [Fecha de consulta: 20/02/2020].
- ⁷ Para clasificar los ayuntamientos a nivel ideológico, se recodificó la variable multicategórica inicial «partido político» en otra dicotómica llamada «ideología de gobierno», cuyos nuevos valores son: 1 = izquierda y 2 = derecha. Se incluyeron en la izquierda al Partido Socialista, Partido Comunista y Partido Ecologista, y a los Grupos de Ciudadanos independientes; mientras que, a la derecha, fueron considerados el Partido Social Demócrata y el Partido Popular.
- ⁸ La «calidad percibida del sitio web» ($\alpha_k = 0.84$) es un ítem inherente al estudio y se empleó como una suerte de variable control. Se rige mediante una escala tipo Likert, donde: 1 = pésimo, 2 = mediocre, 3 = correcto, 4 = bueno y 5 = óptimo (Rocha, 2012). Su promedio es: $M = 2.67$ ($DT = 0.79$), significativamente inferior a la media teórica (valor «3») de su escala [$t(308) = -7.37, p < 0.001$].

- Hayes, A. F.; Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1 (1), 77-89. <https://doi.org/10.1080/19312450709336664>
- Helbig, N.; Gil-García, J. R.; Ferro, E. (2009). Understanding the complexity of electronic government: Implications from the digital divide literature. *Government Information Quarterly*, 26 (1), 89-97. <https://doi.org/10.1016/j.giq.2008.05.004>
- Ingrams, A.; Manoharan, A.; Schmidhuber, L.; Holzer, M. (2018). Stages and Determinants of E-Government Development: A Twelve-Year Longitudinal Study of Global Cities. *International Public Management Journal*, 0 (0), 1-39. <https://doi.org/10.1080/10967494.2018.1467987>
- Johnson, B. T.; Scott-Sheldon, L. A. J.; Snyder, L. B.; Noar, S. M.; Huedo-Medina, T. B. (2008). Contemporary approaches to meta-analysis in communication research. En: Hayes, A. F.; Slater, M. D.; Snyder, L. B. (Eds.), *The SAGE Sourcebook of Advanced Data Analysis Methods for Communication Research*, pp. 311-347. Thousand Oaks, California: Sage. <https://doi.org/10.4135/9781452272054.n11>
- Joshi, P. R.; Islam, S. (2018). E-Government Maturity Model for Sustainable E-Government Services from the Perspective of Developing Countries. *Sustainability*, 10 (6), 1882. <https://doi.org/10.3390/su10061882>
- Karabasevic, D.; Stanujkic, D.; Maksimovic, M.; Popovic, G.; Momcilovic, O. (2019). An Approach to Evaluating the Quality of Websites Based on the Weighted Sum Preferred Levels of Performances Method. *Acta Polytechnica Hungarica*, 16 (5), 195-215. <https://doi.org/10.12700/APH.16.5.2019.5.11>
- Krippendorff, K. (2011). Agreement and information in the reliability of coding. *Communication Methods and Measures*, 5 (2), 93-112. <https://doi.org/10.1080/19312458.2011.568376>
- Kuk, G. (2002). The digital divide and the quality of electronic service delivery in local government in the United Kingdom. *Government Information Quarterly*, 20, 353-363. <https://doi.org/10.1016/j.giq.2003.08.004>
- Lacy, S.; Watson, B. R.; Riffe, D.; Lovejoy, J. (2015). Issues and best practices in content analysis. *Journalism & Mass Communication Quarterly*, 92 (4), 791-811. <https://doi.org/10.1177/1077699015607338>
- López-López, V.; Iglesias-Antelo, S.; Vázquez-Sanmartín, A.; Connolly, R.; Bannister, F. (2018). E-Government, Transparency & Reputation: An Empirical Study of Spanish Local Government. *Information Systems Management*, 35 (4), 276-293. <https://doi.org/10.1080/10580530.2018.1503792>
- Manoharan, A. P.; Ingrams, A. (2018). Conceptualizing E-Government from Local Government Perspectives. *State and Local Government Review*, 50 (1), 56-66. <https://doi.org/10.1177/0160323X18763964>
- Manoharan, A. P.; Zheng, Y.; Melitski, J. (2017). Global comparative municipal e-governance: factors and trends. *International Review of Public Administration*, 22 (1), 14-31. <https://doi.org/10.1080/12294659.2017.1292031>
- Mota, G.; Losada, N. (2018). Promoção Turística nos Websites Municipais: O Caso da Região do Douro (NUT III). *Revista Portuguesa de Estudos Regionais*, 47, 49-71.
- Neuendorf, K. A. (2017). *The content analysis guidebook* (2nd Ed.). Thousand Oaks, California: Sage. <https://doi.org/10.4135/9781071802878>
- Nevado-Gil, M. T.; Gallardo-Vázquez, D. (2016). Información sobre Responsabilidad Social contenida en las páginas webs de los ayuntamientos. Estudio en la región del Alentejo. *Revista Española de Documentación Científica*, 39 (4), e150. <https://doi.org/10.3989/redc.2016.4.1353>
- Nevado-Gil, M. T.; Gallardo-Vázquez, D. (2019). Local government social responsibility: empirical evidence in the region of Extremadura. *Investigaciones Regionales. Journal of Regional Research*, 45, 161-179.
- Philip, L.; Cottrill, V.; Farrington, J.; Williams, F.; Ashmore, F. (2017). The digital divide: Patterns, policy and scenarios for connecting the 'final few' in rural communities across Great Britain. *Journal of Rural Studies*, 54, 386-398. <https://doi.org/10.1016/j.jrurstud.2016.12.002>
- Piñeiro-Naval, V.; Igartua, J. J.; Sánchez-Nuevo, L. A. (2017). Identificación de factores externos que predicen el estado de los sitios web municipales: una aproximación regional al caso de España. *Observatorio (OBS*)*, 11 (3), 44-60.
- Piñeiro-Naval, V.; Mangana, R.; Serra, P. (2018). Validación del Índice de Calidad Formal como modelo para la evaluación de websites: el caso de la e-Administración local portuguesa. *Transinformação*, 30 (2), 153-165. <https://doi.org/10.1590/2318-08892018000200002>
- Piñeiro-Naval, V.; Serra, P. (2019a). How Do Destinations Frame Cultural Heritage? Content Analysis of Portugal's Municipal Websites. *Sustainability*, 11 (4), 947. <https://doi.org/10.3390/su11040947>
- Piñeiro-Naval, V.; Serra, P. (2019b). Las instituciones municipales portuguesas y su adopción de las tecnologías 2.0: propuesta de un Indicador de Web Social. *Investigación Bibliotecológica*, 33 (78), 13-28. <https://doi.org/10.22201/iibi.24488321xe.2019.78.57957>
- Piñeiro-Naval, V.; Serra, P.; Mangana, R. (2017). Desarrollo local y turismo. El impacto socioeconómico de la comunicación digital en Portugal. *Revista Latina de Comunicación Social*, 72, 1515-1535. <https://doi.org/10.4185/RLCS-2017-1232>
- Piñuel, J. L. (2002). Epistemología, metodología y técnicas del análisis de contenido. *Estudios de Sociolingüística*, 3 (1), 1-42.
- Pirni, A.; Giampellegrini, P.; Raffini, L. (2019). Digital transformation and e-government. For a research agenda on the Liguria Region. *OBETS. Revista de Ciencias Sociales*, 14 (2), 471-490. <https://doi.org/10.14198/OBETS2019.14.2.07>
- Rocha, Á. (2012). Framework for a Global Quality Evaluation of a Website. *Online Information Review*, 36 (3), 374-382. <https://doi.org/10.1108/14684521211241404>
- Salemink, K.; Strijker, D.; Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. *Journal of Rural Studies*, 54, 360-371. <https://doi.org/10.1016/j.jrurstud.2015.09.001>
- Soares, D. S.; Amaral, L.; Ferreira, L. M. (2017). *Presença na Internet das Câmaras Municipais Portuguesas em 2016: Estudo sobre Local e-Government em Portugal*. Universidade do Minho, Guimarães: Gávea-Observatório da Sociedade da Informação.

- Transparência e Integridade – Associação Cívica (2017). Índice de Transparência Municipal (ITM). Disponible en: <https://transparencia.pt/itm> [Fecha de consulta: 20/09/2019].
- Transparencia Internacional España (2017). Índice de Transparencia de los Ayuntamientos (ITA). Disponible en: <https://transparencia.org.es/ita-2017> [Fecha de consulta: 21/04/2020].
- Várallyai, L.; Herdon, M.; Botos, S. (2015). Statistical analyses of digital divide factors. *Procedia Economics and Finance*, 19, 364-372. [https://doi.org/10.1016/S2212-5671\(15\)00037-4](https://doi.org/10.1016/S2212-5671(15)00037-4)
- Zhao, X.; Liu, J. S.; Deng, K. (2013). Assumptions behind Inter-coder Reliability Indices. *Annals of the International Communication Association*, 36 (1), 419-480. <https://doi.org/10.1080/23808985.2013.11679142>

ESTUDIOS / RESEARCH STUDIES

La transformación digital en los métodos y temas de la investigación brasileña de Información y Documentación 2010-2019

Francisco Carlos Paletta*, José-Antonio Moreiro-González**

*Universidade de São Paulo, Departamento de Biblioteconomia y Documentación. Universidad Carlos III de Madrid
Correo-e: fcpaletta@usp.br | ORCID iD: <http://orcid.org/0000-0002-4112-5198>

**Departamento de Biblioteconomia y Documentación. Universidad Carlos III de Madrid
Correo-e: jamore@bib.uc3m.es ORCID | iD: <http://orcid.org/0000-0002-8827-158X>

Recibido: 23-03-20; 2ª versión: 24-04-20; Aceptado: 07-05-20; Publicado: 04-05-2021

Cómo citar este artículo/Citation: Paletta, F. C.; Moreiro-González, J. A. (2021). La transformación digital en los métodos y temas de la investigación brasileña de Información y Documentación 2010-2019. *Revista Española de Documentación Científica*, 44 (2), e293. <https://doi.org/10.3989/redc.2021.2.1763>

Resumen: A partir de la base de datos *Social Sciences Citation Index* de WoS se identifican los artículos científicos brasileños que atienden a temas y métodos relacionados con la transformación digital en Información y Documentación, en el intervalo 2010-2019. Con los títulos y palabras clave de los 1039 artículos recuperados se elaboran tablas dinámicas. Luego los títulos se categorizan por expertos, mientras que para las *keywords* se usa el modelo de n-grams con el filtro de Fruchterman-Reingold. Sus coocurrencias se visualizan en clústeres elaborados con el método Louvain de detección de comunidades. Los resultados se exponen en tablas y gráficos comentados sobre la internacionalización, coautoría e interdisciplinariedad de los artículos, junto a las técnicas de investigación seguidas en ellos. Se estudia la progresión temática de los que se refieren al ambiente y la transformación digital. Se evidencia una aportación numerosa y significativa de los autores brasileños al estudio de los efectos que los escenarios digitales causan en el sector de la información.

Palabras clave: análisis de materias; análisis de contenido; artículos de revista; Información y Documentación; transformación digital; SSCI-WoS; Brasil.

Digital Transformation in the methods and subjects of the Brazilian Library and Information Science research (2010-2019)

Abstract: From the *Social Sciences Citation Index* database, those Brazilian scientific articles that address issues, methods and consequences of digital transformation into LIS, in the 2010-2019 interval are identified. Dynamic tables are elaborated with the titles and keywords of the 1039 retrieved articles. The categorization of the titles is done by experts, while for the keywords, the n-grams model is used along with the Fruchterman-Reingold filter. Their co-occurrences are visualized in clusters elaborated with the Louvain method of community detection. The results are presented in tables and commented graphics on the internationalization, co-authorship and interdisciplinarity of the articles, together with the research techniques followed in them. In The thematic progression of those that refer to the Digital environment is studied. An abundant and significant contribution of the Brazilian authors to the study of the effects that the digital scenarios cause in the information sector is observed.

Keywords: subject analysis; content analysis; SSCI-WoS; Brazil; journal articles; Library and Information Science; digital transformation.

Copyright: © 2021 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1. INTRODUCCIÓN

Entorno y transformación digitales son dos conceptos provenientes del mundo empresarial, donde la actividad digital les da ventajas por la inmediatez, cantidad, ubicuidad y uso multimedia de información. El impulso de las Tecnologías de la Información y Comunicación (TIC) ha trascendido desde las empresas a la salud, el trabajo, los desafíos sociales, la cultura, la educación y la vida personal. Asimismo, su continuo avance afecta al intercambio continuado de bienes del conocimiento, pues de acuerdo con Fleischmann (2007), ninguna actividad es ajena a la cantidad de información accesible, ni a los cambios en su creación, difusión y uso. Aquí está la mayor diferencia de nuestro mundo con los periodos históricos precedentes, que llegan a parecernos carentes de información. Igual que sucede en las empresas, la integración de las tecnologías en todos los aspectos del funcionamiento y gestión de los centros de información hace que su intervención alcance a los procedimientos técnicos, a la atención y comunicación con los usuarios y con otros profesionales, incluso a la oferta de servicios y a los nuevos compromisos y tareas gracias al análisis de datos (Matt y otros 2015). De modo que la transformación no se limita a la explotación e integración de las TIC, pues afecta a todos los aspectos de la vida profesional y tiene consecuencias en los métodos que se emplean para investigar y trabajar.

Las consecuencias que esta transformación tiene en *Library and Information Science* (LIS) se deben de estudiar desde una triple perspectiva. Primero, desde la investigación que la analiza, genera nuevo conocimiento y lo aplica a la solución de los más variados problemas; pasando por la docencia para transmitir comprensión de la situación y preparar o actualizar a los profesionales en las competencias y destrezas requeridas para desarrollar sus cometidos; hasta llegar, de manera especial, al rastreo e identificación de las circunstancias y condiciones de la actuación profesional. Consideramos que investigar los efectos de la transformación digital permite percibir, antes que nada, los avances y aplicaciones consecuentes a ella. Sin duda, los investigadores son los primeros en articular con flexibilidad los conocimientos cuando la condición transdisciplinar afecta a la propia naturaleza de la información y a su empleo y manejo, lejos de cualquier división operativa y clasificatoria del conocimiento. Otra razón para indagar se deriva de que en Brasil se impulse desde 1970 la calidad de las investigaciones y su comunicación científica con la mejora de sus postgrados (Población y De-Oliveira, 2006). Lo que, sin duda, está mediatizado por el contexto digital.

2. REVISIÓN DE LITERATURA

Diferentes estudios argumentan el empleo del análisis de contenido para conocer las características de la producción científica por categorías académicas y por países. Desde luego, como método que cuida la selección de los documentos a analizar, el establecimiento posterior de categorías para organizar los resultados y la necesaria garantía de confiabilidad. Así lo confirman Piñuel (2002) o White y Marsh (2006). Mientras que son antecedentes del empleo de artículos de revista para conocer los temas investigados en LIS las aportaciones de Järvelin y Vakkari (1993) o de Ferrán y otros, (2017) que revelan, desde el análisis de contenido, los cambios de interés que orientan la investigación en Biblioteconomía y Ciencia de la Información. O la de Pettigrew y McKechnie (2001) para saber cómo emplean la teoría los autores de los artículos y comprobar que la mayoría de las teorías manejadas provienen de las ciencias sociales, seguido de la propia LIS, las ciencias y las humanidades. Mientras que, con el mismo fin, pero mediante análisis de citas, Åström (2007) estudia los cambios habidos durante quince años en los frentes de investigación, que estaban causados por la creciente dependencia de la tecnología que llevaría a LIS hacia la integración de sus diferentes espacios de investigación. Ya en el caso brasileño, Pinheiro y otros (2005) realizan análisis de contenido de los artículos de la revista *Ciência da Informação*, mientras que Bufrem y otros (2007) observan la temática atendida en las revistas brasileñas. Indirectamente, desde los currículos alojados en la Plataforma Lattes, Arboit y otros (2011) estudian los artículos, libros, capítulos y comunicaciones a congresos hechos en el exterior por investigadores brasileños, para conocer el contexto de la colaboración institucional. Más cercano a nuestros propósitos se situó el estudio de la producción científica brasileña en las revistas de Información y Documentación indizadas en WoS para identificar las más relevantes (Pinto y otros, 2016).

La comunicación académica tiene que manifestar con claridad que los cambios en las líneas y temas de investigación se pueden corresponder con los principales enfoques de la práctica profesional (Hurd, 2000). Consideramos que los artículos proporcionan conceptos fundamentados sobre los que se relacionan mejor los conocimientos y habilidades requeridos en la actuación profesional, e incluso muestran términos que identifican actividades y procesos, hasta reflejar la consolidación de tendencias.

3. OBJETIVOS

Se quieren seguir los efectos de la transformación digital desde los artículos de publicaciones periódicas. Para conseguirlo, se propone:

- Identificar y seleccionar la contribución de los investigadores brasileños al estudio de los entornos digitales de la información, tanto desde el punto de vista teórico como práctico.
- Percibir los enfoques con que los estudiosos analizan e interpretan la complejidad de la realidad enlazada.
- Valorar los datos del análisis para comprobar la repercusión de la información digital, así como la interdisciplinariedad e internacionalización de sus estudios.
- Desarrollar, en fase posterior, los resultados con las consecuencias que la transformación digital tiene para el mercado de trabajo y la formación universitaria.

4. METODOLOGÍA

4.1. Fuente de los datos

Se maneja información obtenida en la colección principal de la base de datos bibliográfica *Social Sciences Citation Index* (SSCI) de la *Web of Science* (WoS), a partir de una búsqueda que combina la producción en el área de *Information Science & Library Science* con la limitación geográfica de Brasil como país de los investigadores participantes en la elaboración de los artículos indizados entre los años 2010 y 2019.

Utilizar los índices de citas de SSCI-WoS para analizar la productividad brasileña en Información y Documentación proporciona seguridad en la relevancia de los documentos extraídos, pues los contenidos de las revistas indizadas fueron antes revisados por pares, por lo que resulta un buen muestrario para valorar los ajustes que la economía y los procesos de digitalización han supuesto en la conceptualización, las aplicaciones y las tareas de la información (De-Oliveira, 2018). Por otra parte, los artículos de revistas indizados en WoS aseguran una comunicación científica de prestigio e impacto (Jacsó, 2009), lo que aumenta las posibilidades de aprovechamiento en los trabajos de carácter aplicado. Debido a la propagación del análisis de revistas con manifiesta calidad, SSCI es un medio oportuno de acceso a la producción en revistas de las diferentes especialidades al permitir su selección por disciplinas y aportar información sobre la adscripción institucional de todos los autores de los artículos (De-Filippo, 2013), por lo que se establece como el procedimiento preferible para comprobar el estado de la actividad investigadora en Brasil (Marteletto, 2009) a la hora de reflejar la transcendencia epistemológica y práctica que está teniendo lo digital. Que el trabajo se limite a los artículos indizados manifiesta que los resultados obtenidos no son exhaustivos respecto al objeto

estudiado. Ciertamente es que sería aconsejable prestar atención a libros, comunicaciones, tesis doctorales y otros trabajos de fin de ciclo. Pero hacer esto sería una tarea más larga y no aseguraría la exactitud de la consulta que nos proporcionan las fuentes empleadas en este análisis.

4.2. Procesamiento de la información

La descarga se efectúa el día 18/06/2019. Se obtienen 1039 artículos de los que se extraen los títulos y palabras clave asignadas por los autores. No así las *KeyWords Plus* ofrecidas por el sistema de indización, por tratarse de palabras de interés secundario, que aparecen con frecuencia en los títulos de las referencias de los artículos, cuando una parte de nuestro análisis atiende a los propios títulos de los artículos. Además, las *KeyWords Plus* suelen ser palabras simples, muchas veces de carácter genérico con bastante ambigüedad. Para limpiar los títulos se contrastan con una lista de palabras vacías y otra de genéricas. Luego, los términos se fijan como *n_grams*, bigrams y trigrams. Los listados obtenidos se vacían en tablas dinámicas que permiten comparar el contexto de aparición de los términos en la muestra seleccionada. Para estos procesos se usa el gestor de datos Voyant (voyant-tools.org).

Ante la diversidad de las denominaciones causada por la libre asignación tanto de los títulos como de las palabras-clave y, para evitar las posibles indefiniciones, los términos extraídos se validan mediante contraste de experto con las preferencias y condiciones buscadas, lo que supone realizar una extracción híbrida (Vivaldi y Rodríguez, 2007). Se hace un recuento distinto para los títulos y para las palabras clave. El cálculo de aparición de éstas tiene en cuenta su mayor número o superordenación. Para precisar su significación hay que identificar los sinónimos, evitar las variaciones gráficas y reducir las variables al término canónico. Se excluyen del recuento los términos específicos cuando coinciden en el mismo título con un término genérico. Asimismo, cuando la combinación de los términos contenidos en los títulos y en las palabras clave no ofrece información decisoria respecto a su relevancia, se consultan los resúmenes de los artículos e incluso su contenido completo. Al tiempo que se comprueba el contexto en que se sitúan los términos simples. Las palabras-clave referentes a métodos y técnicas de investigación se clasifican de manera inductiva en las tablas VI y VII hasta formar las categorías inclusivas. El contexto transdisciplinar de las aplicaciones favorece que muchos términos ofrezcan una dependencia polijerárquica, y que se hayan subordinado al genérico más habitual. Mientras que la caracterización de los títulos

dentro del entorno digital se efectúa mediante su consulta individualizada. Se seleccionan los títulos cuya materia, método o herramientas empleados corresponden al entorno digital y se disponen en categorías prácticas para ponderar su impacto.

Con los listados de términos se elaboran grafos mediante la técnica de análisis de redes sociales (SNA) (Wasserman y Faust, 1997). Los enlaces marcan la relación existente entre dos nodos cuando se utilizan simultáneamente en el mismo artículo. El grosor de los enlaces está escalado en función del número de coocurrencias que, en la base de datos, ofrece un máximo de 8; Como sucede con los pares redundantes *bibliometría - cientometría*, *interdisciplinariedad - Ciencia de la Información y epistemología - Ciencia de la Información*. La red se filtra con el algoritmo de Fruchterman-Reingold (Gajdoš y otros 2016) para representar solo aque-

llos nodos con un peso de arista mayor o igual a 2, por lo que se eliminan los pares que aparecen solo una vez. Para calcular el modularidad de la red se emplea el Método Louvain de detección de comunidades que facilita colorear los grupos de términos entre los que existe mayor relación (Papadopoulos y otros, 2012). Finalmente, se utiliza el análisis de datos de redes sociales egocéntricas para producir los grafos de un grupo de términos.

5. ANÁLISIS DE LOS RESULTADOS GENERALES DE LA RECUPERACIÓN EFECTUADA

Las tres revistas brasileñas indizadas en JCR son las que publican más artículos (tabla I). Se destaca *Informação&Sociedade: Estudos* que duplica a la que ocupa el segundo lugar, *Transinformação*. Estas dos revistas, junto con *Perspectivas em Ciência da Informação*, aportan el 69% de la muestra, 614

Tabla I. Títulos de las revistas.

Título de la revista	ISBN	Nº artículos
Informação&Sociedade. Estudos	0104-0146	342
Transinformação	0103-3786	169
Perspectivas em Ciência da Informação	1413-9936	103
Scientometrics	0138-9130	101
Knowledge organization	0943-7444	49
Investigación bibliotecológica	0187-358X	29
Information processing and management	0306-4573	23
J of the association for Information science and technology	2330-1635	20
Journal of knowledge management	1367-3270	17
International journal of information management	0268-4012	13
Journal of informetrics	1751-1577	13
Knowledge management research & practice	1477-8238	11
Qualitative health research	1049-7323	10
Telematics and Informatics	0736-5853	10
J American society for information science and technology	1532-2882	9
International journal of geographical information science	1365-8816	8
Information development	0266-6669	7
Journal of global information management	1062-7375	7
Research evaluation	0958-2029	7
Revista Española de Documentación Científica	0210-0614	7
Social science information	0539-0184	6
Government information quarterly	0740-624X	5
Information research	1368-1613	5
Information technology for development	0268-1102	5
El profesional de la información	1386-6710	5
		TOTAL 981

Fuente WoS.

artículos. La cuarta, *Scientometrics*, supera también las 100 apariciones y duplica en el listado a la siguiente.

En otras 35 revistas se publican cuatro artículos o menos: dos revistas con 4 artículos; otras dos con 3; 10 con 2 y 23 con uno solo. Estas revistas no aparecen en el listado por razones de espacio, si bien muestran una interdisciplinariedad muy dispersa, al corresponderse muchas de ellas a dominios de colaboración menos frecuente para investigadores de LIS. A la hora de publicar es notoria la preferencia por revistas latinas, pues si a las 3 revistas brasileñas se añaden los 29 artículos aparecidos en *Investigación Bibliotecológica* y los 12 entre la *Revista Española de Documentación Científica* y *El Profesional de la Información* se alcanzan 715 originales, casi el 70%. Todas ellas facilitan el acceso abierto a sus contenidos, aunque la última en modo diferido.

5.1. Internacionalización, coautoría e interdisciplinariedad

La internacionalización se deriva tanto de las colaboraciones en la autoría de los artículos seleccionados como del idioma en que se publicaron (tabla II).

Tabla II. Idioma de publicación de los artículos

Idioma	Nº	Nº de citas
Portugués	552	565
Inglés	429	2338
Español	55	46
Francés	2	0
Húngaro	1	0
Total	1.039	2.949

Fuente WoS.

Se observa que los artículos escritos en inglés obtienen mayor difusión internacional (De-Souza, 2018) y, en consecuencia, son más citados. Estos datos confirman la observación de Garfield para mejorar la comunicación científica (1983).

Tabla III. Tipo de colaboración.

Colaboración	Nº	%	Nº de citas
Nacional	398	38,30	1.900
Internacional	293	28,20	485
Sin colaboración	348	33,49	564
Total	1.039		2.949

Fuente WoS.

La complejidad creciente de las investigaciones se afronta cada vez más en equipo, pues el 66,5% de la muestra se ha producido en colaboración (tabla III). Cuando, además, la colaboración es nacional, el 38,30%, recibe un acusado mayor porcentaje de citas que alcanza al 64,42%. Su desarrollo requiere la contribución de personas con competencias interdisciplinarias, en especial de las áreas que se señalan en la tabla V.

Tabla IV. Índice de coautoría.

Año	Autores	Artículos	Índice de coautoría
2010	266	107	2,49
2011	274	106	2,58
2012	304	128	2,38
2013	225	79	2,85
2014	257	93	2,76
2015	257	97	2,65
2016	392	130	3,02
2017	369	133	2,77
2018	324	121	2,68
2019	161	45	3,58
Total	2.572	1.039	2,77

Fuente WoS.

La participación media de 2,77 colaboradores por artículo es discreta ante las variadas tareas que requiere la elaboración de los artículos de carácter aplicado, la mayoría en el listado obtenido (tabla IV). Esa participación se sitúa por debajo de la mediana de autores por documento que en España es de 3 autores en la categoría *Information Science & Library Science* para el periodo 2007-2016 (IUNE, 2019). Sin olvidar que en un área de colaboración tan alta como *Computer Science* la mediana sube a 4 autores.

La interdisciplinariedad de LIS se basa en la inevitable intervención de las tecnologías informáticas y las telecomunicaciones en el tratamiento de la información. Esta circunstancia se ha fortalecido en el ambiente digital al aumentar la participación de las TIC, lo que supone abrir las actuaciones tradicionales de la información en cualquier centro o servicio de información hacia espacios mixtos (Bauer y otros, 2016), (Dias, 2008). Lo que refleja bien la tabla V al destacar las áreas de Informática y de Gestión como las de mayor participación. En especial aquella, que supone casi el 20% de las colaboraciones efectuadas por los investigadores de Información y Documentación. Continuando así la valoración que hicieron Humphreys y otros

Tabla V. Interdisciplinariedad de los artículos

Categorías WoS	Frecuencia	%
Library and Information Science	1.039	100
Computer Sci. interdisciplinary Applications	118	11,35
Computer Sci. Information Systems	83	7,98
Management	43	4,13
Social Sci. Interdisciplinary	17	1,63
Communication	10	0,96
Social Sci. Biomedical	10	0,96
Geography	8	0,76
Geography Physical	8	0,76
Telecommunications	3	0,28
Health Care Sci. Services	2	0,19
Medical Informatics	2	0,19
Education Educational Research	1	0,09
History	1	0,09
History of Social Sci.	1	0,09
Humanities Multidisciplinary	1	0,09
Multidisciplinary Sci.	1	0,09
Total	1348	

Fuente: WoS.

(1998) para la National Library of Medicine, en particular entre Informática, Información - Documentación y Lingüística. Y que en este caso se acrecienta, sin duda, por la aplicación de los más variados sistemas de información y gestión de TIC en empresas e instituciones.

El 4,13% que alcanza Administración y Gestión contrasta con su notable presencia en la descripción del contenido en los títulos y palabras-clave de los artículos, que luego se atenderá. Y nos hace pensar en una participación real de investigadores de ese dominio bastante inferior a la de aplicación de los asuntos de gestión a la información y sus servicios. Otro resultado de relación escasa es el de Comunicación, dominio del que, a priori, cabría suponer una mayor colaboración (Garvey, 1979) que el 0,96% reflejado en la muestra, pues la Teoría de la Comunicación viene participando decisivamente desde que se formó el concepto de *Information Science* en los años 60 (Georgia Institute of Technology, 1962). Esta baja presencia reitera lo observado por Sánchez-Vigil y otros (2015) respecto al número de artículos de Comunicación en las revistas españolas de Información y Documentación, tan solo un 8,73% del total. Aunque estas cifras podrían aumentar, pues hasta 16 artículos atienden a los medios de comunicación de las re-

des sociales y a los de comunicación social si se suman aquellos con perspectiva empresarial y de investigación social.

De los documentos recuperados, 55 no ofrecen palabras clave. Aunque publicaciones como *Information Research*, *JASIST* o *Restaurator* no las solicitan y plataformas como De Gruiter no las ofrecen en su acceso libre limitado en ocasiones a título y *abstract*. La media de palabras clave por documento es de 4,44, en gran contraste entre las 17 del artículo más generoso y los que tienen tan solo una o ninguna. De las 4.373 palabras clave asignadas por los autores 2.784 solo aparecen una vez y 265 dos veces.

5.2. Métodos y técnicas de investigación empleados

La metodología y las técnicas de investigación utilizadas se declaran en los títulos o palabras-clave de 688 artículos, el 69,91% de los 984 que ofrecían palabras clave (tabla VI).

Por su relieve se destacan las 197 presencias relativas a los estudios métricos de la información. Un 26,45% de los artículos emplearon sus métodos de análisis, indicadores, técnicas o herramientas dentro de cualquiera de los aspectos relacionados con la descripción de la actividad y comunicación científica (Bauer y otros, 2016). El peso de estas contribuciones se entiende bien al considerar que solo en *Scientometrics* se publicaron 101 de los artículos. El uso de modelos estadísticos es la base de los métodos métricos de la información, potenciados además por las tareas de computación a la hora de procesar los datos (Gorbea, 1998). Pero su empleo va más allá al atender a los datos acumulados en las organizaciones, sobre todo en empresas e industrias. Como también a los derivados del procesamiento lingüístico y textual o de la recuperación de la información, en especial los de minería de textos para buscar información en grandes colecciones de documentos no estructurados (Feldman y Sanger, 2007). De modo que los métodos, modelos y algoritmos de aplicación estadística, además de estar muy unidos a los estudios métricos, se aplican en 99 ocasiones a todos los dominios de referencia de los artículos a la hora de recopilar, comparar y evaluar. Hay que considerar que el estudio métrico de la información sigue usándose con los documentos tradicionales, pero se ha abierto a los objetos de contenido electrónicos disponibles en Internet y en la web (Leydesdorff y Wouters, 1999). Sin olvidar que el entorno digital está cambiando en profundidad las formas de comunicación científica.

Tabla VI. Métodos y técnicas cuantitativas y tecnológicas de investigación.

Estudios métricos de información 197 frecuencias	
Disciplinas	Bibliometría 78, cientometría 27, estudios métricos 7, altmetrics 6, informetría 4, estudios de usuarios 4, webometría 2, mapas de ciencia 2, cibermetría, métricas de relevancia
Métodos	Análisis de citas 34, análisis de cocitas 7, indicadores bibliométricos 2, análisis de coenlaces, factor de impacto, factor de impacto de las revistas, citas entre patentes
Leyes métricas	Zipf 4, Benford 2, Bradford 2, punto de transición de Goffman 2, índice h 2, coseno de Salton, Lotka
Herramientas	De análisis bibliométrico: Proknow c 3; redes de citaciones: main path analysis
Análisis estadístico 99 frecuencias	
Multivariante	Análisis de clúster 14, modelos de ecuaciones estructurales 7, análisis multivariante 5, análisis factorial 4, índice de Nihans
Descriptivo	Análisis de datos 9, análisis cuantitativo 8, minería de datos 3, análisis de big data, indicadores analíticos del almacén de datos
Otros	21 casos de aparición singular: algoritmos bivariantes, genéticos, de correlación y de proximidad
Redes y grafos	Redes complejas 10, análisis de redes sociales 6, fusión basada en gráficos 2, gráficos y redes 2, intermediación, representación de trayectoria elíptica, análisis de datos de la trayectoria, similitud de trayectoria, grafos genealógicos
Métodos de evaluación 79 frecuencias	
Indicadores, evaluación, medidas	Aplicados a: rendimiento, sistemas de recuperación, competencias, servicios electrónicos, impacto, productividad, análisis semiótico, modelos de servicio, modelos de negocio, circulación, descripción, organización; imagen o impacto social de bibliotecas, archivos, museos, servicios de información; sitios web o portales
Métodos informáticos 56 frecuencias	
Inteligencia artificial	Aprendizaje automático 12, programación genética 2, reconocimiento de patrones 2, análisis del árbol de fallas, flujos lingüísticos formales e informales, metodología de sistemas inteligentes, razonamiento espaciotemporal, red temporal
Procesamiento del lenguaje natural	PLN 7, Análisis de sentimiento 7, minería de textos 6, medidas de similitud 2, derivaciones, minería de contenidos, n-gram, procesamiento multi-documental, reconocimiento de patrones sintácticos
HCI	Interacción persona-máquina 3 (disciplina)
Software	Simulación por ordenador (modelos matemáticos de ordenador) 4, computación ubicua

Con otras 79 presencias aparecen diferentes métodos de análisis y evaluación que aplican indicadores, evaluaciones o mediciones a los sistemas, servicios y procesos informativos. La mayoría son inseparables de los procedimientos informáticos en los que se apoyan, como sucede en la mayor parte de los apartados que estamos mostrando. Lo que no impide reconocer 56 técnicas como características de la investigación en Informática, aunque muchas de las incluidas en Inteligencia artificial sean de base estadística y las de Procesamiento del lenguaje natural muestren interdisciplinariedad inmediata con organización y representación del conocimiento (tabla VII).

Los métodos de carácter inductivo, narrativo y sintético, empleados para identificar, evaluar críticamente e integrar los resultados, en apreciación de Cooper (2003), aparecen 54 veces. La presencia de mapeo entre ellos debe entenderse en el sentido de complemento a las revisiones sistemáticas, no como interoperabilidad (Sayão y Marcondes, 2008). Mientras que en 49 ocasiones se utiliza la investigación social para analizar los hechos de

la información, obtener datos cualitativos, conocer los gustos de los usuarios o los índices de satisfacción. Este grupo se completa con las 10 presencias de análisis y cálculos de carácter geográfico, geo-morfométrico o topográfico y con el mismo número de propuestas metodológicas aplicadas a educación y formación.

Por su parte, epistemología, paradigmas y filosofía de LIS resalta en 41 ocasiones la función de la teoría en la investigación. Son cercanas las 19 apariciones de técnicas y teorías relacionadas con el análisis y la interpretación histórica de los hechos documentales o de la evolución de los temas de investigación.

Destaca el empleo de métodos y técnicas propias de la investigación en administración y economía de la empresa, junto a la gestión de sistemas de información, en especial los contables, con presencia en 65 casos. Finalmente, en 60 ocasiones se aplican métodos y técnicas semánticas e interpretativas a la representación y organización del conocimiento. Son, sobre todo, técnicas de recu-

Tabla VII. Métodos y técnicas cualitativas de investigación.

Métodos inductivos, narrativos y sintéticos	
<i>Métodos inductivos</i> 54 frecuencias	
Estudio de caso 17, mapeo sistemático 15, evaluación cualitativa 13, modelos conceptuales 8, mediación cultural 6, revisión sistemática de la literatura 4, método de ordenación	
<i>De Ciencias Sociales</i> 49 frecuencias	
Investigación mediante encuestas 13, análisis de contenido 10, grupos focales 6, paradigmas y modelos sociales 5, protocolo verbal 4, entrevista cualitativa 3, interaccionismo simbólico 3, triangulación metodológica 2, teoría del actor-red, enfoque narrativo, netnografía	
<i>Geográficos y Topográficos</i> 11 frecuencias	
Cálculo Geomorfométrico 3, análisis de usos del suelo 2, mapas de riesgos 2, doble ambigüedad cartográfica, métodos de reconstrucción empírica, estudio cartográfico, análisis espacial	
<i>Educación e instrucción</i> 10 frecuencias	
Análisis de colaboración 2, modelos de objetos de aprendizaje 2, diagnóstico del alumno, integración entre teoría y práctica, método integrador, ludificación, método de proyectos, tutoría	
<i>Administración y gestión</i> 65 frecuencias	
Análisis empresarial 13, indicadores de gestión 11, modelos y procesos de toma de decisiones 9, sistemas difusos de apoyo a la decisión 5, mapas de competencias 3, análisis de riesgos 3, cadena de suministro 3, modelo de madurez de capacidades 3, comportamiento planeado 2, adhocracia, benchmarking, análisis de eficiencia, FMEA, modelo de gravedad, gestión de ideas, modelo integrado de gestión del conocimiento, modelo de gestión, expansión multimodal de consultas, cambio organizacional, regresiones aparentemente no relacionadas, isoperms de Sebera, modelo de aceptación de tecnología	
<i>Organización y representación del conocimiento</i> 60 frecuencias	
Métodos y técnicas semánticas: análisis del discurso 12, análisis de dominios 7, selección de frases nominales 6, análisis de textos 4, análisis facetado 2, aparato formal de la enunciación, semántica formal, compatibilidad semántica Recuperación de información: expansión de la búsqueda 2, procesamiento de la búsqueda 2, métodos asociativos, método de búsqueda semiótica, modelo de categorización, modelo vectorial. Desarrollo de vocabularios: Desarrollo de taxonomías u ontologías 9, interoperabilidad por mapeo 2, estándares de metadatos 2, SKOS 2, matriz del lenguaje, Protégé, OWL	
<i>Fundamentación teórica</i>	
<i>Epistemología de LIS</i> 41 frecuencias	Epistemología 26, filosofía de LIS 8, paradigmas 5, postura del pensamiento jurídico, método dialéctico
<i>Historiografía</i> 19 frecuencias	Métodos históricos y sociohistóricos 19

peración y de desarrollo de vocabularios, estrechamente relacionadas con la Informática, en especial con la Lingüística computacional aplicada al PLN, y con la estadística.

Para las fronteras de los métodos del conocimiento seguidos en un área tan pequeña como LIS no es ningún obstáculo que esta visión presente tanta permeabilidad como ya reflejara Cronin (2012).

6. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS DESDE LA TRANSFORMACIÓN DIGITAL.

6.1. Subcategorías de los títulos del corpus.

Se comprueba que 510 títulos se corresponden con la transformación digital (tabla VIII), casi el 50% de la muestra, una vez contrastados los no pertinentes. Se distribuyen en categorías inducidas durante las tareas de identificación y selección, sin poder evitar el influjo de las redes de términos que se estaban elaborando con las palabras clave.

Una primera consideración proviene del uso no siempre preciso de datos, información o conocimiento como palabras clave. Al revisarse los contenidos se verifica que en ocasiones se manejan como equivalentes (Liew, 2007). En especial con

Tabla VIII. Categorías de los títulos

Categorías	Nº Títulos	%
Centros y servicios de información (InCS)	122	23,92
Internet-Web (InWeb)	94	18,43
Gestión del conocimiento (GC)	83	16,27
Sociedad digital (DG)	78	15,29
Tecnologías de la información y comunicación (TIC)	71	13,92
Organización del conocimiento (OC)	62	12,15
Total	510	99,98

su uso indistinto como elementos componentes de organización del conocimiento y representación del conocimiento, que dividen su aparición con organización de la información y representación de la información, con ligera ventaja para las primeras de cada pareja.

En primer lugar, los títulos incluidos en la categoría Centros y servicios de información son un reflejo directo de la digitalización de los objetos de contenido y de las consecuencias que ha tenido en las formas de hacer su tratamiento, acceso y conservación. 24 títulos recorren todo tipo de archivos, bibliotecas y museos digitales, en los que la comunicación de los más variados objetos digitales depende de los ordenadores en red, en 11 ocasiones. En especial desde los problemas abordados por el movimiento abierto. Aunque ciencia abierta no sale ninguna vez en el listado de términos, si lo hacen y mucho los elementos que se integran en sus principios y recursos de conocimiento. Otros 12 títulos se ocupan de los artículos científicos digitales, las revistas electrónicas y hasta de magazines virtuales. Los medios sociales se especifican en Twitter, Orkut, Flickr, Instagram, Youtube, sobre todo en Facebook, para difundir informaciones, pero también con la presencia de blogs y hasta de *likes* para valorar lo emitido. Son muchos los centros que utilizan estos medios sociales para promocionar o comercializar sus actividades y recursos multimedia, en una iniciativa emprendida por los profesionales de la información para realizar su primer papel, el de comunicar. Sin olvidar los 20 títulos alrededor de los repositorios institucionales digitales que proporcionan la mejor plataforma para archivar contenidos propios, datos de proyectos u objetos de aprendizaje, con beneficio inmediato para los usuarios cuyo comportamiento en la búsqueda y uso de la información también se considera. Igual que interesan los problemas de la preservación digital o de las publicaciones fraudulentas. Es notable el avance que ha habido desde los trabajos iniciales, casi veinte años atrás, cuando se centraban en los procesos de digitalización de los acervos custodiados en soportes físicos.

Por su parte, los títulos incluidos en TIC y en Internet-Web son inmediatos a los aspectos transdisciplinarios de investigación, aplicación y funcionalidad intermediados por lo digital, pues casi siempre sobrevuelan las investigaciones conjuntas de LIS con otros dominios. En especial los de Internet-Web cuyas apariciones son complicadas de deslindar, pues interfieren con cualquier ámbito de aplicación. Si agrupamos los 189 títulos válidos de ambas se percibe con claridad la importancia de los factores tecnológico-digitales en LIS. La contribución de los sistemas y procesos digitales es imprescindible para los centros de información, pero más aún para ges-

tionar el conocimiento en empresas y organizaciones. La categoría de Internet-Web incluye 94 documentos que abarcan todo el espectro de internet desde red informática hasta internet de las cosas, junto a la información que permite compartir y que alberga la web. Con presencia acusada de redes sociales, de cada una de sus aplicaciones y herramientas colaborativas, y de medios electrónicos. También en este caso se percibe el empleo confuso de redes sociales en el sentido de sociedad en red o de comunidad virtual frente a apariciones que la relacionan con estudios de producción científica. En los títulos referentes a la web se da mucha atención a la web 2.0 y sus recursos; a la web semántica, sus motores, tecnologías y especificaciones; a la arquitectura de la información, protección de datos, *big data*, datos enlazados y datos abiertos; búsquedas y navegación en la web, web invisible y sitios web de museos.

Las categorías gestión del conocimiento y organización del conocimiento reflejan la incidencia de la intervención tecnológica y los cambios habidos en la gestión y en el tratamiento del conocimiento y de los objetos de contenido. Sin duda, las empresas e instituciones emplean la capacidad gestora que les ofrecen las TIC para potenciar el valor económico del conocimiento en su proceso productivo y organizativo (Risso, (2012). Por lo que dependen cada vez más del manejo tecnológico para favorecer su agilidad, su adaptación a los procesos en red y la gestión de su conocimiento (Del-Canto y otros, 2014). El grupo de 83 títulos de la categoría gestión del conocimiento tiene que ver, por una parte, con investigaciones focalizadas en el empleo de técnicas de la información para *e-commerce*, portales de negocios, portales de empresa, conocimiento tácito, toma de decisiones y *márketing*, como identificaran Kroll y Pinto (2018). Por otra, los modelos de gestión del conocimiento se aplican de manera creciente al buen desempeño de los sistemas y centros de información. En especial a los procedimientos destinados a satisfacer la atención que, cada vez más individualizada, exigen los usuarios. Entre otras razones, porque la implicación del *márketing* mejora la prestación de servicios para atender las solicitudes de los usuarios, así como para conocer la satisfacción que alcanzan (Do-Amaral, 2011). En este sentido, son apreciables los 12 títulos referentes a gestión en las bibliotecas y a calidad de los servicios de información, o los 8 de *márketing* de información digital. Asimismo, se destaca la presencia del conocimiento compartido y la transferencia de conocimiento. En relación con la estrategia digital y la gestión de negocio no se muestran, como sería de esperar, referencias a SEO, cuando Brasil es el país de América latina con mayor empleo de estos

especialistas (Romero, 2019); ni a construcción de enlaces, análisis web o proyectos web. Lo que indica que su atención se deja en manos de Administración y gestión o de Comunicación. En los perfiles combinados de estos ámbitos no se identifica ninguna aparición de gestión de comunidades, gestión de contenidos digitales o edición de contenido en el sentido que les da FTI-AMETIC (2012). Sí hay bastantes de medios de comunicación de las redes sociales, hasta 12. Muchas de ellas en aplicaciones métricas al estudio de sus contenidos, junto a su empleo en el marketing digital de la información o de las bibliotecas antes referido (Dueñas y otros, 2005). De nuevo se manifiesta la escasa investigación conjunta entre sectores que deben de colaborar ante la ausencia de estudios transdisciplinares de TIC, LIS y Comunicación (Moreiro, 2018).

No todos los contenidos y técnicas atendidas por los títulos de gestión y de organización del conocimiento pueden entenderse siempre dentro del entorno digital. Por lo que no se contabilizaron 73 de los artículos cuyo tema les es propio pero que presentan continuidad temática, técnica o metodológica con los usos precedentes, como sucede también al grupo de Organización del conocimiento (OC) y, en realidad, a todas las categorías.

El rango de la Sociedad digital hace referencia inmediata al impacto de las TIC en la cultura, el desarrollo de las ciudades, el e-gobierno, el empleo, la ciencia, la educación y la organización social. Se muestra en tres marcos de representación. El primero en torno a las consecuencias inmediatas que ha tenido la transformación digital a través de denominaciones variadas: postmodernidad o era digital, sociedad del conocimiento, sociedad red, ciberespacio o cultura digital. El segundo en torno al empleo de internet para proporcionar información y servicios públicos o para facilitar la comunicación y la colaboración entre el gobierno y los ciudadanos sobre las que apoyar la planificación y la toma de decisiones, con atención al uso de internet en la administración, las ciudades o los colectivos inteligentes mediante la democracia y la participación electrónicas, los movimientos sociales en línea o el e-gobierno. Por fin, los artículos que se han elaborado en torno a los aspectos sociológicos de las redes sociales y los comportamientos de las nuevas generaciones. Aparecen bien representados los aspectos tocantes a la inclusión digital, también a la brecha digital. De otros campos, tienen especial repercusión por su número, los títulos interdisciplinares sobre e-Salud y tecnología asistencial, junto a los de banca móvil, webs bancarias y microfinanzas, o de los móviles, hasta alcanzar a las plataformas de música, los detectores de humo de tabaco en el metro o el petróleo y el gas. Con presencia de los

problemas éticos y jurídicos causados por el uso indebido de lo digital, en especial el plagio y las copias fraudulentas. Llama la atención la representación relacionada con educación desde las habilidades y derechos básicos en competencias de información y formación permanente. Y desde luego, los entornos virtuales de aprendizaje, los objetos de aprendizaje y la educación a distancia. Sobre formación y enseñanza en LIS hay otro grupo significativo, pero está poco relacionado con los ambientes digitales, fuera de su modalidad de impartición a distancia.

La clase TIC agrupa 71 títulos, entre los que aparecen en más de diez ocasiones el propio término TIC (De-Araújo y Oliveira, 2015), junto a recuperación de información y organización e industria de TI. Con más de cinco se sitúan visualización de información; inteligencia artificial; motores y robots de búsqueda; ingeniería del conocimiento; uso, transferencia y convergencia de la tecnología; y desarrollo e industria de programas. Junto a una gran variedad de menor presencia que va desde profesional de TI o nanotecnología, pasando por redes cooperativas, ética de la información digital, lectura digital o infografía, hasta TI inalámbrica o ciber-infraestructura de investigación, que permiten representar con rigor el papel de las aplicaciones tecnológicas en LIS.

A los resultados del análisis de Organización del conocimiento se les da mayor atención al atender luego a los clústeres de términos. Su contexto es inseparable de la vinculación a la Web semántica, con la vigencia de los nuevos estándares o, incluso, de los vocabularios y esquemas elaborados por empresas e industrias para organizar y gestionar su información. Adquieren ahí un vigoroso espacio las ontologías, con 24 apariciones, junto a 9 de las taxonomías. Sin olvidar los 37 artículos correspondientes a Sistemas de organización del conocimiento (SOC) y vocabularios semánticos.

6.2. Clúster de las palabras-clave de autor

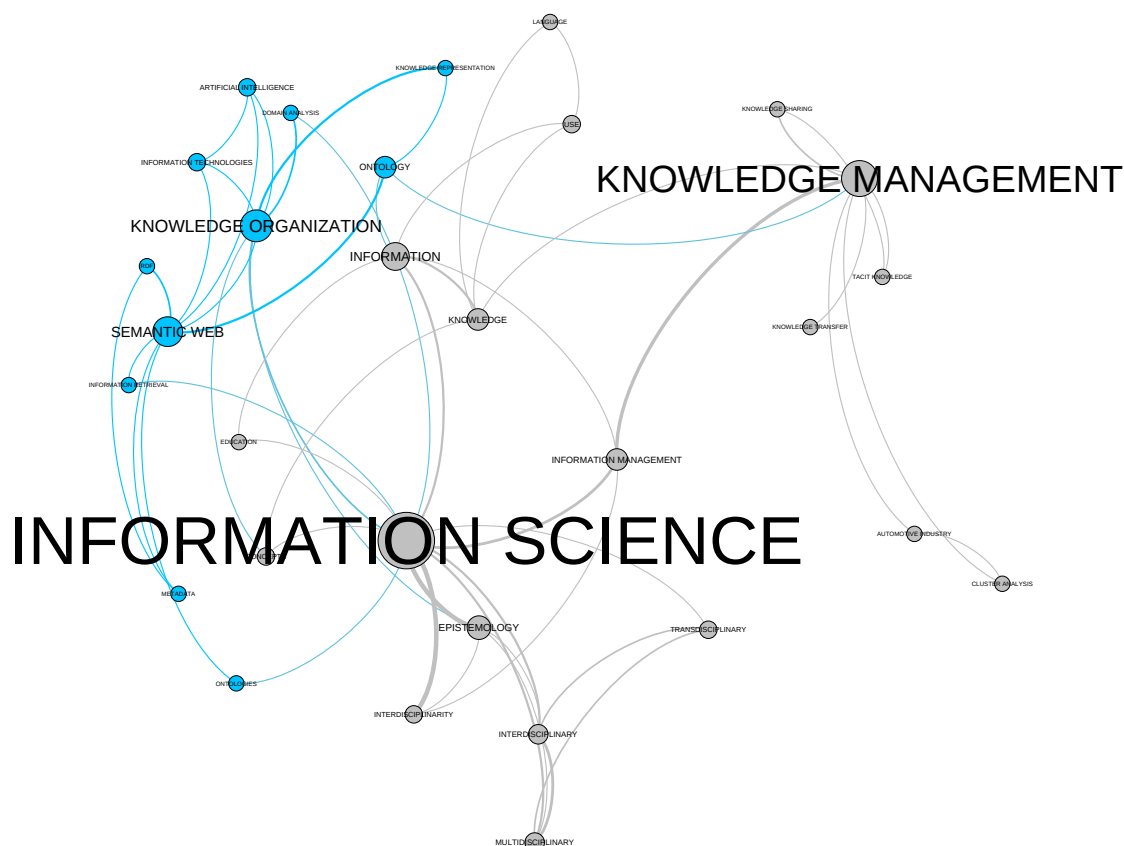
El clúster de las palabras clave contenidas en los 984 documentos representa todo el léxico utilizado por los autores para indizar sus artículos. Por lo que incluyen muchos conceptos comentados antes al tratar la metodología utilizada en 688 de los artículos. El tamaño de los nodos depende del grado de enlaces de un término con los del resto del grafo. Mientras que el grosor de los enlaces está escalado en función del número de apariciones simultáneas de las palabras clave en el mismo artículo. El grosor de las aristas en la base de datos va de 1 a 8, el máximo número de veces que un par de palabras clave coinciden. Como, por ejemplo, en los pares antes señalados de *bibliometría - cientometría*; *interdisci-*

Se comprueba que cinco de los nodos con mayor frecuencia se vinculan directamente con las innovaciones tecnológicas y la transformación digital. Si bien, esa vinculación va más allá de la correspondencia directa entre nodos, al darse una proximidad contextual al resto de nodos e islas. Como muestra que tareas profesionales y competencias enlacen con Internet que, a su vez, lo hace con gestión del conocimiento. No se incluyen en la tabla, pero consiguen cuatro arcos: redes sociales, multidisciplinariedad, interdisciplinariedad y transdisciplinariedad. Las dos islas que se observan en el clúster están compuestas por los nodos: tendencias TIC para el futuro digital, presupuestos de TIC, estructura organizativa de TIC, problemas de gestión y principales aplicaciones y tecnologías, la primera; mientras que la segunda se integra por gobernanza de Internet, regulación de Internet y tipo cero. La relación de ambas islas es inmediata con las TIC. Igual que sucede con la cuasi-isla de relaciones en torno a interoperabilidad-metadatos.

El grupo de términos referentes a gestión del conocimiento está integrado por 117 pares de términos totales que suponen el 8,2%, lo que habla de

su relevancia (figura 2). La interdisciplinariedad en la práctica metodológica de Gestión del conocimiento se refleja no solo por los dominios de aplicación, sino también por su empleo simultáneo con TIC y con SOC. Se muestra con claridad desde los enlaces de ontología y de taxonomía. Pues ambas organizan la información e integran los nuevos datos, a partir de los cuales se toman decisiones y se realizan operaciones comerciales. De modo que una parte representa el conocimiento sobre los objetos de contenido y otra de las empresas que los sirve Su situación es intermedia entre TIC, Web semántica y gestión del conocimiento. Este clúster enlaza a través de los nodos Información y gestión de la información con organización del conocimiento. Los nodos de ontologías, con 24 apariciones en la red total, y los 9 de taxonomías, aparecen conectados con gestión del conocimiento. También, pero con menor presencia, los 8 de tesauros y los 3 de folksonomías. Son 37 los nodos correspondientes a SOC, PLN, indexación automática, metadatos, interoperabilidad y vocabularios, entre otros. Si bien con presencia en el clúster de 10 enlaces más que muestran continuidad en nombres, métodos y procesos que son ajenos al ambiente digital.

Figura 2. Clúster de gestión del conocimiento y organización del conocimiento



Como última consideración, destaca la presencia del nodo TIC cuya importancia se muestra en los 201 enlaces que protagoniza y en los 173 documentos que lo contienen. Si añadimos los aportados por los nodos Internet y Web que afectan a 186 enlaces, siendo el asunto principal de 71 títulos de documentos, se percibe la pujanza de los asuntos relacionados con las tecnologías que fundamentan la transformación digital.

En las palabras clave se reflejan muchos nombres de estados y ciudades brasileñas, de universidades y de otras instituciones; también de personas académicas.

7. CONCLUSIONES

Respecto al objetivo general, se identifica una intensa participación de los investigadores brasileños en el escenario de la transformación digital en LIS, donde han cambiado los modelos y el manejo de la información. Las herramientas y tecnologías digitales apoyan la investigación colaborativa, la extracción, análisis y asociación de los datos y su posterior visualización, comunicación y difusión. LIS, en cuanto ciencia social, además del ubicuo procesamiento informático, refleja su estrecha colaboración con los procesos estadístico-métricos, de evaluación, de administración y gestión, así como la aplicación de sus procedimientos a un sin fin de especialidades y de sectores de actuación.

Desde la finalidad del primer objetivo específico se comprueba que el carácter interdisciplinar de LIS lleva al empleo de métodos y aplicaciones variadas. Entre los que destacan los relacionados con los estudios métricos de información y los análisis estadísticos que intervienen en gran parte de las aplicaciones, más allá de las propias métricas. Igual que sucede con la intervención decisiva de los métodos informáticos. Sin olvidar las investigaciones sociales, en especial las derivadas de la Gestión-administración para diseñar, evaluar y mantener los centros y servicios. Así como las de carácter semántico y lingüístico aplicado a organización y representación del conocimiento, con el empleo tan exclusivo en LIS de las técnicas de recuperación de información, junto a otros métodos inductivos. No se confirma un número elevado de investigaciones descriptivas y exploratorias. Tal vez porque las revistas indizadas rechacen, entre otros, los manuscritos que no suponen modificación en el estado de los conocimientos.

El propósito del segundo objetivo se alcanza desde la claridad con que los artículos muestran la influencia de la transformación digital en los cambios metodológicos y en los asuntos estudiados. El enlace con la profesión es inmediato si se añade

que muchos de ellos son aplicaciones prácticas. Se comprueba también desde el índice de coautoría que es comparable con el de la interdisciplinariedad necesaria para desarrollar investigaciones polifacéticas. Se muestra con claridad la colaboración con investigadores de diferentes disciplinas. Las de mayor relación son Informática de sistemas, Inteligencia artificial y Gestión-administración. Así como desde el uso de métodos de carácter automatizado, ahora potenciados por la comunicación y la capacidad tecnológica del ambiente digital, que han transformado todo el análisis y la gestión de la información y que suponen la aparición incesante de nuevos conceptos y procesos. Los sectores de colaboración identificados coinciden con aquellos en que se desarrolla la actividad profesional más reciente, pues si se categoriza se generaliza.

En consistencia con el tercer objetivo específico, es muy representativa la producción realizada en colaboración con investigadores de otros países. Los autores brasileños de LIS muestran una notable colaboración internacional, tanto por la coautoría como por el idioma de publicación. Se observa que, además del portugués, el idioma más empleado es el inglés, a mucha distancia del español. Los otros idiomas manejados, el francés y el húngaro, son testimoniales. Los artículos escritos en inglés obtienen mayor difusión internacional y, en consecuencia, son más citados.

Aunque el cuarto objetivo fija una aplicación de continuidad y, pese a que en las aplicaciones tecnológicas los procesos nunca terminan, los resultados obtenidos se deberían de considerar dentro de una línea de coherencia académica que parte de la investigación, se comunica en la docencia y se aplica en la profesión. En este caso por compromiso metodológico, inmediatez temática, preparación competencial e innovación conceptual. Quien investiga en los artículos analizados es un profesional de la información, en la mayoría de los casos profesor. De forma que habría que comprobar en trabajos posteriores si estos sectores de investigación lo son también de docencia. Si cuando se enseña se aplican las mismas competencias y habilidades que cuando se investiga. También si lo que se investiga se enseña.

Como cabría suponer, en el periodo analizado destaca la atención tocante a las colecciones de documentos digitales o digitalizados desde otros soportes, su procesamiento, preservación, conservación y curaduría, con énfasis en los servicios de acceso, gestión y procesamiento digital, en el estudio de los bancos y bases de datos, los repositorios, el acceso abierto, la organización del conocimiento y la satisfacción de las necesidades de los usuarios. La acumulación de estudios correspondientes a estos

elementos expresa la penetración de la transformación digital en las tareas de investigación y en las actividades profesionales, si bien se observa alguna desaparición nominal como la de las bibliotecas híbridas. A la vez, se atiende al carácter relacional de la gestión y reutilización del conocimiento, sobre todo en las esferas comercial y jurídica. De hecho, el análisis de datos es uno de los resaltes, ampliado además con los *big data*, como consecuencia de la necesidad de emplear las TIC en las empresas, los centros de enseñanza superior y los organismos públicos para administrar, analizar y procesar datos. Siempre en relación con las técnicas estadísticas de minería, extracción de valor y análisis de corpus o con la gestión de datos de investigación. Vinculado a la experiencia en el análisis y la gestión de datos asoma, aún con cierta timidez, el uso de la inteligencia artificial, la ingeniería del conocimiento y el aprendizaje automático.

9. REFERENCIAS

- Arboit, A. E.; Bufrem, L. S.; Moreiro-González, J. A. (2011). A produção brasileira em Ciência da Informação no exterior como reflexo de institucionalização científica. *Perspectivas em Ciência da Informação*, 16 (3), 75-92. DOI: <https://doi.org/10.1590/S1413-99362011000300005>
- Åström, F. (2007). Changes in the LIS research front: Time-sliced cocitation analyses of LIS journal articles, 1990–2004. *Journal of the American Society for Information Science and Technology*, 58 (7), 947-957. <https://doi.org/10.1002/asi.20567>
- Bauer, J.; Leydesdorff, L.; Bornmann, L. (2016). Highly cited papers in Library and Information Science (LIS): Authors, institutions, and network structures. *Journal of the Association for Information Science and Technology*, 67 (12), 3095-3100. DOI: <https://doi.org/10.1002/asi.23568>
- Bufrem, L.S.; Silva, H. N.; Fabian, C.; Sorribas, T. (2007). Produção científica em Ciência da Informação: análise temática. *Perspectivas em Ciência da Informação*, 12, (1), 38-49. <https://doi.org/10.1590/S1413-99362007000100004>
- Cooper, H. (2003). Editorial. *Psychological Bulletin*, 129, (1), 3-9. DOI: 10.1037/0033-2909.129.1.3. <https://doi.org/10.1037/0033-2909.129.1.3>
- Cronin, B. (2012). The waxing and waning of a field: reflections on information studies education. *Information Research: An International Electronic Journal*, 17, (3), paper 529. Disponible en: <http://InformationR.net/ir/17-3/paper529.html> [Fecha de consulta: 18/08/2019].
- De-Araújo, R. Ferreira; Oliveira, M. (2015). A tecnologia na pesquisa no ensino da Biblioteconomia e Ciência da Informação no Brasil: representações, conteúdos e práticas. *Pesquisa brasileira em Ciência da Informação e Biblioteconomia*, 10, (2), 64-72. Disponible en: <http://www.periodicos.ufpb.br/ojs/index.php/pbcib/article/view/24932> [Fecha de consulta: 17/06/2019].
- La misma razón subyace en la atención a la mudanza de la comunicación en empresas, administraciones y servicios de información a la hora de aprovechar la utilidad de las redes sociales para facilitar sus contenidos, servicios y tareas en beneficio de los usuarios. Otro sector de interés dinámico es el que marcan los estudios de la información en el gobierno y la democracia electrónicos, así como en las ciudades inteligentes o respecto a la enseñanza e inclusión digitales.
- ## 8. AGRADECIMENTOS
- Trabajo financiado por el Proyecto de investigación nº 2018-02595-8 de la FAPESP
- ## ACKNOWLEDGEMENTS
- This work has been funded by FAPESP Research Project No. 2018-02595-8
- De-Filippo, D. (2013). Spanish scientific output in Communication sciences in WOS. The Scientific Journals in SSCI (2007-12). *Comunicar*, 21, (41), 25-34. [Fecha de consulta: 20/02/2020]. <https://doi.org/10.3916/C41-2013-02>
- Del-Canto, E.; Sánchez, P.; Martins, V. F.; Soares, A. B. (2014). Agentes generadores de innovación en las empresas: gestión del conocimiento y TIC. *Revista de Auditoria, Governança e Contabilidade*, 2, (3), 72-88. Disponible en: <http://fucamp.edu.br/editora/index.php/rac/article/viewFile/465/334> [Fecha de consulta: 21/04/2020].
- De-Oliveira, E.T. (2018). *Estudos Métricos da Informação no Brasil: Indicadores de produção, colaboração, impacto e visibilidade*. Marília: Editora da UNESP.
- De-Souza, C.-D. (2018). *Impacto de las políticas brasileñas de ciencia y tecnología en la actividad investigadora de las universidades federales: un estudio cuantitativo del período 2003-2015*. Madrid: Universidad Carlos III de Madrid [Tesis doctoral]. Disponible en: https://e-archivo.uc3m.es/bitstream/handle/10016/27815/souza_tesis_2018.pdf?sequence=1&isAllowed=y [Fecha de consulta: 11/06/2019].
- Dias, F. S. (2008). Análise das relações interdisciplinares das pesquisas científicas em sistemas de informação. *Perspectivas em Ciência da Informação*, 13, (1), 199-215. Disponible en: <http://www.scielo.br/pdf/pci/v13n1/v13n1a13> [Fecha de consulta: 08/08/2019]. <https://doi.org/10.1590/S1413-99362008000100013>
- Do-Amaral, S. A. (2011). Marketing da informação: abordagem inovadora para entender o mercado e o negócio da informação. *Ciência da Informação*, 40, (1), 85-98. Disponible en: <http://revista.ibict.br/ciinf/article/view/1327> [Fecha de consulta: 11/07/2019]. <https://doi.org/10.1590/S0100-19652011000100007>
- Dueñas, J. C.; Burillo, V.; Ruiz, J. L. (2005). Professional ICT Profiles to Implement Services and Digital Content. Madrid: Colegio Oficial de Ingenieros de Telecomunicación. PAFET 4. Disponible en: <https://www.coit.>

- es/informes/pafet-4-perfiles-profesionales-tic-para-la-implantacion-de-servicios-y-contenidos-digital-0 [Fecha de consulta: 03/06/2019].
- Feldman, R.; Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. New York, Cambridge University Press. <https://doi.org/10.1017/CBO9780511546914>
- Ferrán-Ferrer, N.; Guallar, J.; Abadal, E. Server, A. (2017). Research methods and techniques in Spanish Library and Information science journals (2012-2014). *Information research*, 22, (1), paper 741. <http://InformationR.net/ir/22-1/paper741.html>
- Fleischmann, K. (2007). Digital libraries with embedded values: Combining insights from LIS and science and technology studies. *The Library Quarterly*, 77, (4), 409-427. <https://doi.org/10.1086/520997>
- FTI-AMETIC (2012). *Estudio de Perfiles Profesionales más demandados en el ámbito de los Contenidos Digitales en España 2012-2017. Profesionales TIC*. Madrid: Fundación de Tecnologías de la Información-AMETIC. Disponible en: <https://www.redalyc.org/articulo.oa?id=93528051014> [Fecha de consulta:02/02/2020].
- Gajdoš, P.; Jeřowicz, T.; Uher, V.; Dohnálek, P. (2016). A parallel Fruchterman-Reingold algorithm optimized for fast visualization of large graphs and swarms of data. *Swarm and Evolutionary Computation*, 26, (February), 56-63. <https://doi.org/10.1016/j.swevo.2015.07.006>
- Garfield, E. (1983). Mapping science in the third world. *Science and Public Policy*, 10, (3), 112-127. <https://doi.org/10.1093/spp/10.3.112>
- Garvey, W. (1979). *Communication: the essence of science: facilitating information exchange among librarians, scientists, engineers and students*. Oxford: Pergamon I. Library.
- Georgia Institute of Technology. (1962). *Proceedings of the conference on Training Science Information specialists, October 12-13, 1961*. Atlanta. Sponsored by the National Science Foundation Dorothy M. Crosland, General Chairma 139 pp.
- Gorbea Portal, S. (1998). Modelación matemática de la actividad bibliotecaria: una revisión. *Investigación Bibliotecológica*, 12, (24), 5-23. DOI: <http://dx.doi.org/10.22201/iibi.0187358xp.1998.24.3869>
- Humphreys, B.; Lindberg, D.; Schoolman, H.; Barnett, O. (1998). The unified medical language system: an informatics research collaboration. *Journal of the American Medical Informatics Association*, 5, (1), 1-11. <https://doi.org/10.1136/jamia.1998.0050001>
- Hurd, J. M. (2000). The transformation of scientific communication: A model for 2020. *Journal of the American Society for Information Science*, 51, (14), 1279-1283. [https://doi.org/10.1002/1097-4571\(2000\)9999:9999<::AID-ASI1044>3.0.CO;2-1](https://doi.org/10.1002/1097-4571(2000)9999:9999<::AID-ASI1044>3.0.CO;2-1)
- IUNE. 2019. Mediana de autores por documento por categoría WoS y año (SUE 2008-2017). Madrid; IUNE. Disponible en: http://www.iune.es/es_ES/informes [Fecha de consulta:05/07/2019].
- Järvelin, K.; Vakkari, P. (1993). The evolution of library and information science 1965-1985: A content analysis of journal articles. *Information Processing & Management*, 29, (1), 129-144. DOI: [https://doi.org/10.1016/0306-4573\(93\)90028-C](https://doi.org/10.1016/0306-4573(93)90028-C)
- Jacsó, P. (2009). Five-year impact factor data in the Journal Citation Reports. *Online Information Review*, 33, (3), 603-14. DOI: <https://doi.org/10.1108/14684520910969989>
- Kroll-do-Prado, J. M.; Pinto, A. L. (2018). Marketing in Brazilian Information Science: four decades of production. *Informação&Sociedade: estudos*, 28, (3), 53-68. DOI: <https://doi.org/10.22478/ufpb.1809-4783.2018v28n3.37935>
- Leydesdorff, L.; Wouters, P. (1999). Between texts and contexts: advances in theories of citation? *Scientometrics*, 44, (2), 173-192. Disponible en: <https://link.springer.com/content/pdf/10.1007/BF02457378.pdf> [Fecha de consulta:08/08/2019]. <https://doi.org/10.1007/BF02457378>
- Liew, A. (2007). Understanding data, information, knowledge and their inter-relationships. *Journal of knowledge Management Practice*, 8, (2), 1-16. Disponible en: <https://socygu.ga/1549290980.pdf> [Fecha de consulta:15/08/2019].
- Marteletto, R. (2009). A pesquisa em Ciência da Informação no Brasil: marcos institucionais, cenários e perspectivas. *Perspectivas em Ciência da Informação*, 14, (3), 19-40. <https://doi.org/10.1590/S1413-99362009000400003>
- Matt, C.; Hess, T; Benlian, A. (2015). Digital Transformation Strategies. *Business & Information Systems Engineering*, 57, 339-343. <https://doi.org/10.1007/s12599-015-0401-5>
- Moreiro-González, J. A. (2018). Reciprocidad laboral, formativa y difusora entre Información-Documentación y Comunicación social. *Anuario ThinkEPI*, 12, 90-95. DOI: <https://doi.org/10.3145/thinkepi.2018.08>
- Papadopoulos, S.; Kompatsiaris, Y.; Vakali, A.; Spyridonos, P. (2012). Community detection in social media. *Data Mining and Knowledge Discovery*, 24, (3), 515-554. <https://doi.org/10.1007/s10618-011-0224-z>
- Pettigrew, K.; McKechnie, L. (2001). The use of theory in information science research. *Journal of the American Society for Information Science and Technology*, 52, (1), 62-73. DOI: [https://doi.org/10.1002/1532-2890\(2000\)52:1<62::AID-ASI1061>3.0.CO;2-J](https://doi.org/10.1002/1532-2890(2000)52:1<62::AID-ASI1061>3.0.CO;2-J)
- Piñuel Raigada, J. L. (2002). Epistemología, metodología y técnicas del análisis de contenido. *Sociolinguistic Studies* 3 (1), 1-42. <https://doi.org/10.1558/sols.v3i1.1>
- Pinheiro, L. V. R.; Brascher, M.; Burnier, S. (2005). Ciência da Informação: 32 anos (1972-2004) no caminho da história e horizontes de um periódico científico brasileiro. *Ciência da Informação*, 34, (3), 38-49. Disponible en: <http://revista.ibict.br/ciinf/article/viewFile/1084/1187> [Fecha de consulta:27/06/2019]. <https://doi.org/10.1590/S0100-19652005000300003>
- Pinto, A. L.; Matias, M.; Moreiro-González, J. A. (2016). Produção brasileira da Ciência da Informação na Web of Science entre 1994 e 2013 e a lista Qualis/Capes da Área. *Ibersid*, 10, (1), 51-61. Disponible en: <https://www.ibersid.eu/ojs/index.php/ibersid/article/view/4252> [Fecha de consulta: 08/08/2019].
- Población, D. A.; De-Oliveira, M. (2006). Input e output: insumos para o desenvolvimento da pesquisa. En: Población, D. A.; Witter, G.; Da Silva, J. F. M. *Comunicação & produção científica: contexto, indicadores e avaliação*. São Paulo: Angellara, 57-80.

- Risso, V. G. (2012) Aproximación teórica a la relación entre los términos gestión documental, gestión de información y gestión del conocimiento. *Revista Española de Documentación Científica*, 35, (4), 531-554. <https://doi.org/10.3989/redc.2012.4.869>
- Romero, E. (2019). Países de América con mayor cantidad de profesionales SEO. *Karai digital*. Disponible en: <https://www.karaidigital.com/blog/profesionales-especialistas-seo-america> [Fecha de consulta:18/08/2019].
- Sánchez-Vigil, J. M.; Marcos-Recio, J. C.; Olivera-Zaldu; M. (2015). Producción científica sobre Comunicación y Medios en las revistas de Documentación (2000-2014). *Revista Española de Documentación Científica*. 38, (4), e108.
- Sayão, L. F.; Marcondes, C. H. (2008). O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *Transinformação*, 20, (2), 133-148. Disponible en: <https://www.redalyc.org/pdf/3843/384334798002.pdf> [Fecha de consulta:21/08/2019]. <https://doi.org/10.1590/S0103-37862008000200002>
- Vivaldi, J.; Rodríguez, H. 2007. Evaluation of terms and term extraction systems: A practical approach. *Terminology*, 13 (2). 225-248. DOI: <https://doi.org/10.1075/term.13.2.06viv>
- Wasserman, S.; Faust, K. (1994). *Social Network Analysis: methods and applications*. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511815478>
- Wellisch, H. (1972). From Information Science to Informatics: a terminological investigation. *Journal of Librarianship*, 4, (3), 157-187. <https://doi.org/10.1177/096100067200400302>
- White, M.; Marsh, E. (2006). Content analysis: a flexible methodology. *Library trends*, 55, (1), 22-45. DOI: <https://doi.org/10.1353/lib.2006.0053>

ESTUDIOS / RESEARCH STUDIES

Tendencias de las búsquedas de información sobre las colecciones SciELO, Redalyc y Dialnet realizadas a través de Google

Cristina Bojo-Canales*, María Sanz-Lorente**, Javier Sanz-Valero***

* Instituto de Salud Carlos III, Biblioteca Nacional de Ciencias de la Salud. Madrid.

Correo-e: cbojo@isciit.es | ORCID iD: <http://orcid.org/0000-0002-0330-171X>

** Universidad Miguel Hernández. Departamento de Salud Pública e Historia de la Ciencia, Campus de Sant Joan d'Alacant. Alicante. Conselleria de Sanidad Universal y Salud Pública de la Generalitat Valenciana. Centro de Salud Pública. Manises.

Correo-e: msanzlor@gmail.com | ORCID iD: <https://orcid.org/0000-0002-8735-0075>

*** Universidad Miguel Hernández. Departamento de Salud Pública e Historia de la Ciencia, Campus de Sant Joan d'Alacant. Alicante. Instituto de Salud Carlos III, Escuela Nacional de Medicina del Trabajo. Madrid. España.

Correo-e: fj.sanz@isciit.es | ORCID iD: <https://orcid.org/0000-0002-8735-0075>

Recibido: 30-03-20; 2ª versión: 19-05-20; Aceptado: 21-05-20; Publicado: 27-04-2021

Cómo citar este artículo/Citation: Bojo-Canales, C.; Sanz-Lorente, M.; Sanz-Valero, J (2021). Tendencias de las búsquedas de información sobre las colecciones SciELO, Redalyc y Dialnet realizadas a través de Google. *Revista Española de Documentación Científica*, 44 (2), e294. <https://doi.org/10.3989/redc.2021.2.1765>

Resumen: El objetivo del presente estudio fue analizar la tendencia de las búsquedas de información sobre las colecciones bibliográficas SciELO, Redalyc y Dialnet, que se realizan a través de Google. A partir de los resultados obtenidos se observó un claro predominio en las búsquedas, en su país de origen y filiación de cada una de las colecciones: SciELO en Brasil, Redalyc en México y Dialnet en España. La colección más buscada fue SciELO en todo el periodo de tiempo analizado y si bien se constató un crecimiento en el volumen de búsqueda relativo (VBR) tanto en las búsquedas sobre Redalyc y Dialnet, no sucedió así en el de SciELO. En ninguna de ellas se pudo observar un hito claramente destacable. Solo se pudo demostrar una moderada asociación entre los VBR de SciELO y Dialnet y leve entre los Redalyc y Dialnet.

Palabras clave: SciELO; Redalyc; Dialnet; volumen de búsqueda relativo; Google Trends.

Trends in the searching information on the collections SciELO, Redalyc and Dialnet conducted through Google.

Abstract: The present research was aimed to analyze the trend of information searches on SciELO, Redalyc and Dialnet bibliography collections conducted through Google. From the obtained results, it was possible to conclude that in their home country and affiliation of each collection it was observed a high incidence in the following searches: SciELO in Brasil, Redalyc in México and Dialnet in Spain. Throughout the analyzed period, the most sought-after collection was SciELO, although, it was verified a relative search volumen (RSV) both in Redalyc and Dialnet searches, this did not befall in the case of SciELO. In none of them has been possible to behold a remarkable landmark. Merely it could be shown a moderate association between the RSV of SciELO and Dialnet and slight between the Redalyc and Dialnet.

Keywords: SciELO; Redalyc; Dialnet; relative search volume; Google Trends

Copyright: © 2021 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1. INTRODUCCIÓN

La idea de que las poblaciones proporcionen datos sobre sus gustos, la búsqueda de servicios e incluso sobre su enfermedad, a través del comportamiento de búsqueda de información en la Web, ya ha sido explorada en los últimos años (Pelat y otros., 2009; Sanz-Lorente, 2020).

Es conocido que el buscador de Google recopila información relativa al comportamiento de los usuarios que navegan por su dominio y, por supuesto, también recoge los términos y expresiones que se utilizan para buscar información en el conocido buscador. Al análisis estadístico de los términos de búsqueda de los usuarios, ya sea en un sitio web particular, en una plataforma de red social o en un buscador web, se le denomina *search analytics* y forma parte del campo de la cibermetría (Orduña-Malea, 2019).

En el campo de la salud, Eysenbach (2009) acuñó el término de «infodemiología» como un conjunto emergente de métodos de información de salud pública para analizar el comportamiento de la búsqueda, la comunicación y la publicación en Internet. Es decir, observar y analizar el comportamiento basado en la Web para conocer la conducta humana a fin de predecir, evaluar, e incluso prevenir problemas relacionados con la salud que surgen constantemente en la vida cotidiana (Mavragani y otros, 2018).

Así pues, está más que demostrado que Google es un motor de búsqueda que facilita información a cualquier persona, que a través de los resultados obtenidos puede acceder fácilmente a los documentos existentes en la Red. Y, aunque no es el útil más conocido de Google, Google Trends (GT) es una herramienta de acceso libre, que informa del volumen de búsquedas realizadas por los usuarios de todo el mundo para exponer qué tan frecuentemente se busca un término y en qué lugares. Los datos de búsqueda en Internet pueden proporcionar información valiosa sobre los patrones de comportamiento de la población (Nuti y otros, 2014; Sanz-Lorente y Wanden-Berghe, 2018).

En los últimos años, los datos obtenidos a través de Google Trends han sido utilizados en investigaciones de diversos ámbitos y con distintas finalidades. Estos datos se han tomado como indicadores para estimar la incidencia de enfermedades (Sanz-Lorente y otros, 2019), para predecir niveles de desempleo (Redondo Caballero, 2013) o para analizar proyecciones de ventas (Belapatiño y Crispin, 2016).

Ahora bien, para los investigadores las revistas académico-científicas son uno de los principales canales de comunicación y difusión de los resultados de la investigación y las revistas más solicitadas depen-

den, en gran medida, de su calidad y su visibilidad. Cuanta más visibilidad tiene una revista, más interés despierta y, por ende, mayor cantidad de trabajos recibe para su publicación. Las bases de datos y otros repertorios de difusión de la literatura publicada desempeñan un rol central en este contexto, como vehículos que contribuyen a incrementar la visibilidad de las revistas y artículos, en tanto que la calidad es el filtro de editores y distribuidores para su inclusión en dichas fuentes (Villamón y otros, 2005).

Las colecciones digitales conservan y preservan la producción científica, favorecen su difusión, accesibilidad y utilización, además de aumentar su visibilidad e impacto en la comunidad científica. Por tanto, las colecciones de revistas serán foco de interés de profesionales e investigadores y en el ámbito iberoamericano hay tres que destacan sobremanera: SciELO, Redalyc y Dialnet.

Las tres plataformas analizadas son portales de revistas científicas de ámbito ibero-latinoamericano y carácter multidisciplinar, pero presentan características diferenciadoras.

SciELO y Redalyc surgen como una respuesta de la Región a la escasa visibilidad de sus revistas en las principales bases de datos, y el bajo impacto de la investigación realizada en estos países en los circuitos principales de la ciencia.

SciELO nace en el año 1998, en Brasil, bajo tres premisas básicas: aumentar la visibilidad de las revistas brasileñas mediante la publicación electrónica y el libre acceso a los textos, proporcionar indicadores bibliométricos sobre la ciencia producida en la región y mejorar la calidad de sus revistas científicas. El modelo, poco a poco, se fue extendiendo por la región y hoy SciELO funciona como una red descentralizada de hemerotecas científicas que abarca 14 países y tres continentes (<http://www.scielo.org>). En conjunto, la red indiza más de 1.440 revistas y da acceso a más de 700.000 artículos, con un incremento medio de más de 40.000 artículos por año (Packer y otros, 2014; Packer y Meneghini, 2015).

Algunos años después, en 2003, la Universidad Autónoma de México, pone en marcha el sitio Redalyc (<https://www.redalyc.org/>) bajo unas premisas muy similares a las de SciELO, pero funcionando como una única hemeroteca virtual de libre acceso, no como una red, y más centrada en reunir la producción más relevante de las Ciencias Sociales, mientras que SciELO nació con un carácter biomédico. Hoy ambas plataformas son multidisciplinarias y presentan cifras muy similares, Redalyc cuenta con 1360 revistas y 670.000 artículos (Aguado-López y otros, 2003).

Dialnet tiene un origen distinto, nace en el año 2001, bajo otros planteamientos. Se crea con la idea de dar un servicio de alertas bibliográficas accesible a cualquier usuario, poco a poco se convierte en una hemeroteca virtual y un servicio de alojamiento de textos, todo ello con vocación integradora en el sentido de almacenar distintos tipos de documentos (científicos, técnicos, divulgativos, etc.), algunos de manera referencial y otros a texto completo. Dialnet recoge artículos de revistas, tesis, actas de congresos, capítulos de libros...etc., publicados en España o en español en cualquier país. En este caso con claro predominio de revistas de Humanidades y Ciencias Sociales (Mateo, 2015).

Desde hace años, los motores de búsqueda son el punto de partida para la navegación web, permitiendo al usuario encontrar contenidos en Internet utilizando una serie de términos. Google, fundado 1997, no necesita presentación, es el motor de búsqueda más utilizado en el mundo, con una cuota de mercado que rebasa ampliamente a otros buscadores como Baidu o Yandex (los más usados en China y Rusia respectivamente) (Amazon Company, sin fecha, Net Market Share, 2020). En este sentido, es conocido que una gran mayoría de los estudiantes universitarios norteamericanos empiezan sus búsquedas de información en buscadores web, en vez de utilizar algún recurso de la biblioteca (Ávila-García y otros, 2015). Actitud que es compartida por muchas personas que, por olvido de la dirección electrónica (*Uniform Resource Locator*), o simplemente por comodidad, utilizan Google como puerta de inicio de su búsqueda.

Con este potencial, como indicó Orduña-Malea (2019), GT puede ser utilizado en tareas de *new-casting* (predecir el presente, es decir, pronosticar valores que están sucediendo al mismo tiempo que se generan los datos), y de *forecasting* (predicción de tendencias futuras). Así, esta herramienta puede ayudar a los responsables de estas 3 plataformas a conocer los productos más buscados, cuáles son las nuevas tendencias, y qué necesidades están demandando los usuarios (Morán, 2016).

En este contexto, el objetivo de este estudio fue analizar la tendencia de las búsquedas de información y los hitos, sobre estas tres colecciones, que se realizan a través de Google.

2. MATERIAL Y MÉTODO

2.1 Diseño

Estudio descriptivo y correlacional de las tendencias de búsqueda de información sobre SciELO, Redalyc y Dialnet, a través de Google y en todo el mundo.

2.2 Fuente de obtención de la información

Los datos sobre búsqueda de información se obtuvieron de la consulta directa, mediante acceso online, a Google Trends (<https://trends.google.es/>).

El periodo a estudio fue del 1 de enero de 2004 hasta el 31 de diciembre de 2019. El ámbito fue todo el mundo y en todas las categorías. El periodo estudiado es común a las tres colecciones, puesto que SciELO se crea en el año 1998, Dialnet en 2001 y Redalyc surge en el año 2002. La fecha de consulta y recolección de los datos fue el 17 de febrero de 2020.

2.3 Tema de búsqueda

Las búsquedas se realizaron con las palabras «SciELO», «Redalyc» y «Dialnet» como «Tema» y tomando los resultados para todo el mundo y en todas las categorías.

Hay que tener en cuenta que en esta herramienta la consulta utilizando un «Término de búsqueda» muestra las coincidencias de todos los términos de la consulta en el idioma especificado (Ejemplo: si se busca «salud» se obtendrán también resultados sobre «salud pública», «ciencias de la salud», etc.). En cambio, si se utiliza el vocablo como «Tema» se obtendrán los resultados de los términos que comparten el mismo concepto en cualquier idioma. (Ejemplo: si se busca «Londres», la búsqueda incluye resultados de temas como los siguientes «capital del Reino Unido», «London» o «London city», entre otros).

2.4 Obtención y almacenamiento de los datos

Los resultados obtenidos fueron descargados en formato normalizado CSV (*comma-separated values*) que permitió su almacenamiento posterior en un archivo Excel.

Las gráficas GT representan con cuánta frecuencia se realiza una búsqueda de un término particular en varias regiones del mundo y en varios idiomas. El eje horizontal de la gráfica representa el tiempo (desde el año 2004), y el eje vertical representa la frecuencia, como volumen de búsqueda relativo (VBR), con la que se ha buscado el término globalmente, permitiendo al usuario comparar el VBR entre dos o más términos (actualmente un máximo de 5).

El control de la calidad de esta información se efectuó mediante dobles tablas, corrigiendo las posibles inconsistencias mediante la consulta con la tabla original descargada. Para el análisis estadístico se utilizó el programa *Statistical Package for the Social Sciences* (SPSS) para Windows versión 25.0.

2.5 Variables a estudio

- Volumen de búsqueda relativo (VBR): resultado mensual ofrecido por Google Trends cuyos valores se normalizan en una escala de 0 (un volumen de búsqueda relativo inferior al 1% del volumen máximo) a 100 (el volumen de búsqueda relativo alcanza su máximo). Por ejemplo, un VBR = 25, representa el 25% de la proporción de búsqueda observada más alta durante el período a estudio.
- Hito: acontecimiento puntual y destacado en el VBR.

2.6 Análisis de los datos

Se obtuvieron las medidas de tendencia central (media y su desviación estándar, la mediana, la amplitud intercuartílica (AIQ) y el máximo y mínimo) para describir las variables a estudio. Se analizó la existencia de asociación entre las variables cuantitativas para comprobar la significación de la diferencia de medias para muestras independientes, utilizando el test de la t de Student. La evolución temporal de las tendencias de búsqueda se examinó mediante el análisis de regresión,

calculándose el coeficiente de determinación (R^2). Para obtener la relación entre variables cuantitativas se usó el coeficiente de correlación de Pearson (medida de dependencia lineal entre dos variables aleatorias cuantitativas).

El nivel de significación utilizado en todos los contrastes de hipótesis fue $\alpha \leq 0,05$.

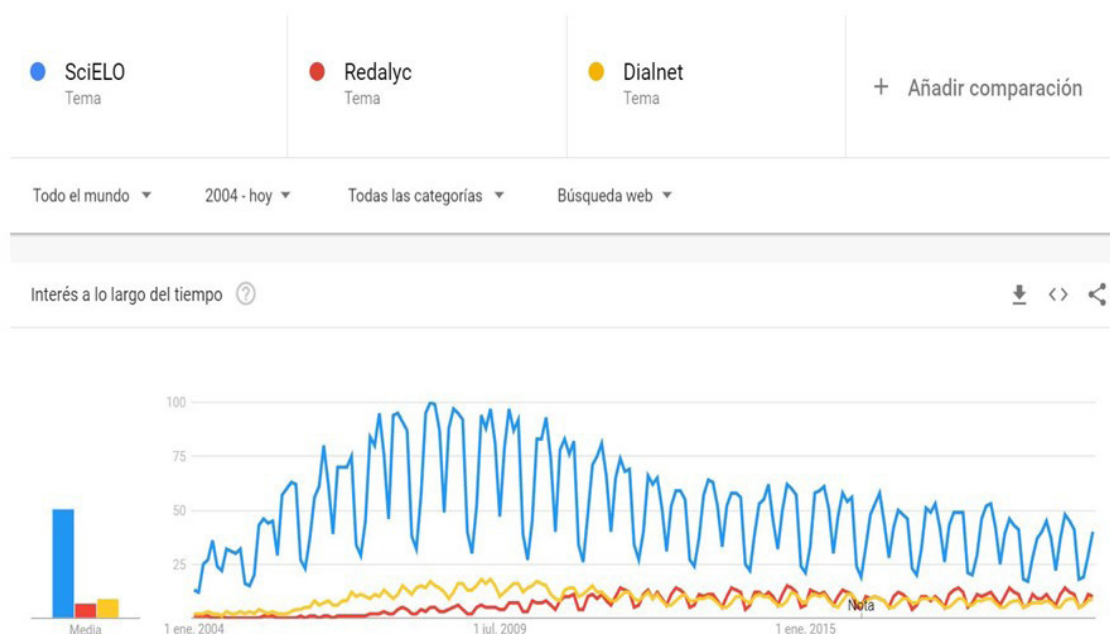
Para analizar la evolución del VBR mensual se segmentó el período de estudio en dos épocas iguales: 1ª época: del 1 enero de 2004 al 30 junio de 2011 y 2ª época: del 1 julio de 2011 al 31 de diciembre de 2019.

Para todos los cálculos relacionados con el VBR se tuvieron en cuenta los valores mensuales de cada uno de los años; ver Anexo I.

3. RESULTADOS

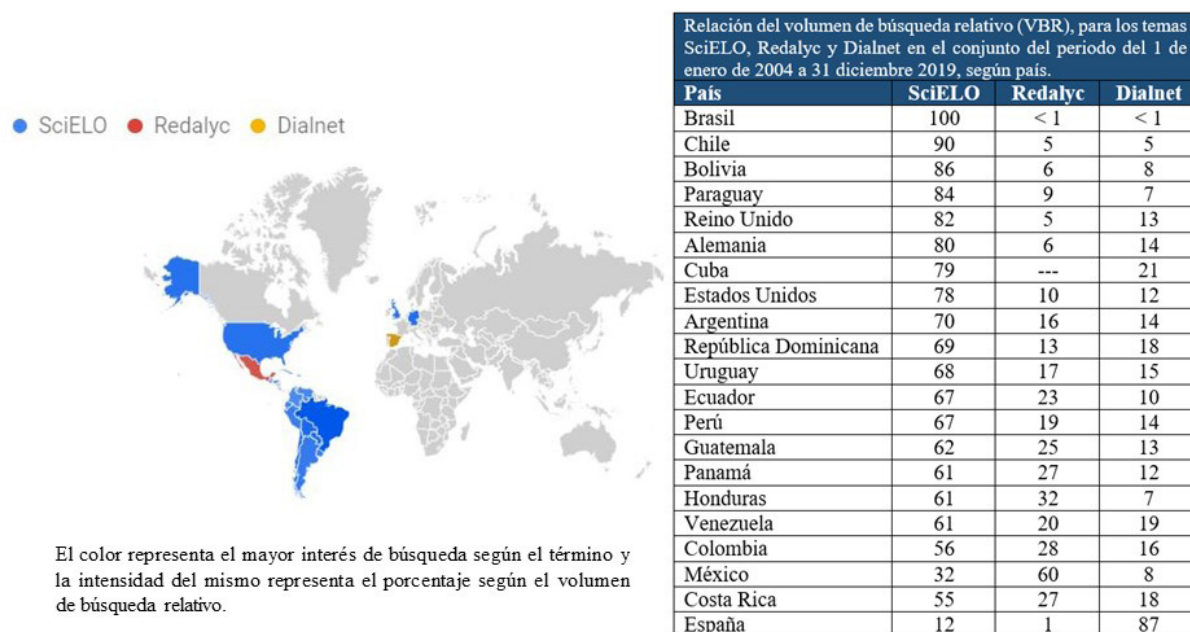
A partir de la consulta efectuada se obtuvieron los valores mensuales de los VBR para los temas «SciELO», «Redalyc» y «Dialnet» obtenidos de Google Trends, ver Anexo I, y las correspondientes

Figura 1. Tendencia de búsqueda en Google Trends para los temas SciELO, Redalyc y Dialnet (desde el 1 ene de 2004 a 31 dic de 2019).



Sumatorio de los volúmenes relativos de búsqueda mensuales, según año, para cada de los temas «SciELO», «Redalyc» y «Dialnet», obtenidos de Google Trends

Año	SciELO	Redalyc	Dialnet	Año	SciELO	Redalyc	Dialnet	Año	SciELO	Redalyc	Dialnet	Año	SciELO	Redalyc	Dialnet
2004	300	12	26	2005	511	12	34	2006	677	12	89	2007	860	30	134
2008	931	46	167	2009	883	60	175	2010	798	86	156	2011	717	113	132
2012	594	120	107	2013	579	115	97	2014	561	125	107	2015	552	102	106
2016	492	102	92	2017	484	112	91	2018	454	118	85	2019	433	118	88

Figura 2. Desglose comparativo por país de los resultados globales de los temas SciELO, Redalyc y Dialnet (desde el 1 ene de 2004 a 31 dic de 2019), obtenidos de Google Trends.

tendencias de búsqueda para cada una de las plataformas; ver captura de pantalla en la Figura 1.

Al mismo tiempo, y tomando la imagen ofrecida por esta herramienta, se obtuvo el desglose comparativo de los resultados globales para los 3 temas a estudio, observándose el interés principal según país; ver captura de pantalla en la Figura 2.

3.1 Hitos

El principal hecho puntual en el VBR se produjo en la consulta sobre «SciELO» en abril de 2008 (VBR = 100). La relación de los VBR en este hito entre «SciELO» *versus* «Redalyc» *versus* «Dialnet» fue: 100/5/17.

3.2 Volumen de búsqueda relativo

Los estadísticos de los VBR para los términos analizados pueden consultarse en la Tabla 1, sien-

do la correspondencia media de los VBR entre «SciELO» *versus* «Redalyc» *versus* «Dialnet» de 51,18/6,73/8,78.

La tendencia de VBR para «SciELO» presentó un gráfico con progresión decreciente y ajuste a un modelo lineal ($R^2 = 0,05$; $p = 0,001$); para «Redalyc» se obtuvo una evolución moderadamente creciente con ajuste lineal ($R^2 = 0,58$; $p < 0,001$); y Dialnet mostró un suave progreso exponencial creciente ($R^2 = 0,06$; $p = 0,001$); ver Figura 3.

Al analizar los datos del VBR para «SciELO», según época, se constató que existían diferencias significativas ($p < 0,001$) a favor de la primera época (valores de las medias: 59,19 *versus* 44,11). Igualmente, se mostraron diferencias significativas ($p = 0,021$) a favor de la primera época en los VBR para «Dialnet» (valores de las medias: 9,49 *versus* 8,16). Por el contrario, los valores del VBR

Tabla 1. Estadísticos del Volumen de Búsqueda Relativo (VBR) de los temas «SciELO», «Redalyc» y «Dialnet», obtenidos de Google Trends

Término	Media	Mediana	Máximo	Mínimo	AIQ
SciELO	51,18 ± 1,60	48,50	100	12	31
Redalyc	6,73 ± 0,31	6	15	0	8
Dialnet	8,78 ± 0,27	9	18	1	5
AIQ = Amplitud Intercuartílica					

Figura 3. Evolución temporal de los resultados mensuales para SciELO, Redalyc y Dialnet (desde el 1 ene de 2004 a 31 dic de 2019), obtenidos de Google Trends.

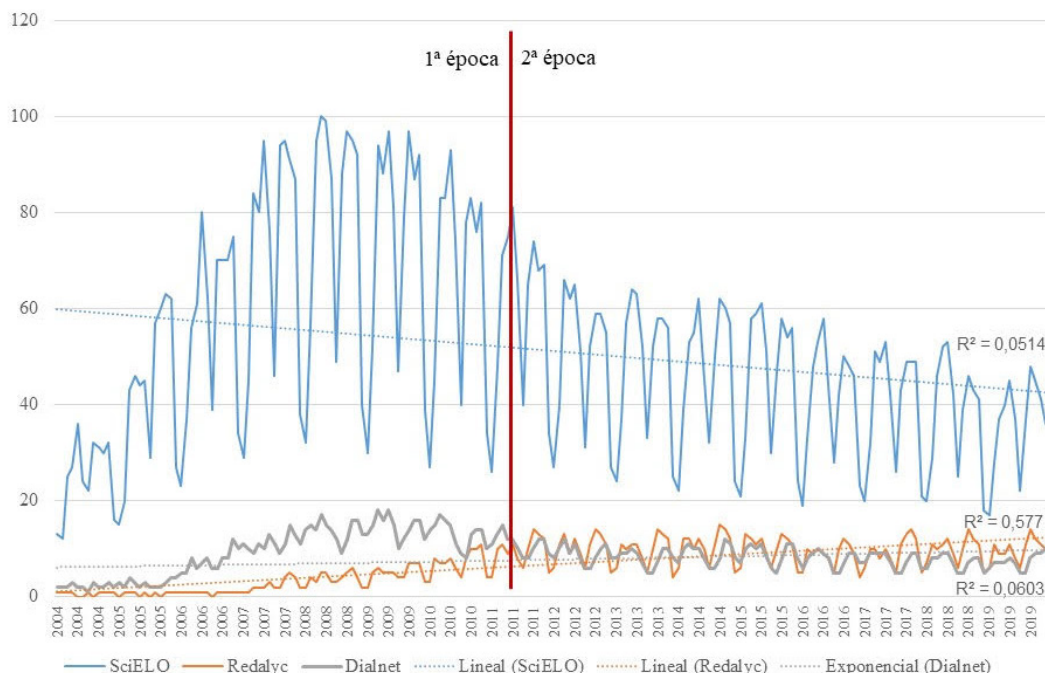
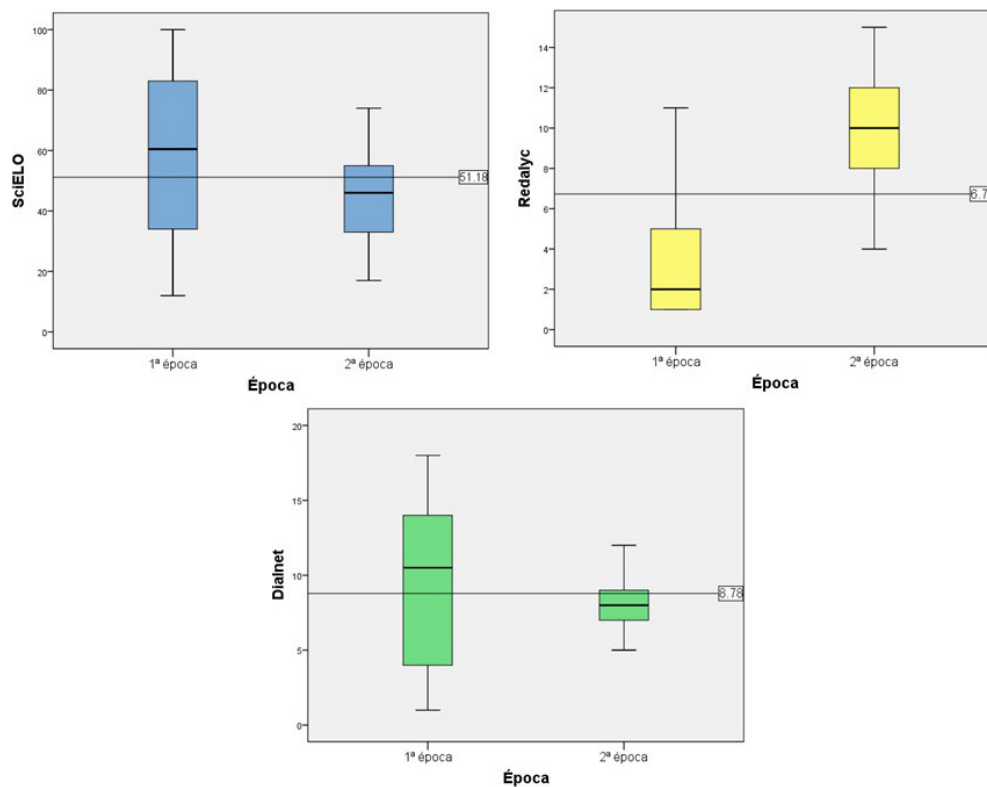


Figura 4. Comparación de las medias del volumen de búsqueda relativo (VBR), para los temas SciELO, Redalyc y Dialnet, obtenidos de Google Trends, según época a estudio (1ª época: 1 de enero de 2004 a 30 de junio de 2011; 2ª época de 1 de julio de 2011 a 31 de diciembre de 2019).



para «Redalyc» dieron diferencias significativas ($p < 0,001$) a favor de la segunda época (valores de las medias: 3,46 *versus* 9,71); ver Figura 4.

3.3 Asociación entre los diferentes VBR

La correlación observada entre los VBR de los temas «SciELO» y «Redalyc» no mostró asociación entre ellos ($R = 0,13$; $p = 0,076$). Por el contrario, si existió buena relación entre los VBR de los temas «SciELO» y «Dialnet» ($R = 0,66$, $p < 0,001$) y leve asociación entre los VBR de temas «Redalyc» y «Dialnet» ($R = 0,26$; $p < 0,001$).

4. DISCUSIÓN

A la vista de los resultados se puede deducir que las tendencias de los motores de búsqueda son una herramienta que puede integrar la vigilancia, en tiempo real, de las necesidades de información de la población. Así, Anderegg y Goldsmith, en 2014, ya afirmaban que Google Trends era una herramienta que se reconocía como indicador sólido y válido para la predicción del patrón de comportamiento en la búsqueda de información (Anderegg y Goldsmith, 2014).

Pero, es importante dejar constancia de que este estudio no pretendió comparar la cobertura de estas plataformas, ni discutir la validez de los contenidos que indizan. El valor del análisis de los VBR propuesto está en conocer cuáles son las búsquedas de los usuarios, de todo el mundo, en un determinado periodo de tiempo y ver cómo evolucionan los comportamientos sociales.

En el desglose comparativo por país de los resultados globales, quedó patente el predominio de cada una de estas colecciones bibliográficas en su país de origen y filiación, situación que confirmó la conjetura de partida y por la correspondencia media de los VBR. Igualmente, era esperable que SciELO tuviera un claro predominio en Latinoamérica. Según Ugarte (2004), la indización en SciELO constituyó, y constituye, la principal herramienta para que esta producción científica saliera de la literatura gris y accediera a los índices de medición bibliométrica clásicos, fundamentalmente por las aportaciones de Brasil, Argentina y Chile (Aguirre-Pitol y otros, 2013).

La evolución temporal de los resultados mensuales para las búsquedas sobre las tres colecciones también mostró un mayor VBR a favor de SciELO, si bien ninguna de las colecciones presentó un claro hito que permitiera fijar un hecho concreto, si acaso el máximo que se dio en 2008, año que puede relacionarse con el décimo aniversario de la creación de SciELO. Lo que más llama la aten-

ción es la forma en diente de sierra que se observa principalmente en el VBR de SciELO, típico de los gráficos de las enfermedades epidemiológicas, con máximos que se centran entre abril y mayo de cada año y mínimos en diciembre y enero. Una posible explicación de esta tendencia podría relacionarse con la actividad académica que durante parte de diciembre y enero disminuye su actividad. Ahora bien, los dientes de sierra también son una visualización que permite reconocer gráficamente la relación entre turnos de un diálogo, una tipología de figuras interaccionales que recoge las diferentes estructuras que se producen en la toma de turnos de la conversación coloquial (sea verbal, escrita o digital) (Espinosa Guerri, 2019).

La tendencia observada en el VBR de Redalyc y Dialnet sigue el progresivo incremento en la producción científica de las revistas del área latinoamericana (Bojo-Canales y Sanz-Valero, 2019). El VBR observado en estas dos colecciones, no alcanzó la llamada «explosión de la información» que lleva a modelos de crecimiento exponencial (Sanz-Valero y Wanden-Berghe, 2017), ya que sus crecimientos solo pueden ser calificados de leves o, como mucho, moderados. De todos modos, estas tendencias no se basan en datos de crecimiento o decrecimiento absolutos, pues el indicador estudiado, VBR, como su mismo nombre indica ofrece valores relativos.

Situación diferente son los VBR para la colección SciELO; la evolución temporal de los VBR dio, sorprendentemente, un resultado de progresión decreciente contrario a lo esperado, teniendo en cuenta el incremento en la producción científica que recoge esta colección (Bojo-Canales y Sanz-Valero, 2019). Una posible explicación sea que al ser una colección ampliamente conocida y, por tanto, como se recoge asimismo en las limitaciones de este estudio, muchos profesionales e investigadores accedan directamente desde la *homepage* de SciELO sin necesidad de localizarla a través de Google. Aun así, hace ya tiempo que los contenidos de las plataformas, como las analizadas aquí, son “recolectados” por Google, de modo que, ante cualquier búsqueda de información un usuario podría encontrar, entre los resultados ofrecidos, un registro que estuviera recogido por una de estas 3 plataformas.

La observación de una moderada correlación entre los VBR de SciELO y Dialnet y leve entre los Redalyc y Dialnet se debería, principalmente, a que en la mayoría de las veces cuando se busca el término correspondiente a una de estas colecciones no se hace con los otros dos. Por otra parte, puede no existir una misma motivación para buscar en las tres ya que no todas las revistas están incluidas a la vez en las tres ni existe un equilibrio temático;

por ejemplo, Redalyc tiene un mayor sesgo hacia las Ciencias Sociales mientras que SciELO lo tiene hacia las Ciencias de la Salud (Miguel, 2011).

5. LIMITACIONES

De acuerdo con Johnson y Mehta (2014), el que Google Trends no facilite datos de uso real e intervalos de tiempo más precisos, disminuye la capacidad de pronóstico. Hay que tener presente que se trataba de un análisis de datos ecológicos y los hallazgos pudieron no ser representativos a nivel individual. Además, como sugieren Cervellin y otros (2017), los resultados que se obtienen mediante esta herramienta pueden estar influenciados por el interés de los medios.

Un hecho importante es que GT ofrece valores relativos, pero no frecuencias absolutas (número total de búsquedas), lo que disminuye la capacidad de pronóstico y de análisis estadísticos más reales. Además, se echa en falta una mayor transparencia, ya que no existe información sobre los métodos específicos que Google utiliza para pronosticar las tendencias, los cuales no han sido divulgados por la compañía (Orellano y otros, 2015).

Por otro lado, puede presentarse un sesgo debido a que la búsqueda se realizó solo en Google y no se contabilizaron las realizadas en otros buscadores o incluso mediante la consulta a las hemerotecas. Más aún, muchos profesionales acceden directamente a estas colecciones sin pasar, previamente, por la búsqueda utilizando Google, por lo que estas búsquedas directas, que determinan el uso y utilidad de dichos portales, no quedan reflejadas en este artículo. No obstante, el estudio de Gardner e Inger (2018), sobre los hábitos de búsqueda de los lectores de literatura científica, señalaba como resultado importante que mientras que las bases bibliográficas seguían siendo la fuente más relevante, su importancia venía disminuyendo desde 2008, perdiendo posiciones frente a los motores de búsqueda, las redes sociales y los servicios de agregadores.

8. REFERENCIAS

- Aguado-López, E.; Sandoval Forero, E.; Chávez-Ávila, S. (2003). La ciencia perdida y las nuevas tecnologías de divulgación del conocimiento: el proyecto Redalyc. *Revista Educación y Ciencia*, 7(13), 11-39. <http://hdl.handle.net/20.500.11799/6618>
- Aguirre-Pitol, M. Á.; Leal-Arriola, M.; Martínez-Domínguez, N. D. (2013). *Cápsulas de Investigación nº3—Análisis comparativo de la cobertura de SciELO y Redalyc* [en línea]. Disponible en: <https://bit.ly/3dv-vWHm> [Fecha de consulta: 15 febrero de 2020]
- Amazon Company. (s. f.). Alexa, The top 500 sites on the Web [en línea]. Disponible en: <http://www.alexa.com/topsites> [Fecha de consulta: 17/05/2020]

6. CONCLUSIONES

El uso de una herramienta como Google Trends puede proporcionar una valiosa información a los gestores y administradores de estas plataformas, sobre las búsquedas no dirigidas ni especializadas, lo que puede ser muy útil para implementar estrategias de visibilización de sus contenidos. Estas tácticas deben estar fundamentadas en datos de investigación que proporcionen la información adecuada para desarrollar el mejor material posible. Además, pueden conocer el interés de la población en unos determinados temas y en un momento determinado.

En este estudio se observó un claro predominio en las búsquedas en su país de origen y filiación de cada una de las colecciones: SciELO en Brasil, Redalyc en México, y Dialnet en España. La colección más buscada fue SciELO en todo el periodo de tiempo analizado y si bien se constató un crecimiento en el VBR tanto en las búsquedas sobre Redalyc y Dialnet, no sucedió así en el caso de SciELO. Y, en ninguna de ellas se pudo observar un hito claramente destacable. Solo se pudo demostrar una moderada asociación entre los VBR de SciELO y Dialnet y leve entre los Redalyc y Dialnet.

7. AGRADECIMIENTOS

A Habiba Chbab, con maestría en inglés y español para fines específicos y estudiante de doctorado en Traducción profesional y audiovisual (rama de investigación: traducción médica), por su colaboración en la traducción al inglés del resumen de este artículo.

ACKNOWLEDGEMENTS

To Habiba Chbab, master's degree in English and Spanish for Specific Purposes and doctoral student in Professional and Audiovisual Translation (Research branch: medical translation), for her collaboration in the English translation of the abstract of this article.

- Anderegg, W. R. L.; Goldsmith, G. R. (2014). Public interest in climate change over the past decade and the effects of the 'climategate' media event. *Environmental Research Letters*, 9(5), 054005. <https://doi.org/10.1088/1748-9326/9/5/054005>
- Ávila-García, L.; Ortiz-Repiso, V.; Rodríguez-Mateos, D. (2015). Herramientas de descubrimiento: ¿una ventanilla única? *Revista Española de Documentación Científica*, 38(1):e077. <http://dx.doi.org/10.3989/redc.2015.1.1178>
- Belapatiño, V.; Crispin, Y. (2016). *Perú—Búsquedas en internet y proyecciones de ventas de departamentos en Lima* [en línea]. Disponible en: <https://bit.ly/3bkLo7i> [Fecha de consulta: 15 febrero de 2020]

- Bojo-Canales, C.; Sanz-Valero, J. (2019). Las revistas de ciencias de la salud de la red SciELO: Un análisis de su visibilidad en el ámbito internacional. *Revista Española de Documentación Científica*, 42(4), 245. <https://doi.org/10.3989/redc.2019.4.1629>
- Cervellin, G.; Comelli, I.; Lippi, G. (2017). Is Google Trends a reliable tool for digital epidemiology? Insights from different clinical settings. *Journal of Epidemiology and Global Health*, 7(3), 185-189. <https://doi.org/10.1016/j.jegh.2017.06.001>
- Espinosa Guerri, G. (2019). Dibujar la conversación en ELE: La enseñanza de la toma de turnos del español. *Foro de profesores de E/LE*, 15. <https://doi.org/10.7203/foroele.15.14860>
- Eysenbach, G. (2009). Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *Journal of Medical Internet Research*, 11(1), e11. <https://doi.org/10.2196/jmir.1157>
- Gardner, T.; Inger, S. (2018). *How readers discover content in scholarly publications* [en línea]. Disponible en: <https://bit.ly/2Ufk3xL> [Fecha de consulta: 15 febrero de 2020]
- Johnson, A. K.; Mehta, S. D. (2014). A comparison of Internet search trends and sexually transmitted infection rates using Google trends. *Sexually Transmitted Diseases*, 41(1), 61-63. <https://doi.org/10.1097/OLQ.000000000000065>
- Mateo, F. (2015). Producción científica en español en humanidades y ciencias sociales: Algunas propuestas desde Dialnet. *El profesional de la Información*, 24(5), 509-515. <https://doi.org/10.3145/epi.2015.sep.01>
- Mavragani, A.; Ochoa, G.; Tsagarakis, K. P. (2018). Assessing the methods, tools, and statistical approaches in Google Trends research: Systematic review. *Journal of Medical Internet Research*, 20(11), e270. <https://doi.org/10.2196/jmir.9366>
- Miguel, S. (2011). Revistas y producción científica de América Latina y el Caribe: Su visibilidad en SciELO, RedALyC y SCOPUS. *Revista Interamericana de Bibliotecología*, 34(2), 187-199.
- Morán, J. (2016). Google Trends: una nueva herramienta para la predicción económica – Aplicaciones para completar el análisis econométrico tradicional [en línea]. Disponible en: <http://hdl.handle.net/10908/12083> [Fecha de consulta: 10 mayo de 2020]
- Net Market Share. (2020). Search Engine Market Share [en línea]. Disponible en: <https://bit.ly/3fYlm8k> [Fecha de consulta: 17/05/2020]
- Nuti, S. V.; Wayda, B.; Ranasinghe, I.; Wang, S.; Dreyer, R. P.; Chen, S. I.; Murugiah, K. (2014). The use of google trends in health care research: A systematic review. *PloS One*, 9(10), e109583. <https://doi.org/10.1371/journal.pone.0109583>
- Orduña-Malea, E. (2019). Google Trends: Analítica de búsquedas al servicio del investigador, del profesional y del curioso de la información. *Anuario ThinkEPI*, 13, e13inf01. <https://doi.org/10.3145/thinkepi.2019.e13inf01>
- Orellano, P.W.; Reynoso, J.I.; Antman, J.; Argibay, O. (2015). Uso de la herramienta Google Trends para estimar la incidencia de enfermedades tipo influenza en Argentina. *Cadernos de Saude Publica*, 31(4):691-700. <https://doi.org/10.1590/0102-311x00072814>
- Packer, A.L.; Cop, N.; Luccisano, A.; Ramalho, A.; Spinak, E. (2014). *SciELO: 15 Años de Acceso Abierto - Un estudio analítico sobre Acceso Abierto y comunicación científica*. Paris, Francia: UNESCO. <http://dx.doi.org/10.7476/9789233012370>
- Packer, A.L.; Meneghini, R. Contribución de SciELO a la Globalización de la Ciencia [en línea]. Disponible en: <https://bit.ly/2ZIF6St> [Fecha de consulta 17/05/2020]
- Pelat, C.; Turbelin, C.; Bar-Hen, A.; Flahault, A.; Valleron, A.-J. (2009). More diseases tracked by using Google Trends. *Emerging Infectious Diseases*, 15(8), 1327-1328. <https://doi.org/10.3201/eid1508.090299>
- Redondo Caballero, J. (2013). *Uso de Google Trends para predecir el nivel y la estructura del desempleo en España* [tesis doctoral] [en línea]. Universitat Politècnica de València, Facultad de Administración y Dirección de Empresas. Disponible en: <http://hdl.handle.net/10251/31028> [Fecha de consulta: 15 febrero de 2020]
- Sanz-Lorente, M. (2020). Tendencias temporales de los patrones de búsqueda de información sobre servicio de asistencia sanitaria domiciliaria en España. *Hospital a Domicilio*, 4(1), 15-23. <https://doi.org/10.22585/hospdomic.v4i1.95>
- Sanz-Lorente, M.; Sanz-Valero, J.; Wanden-Berghe, C. (2019). Tendencias temporales de los patrones de búsqueda de información sobre VIH/sida en España. *Revista Española de Comunicación Salud*, Supl 2, S52-S60. <https://doi.org/10.20318/recs.2019.4554>
- Sanz-Lorente, M.; Wanden-Berghe, C. (2018). Tendencias temporales de los patrones de búsqueda de información sobre cuidado domiciliario "Home Care" u hospitalario "Hospital Care" a través de Google. *Hospital a Domicilio*, 2(3), 93-99. <https://doi.org/10.22585/hospdomic.v2i3.47>
- Sanz-Valero, J.; Wanden-Berghe, C. (2017). Análisis bibliométrico de la producción científica, indizada en MEDLINE, sobre los servicios de salud proporcionados por las unidades de hospitalización a domicilio. *Hospital a Domicilio*, 1(1), 21-34. <https://doi.org/10.22585/hospdomic.v1i1.3>
- Ugarte P, F. (2004). La edición de revistas científicas en Latinoamérica. *Revista Chilena de Pediatría*, 75(6), 509-511. <https://doi.org/10.4067/S0370-41062004000600001>
- Villamón, M.; Devis-Devis, J.; Valenciano-Valcárcel, J. (2005). Análisis de la visibilidad de las revistas científico-técnicas españolas de ciencias de la actividad física y el deporte. *Revista de Psicología del Deporte*, 14(2), 253-267.

ANEXO I

Valores mensuales de los volúmenes de búsqueda relativos para los temas «SciELO», «Redalyc» y «Dialnet», obtenidos de Google Trends

Año	Mes	Volumen de búsqueda relativo		
		SciELO	Redalyc	Dialnet
2004	1	13	1	2
	2	12	1	2
	3	25	1	2
	4	27	1	3
	5	36	1	2
	6	24	1	2
	7	22	1	1
	8	32	1	3
	9	31	1	2
	10	30	1	2
	11	32	1	3
	12	16	1	2
2005	1	1	15	1
	2	2	20	1
	3	3	43	1
	4	4	46	1
	5	5	44	1
	6	6	45	1
	7	7	29	1
	8	8	57	1
	9	9	60	1
	10	10	63	1
	11	11	62	1
	12	12	27	1
2006	1	1	23	1
	2	2	37	1
	3	3	56	1
	4	4	61	1
	5	5	80	1
	6	6	62	1
	7	7	39	1
	8	8	70	1
	9	9	70	1
	10	10	70	1
	11	11	75	1
	12	12	34	1

Año	Mes	Volumen de búsqueda relativo		
		SciELO	Redalyc	Dialnet
2007	1	1	29	1
	2	2	45	1
	3	3	84	2
	4	4	80	2
	5	5	95	2
	6	6	76	3
	7	7	46	2
	8	8	94	2
	9	9	95	4
	10	10	91	5
	11	11	87	4
	12	12	38	2
2008	1	32	2	14
	2	57	4	15
	3	95	3	14
	4	100	5	17
	5	99	5	15
	6	87	3	14
	7	49	3	12
	8	88	4	9
	9	97	5	12
	10	95	6	16
	11	92	4	16
	12	40	2	13
2009	1	30	2	13
	2	52	5	15
	3	94	6	18
	4	88	5	16
	5	97	5	18
	6	81	5	15
	7	47	4	10
	8	79	4	12
	9	97	7	14
	10	87	7	16
	11	92	7	16
	12	39	3	12

Año	Mes	Volumen de búsqueda relativo		
		SciELO	Redalyc	Dialnet
2010	1	27	3	14
	2	45	8	15
	3	83	7	17
	4	83	7	16
	5	93	8	15
	6	74	6	11
	7	40	4	9
	8	78	8	8
	9	83	10	13
	10	76	10	14
	11	82	11	14
	12	34	4	10
2011	1	26	4	11
	2	48	10	13
	3	71	11	15
	4	75	9	12
	5	81	11	12
	6	66	8	10
	7	40	6	8
	8	65	10	8
	9	74	14	10
	10	68	13	12
	11	69	12	12
	12	34	5	9
2012	1	27	6	8
	2	40	11	10
	3	66	13	12
	4	62	9	9
	5	65	12	11
	6	51	9	8
	7	31	6	6
	8	52	11	6
	9	59	14	8
	10	59	13	10
	11	55	11	11
	12	27	5	8

Año	Mes	Volumen de búsqueda relativo		
		SciELO	Redalyc	Dialnet
2013	1	24	6	8
	2	37	11	9
	3	57	10	9
	4	64	11	10
	5	63	11	9
	6	52	8	7
	7	33	5	5
	8	52	10	5
	9	58	14	7
	10	58	13	10
	11	56	12	10
	12	25	4	8
2014	1	22	6	7
	2	39	12	10
	3	53	12	11
	4	55	10	10
	5	62	12	10
	6	45	10	8
	7	32	6	6
	8	50	11	6
	9	62	15	8
	10	60	14	12
	11	57	12	11
	12	24	5	8
2015	1	21	6	7
	2	33	13	10
	3	58	12	11
	4	59	11	10
	5	61	12	11
	6	51	9	8
	7	30	6	6
	8	47	10	5
	9	58	13	8
	10	54	12	11
	11	56	11	11
	12	24	5	8

Año	Mes	Volumen de búsqueda relativo		
		SciELO	Redalyc	Dialnet
2016	1	19	5	6
	2	33	10	8
	3	48	9	9
	4	53	10	10
	5	58	9	9
	6	44	8	8
	7	28	5	5
	8	42	10	5
	9	50	12	7
	10	48	11	9
	11	46	9	9
	12	23	4	7
2017	1	20	6	7
	2	32	10	9
	3	51	10	9
	4	49	8	9
	5	53	10	9
	6	42	8	7
	7	26	5	5
	8	43	11	5
	9	49	13	7
	10	49	14	9
	11	49	12	9
	12	21	5	6

Año	Mes	Volumen de búsqueda relativo		
		SciELO	Redalyc	Dialnet
2018	1	20	7	6
	2	29	11	8
	3	46	10	8
	4	52	11	9
	5	53	12	9
	6	42	9	7
	7	25	6	5
	8	39	10	5
	9	46	14	7
	10	43	12	8
	11	41	11	8
	12	18	5	5
2019	1	17	6	6
	2	28	11	7
	3	37	9	7
	4	40	9	7
	5	45	11	8
	6	37	8	7
	7	22	6	5
	8	37	11	5
	9	48	14	8
	10	45	12	9
	11	41	11	9
	12	36	10	10

ESTUDIOS / RESEARCH STUDIES

Evaluación de la investigación con encuestas en artículos publicados en revistas del área de Biblioteconomía y Documentación

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

* Departamento de Ciencias de la Documentación e Historia de la Ciencia. Universidad de Zaragoza.

Correo-e: jaso@unizar.es | ORCID iD: <http://orcid.org/0000-0001-8568-3098>

Correo-e: gmarco@unizar.es | ORCID iD: <http://orcid.org/0000-0002-7149-6192>

** Departamento de Biblioteconomía y Documentación. Facultad de Ciencias de la Documentación.
Universidad Complutense de Madrid.

Correo-e: caquero@ucm.es | ORCID iD: <http://orcid.org/0000-0002-3097-8734>

Recibido: 21-04-20; 2ª versión: 22-05-20; Aceptado: 22-05-20; Publicado: 27-04-2021

Cómo citar este artículo/Citation: Salvador-Oliván, J. A.; Marco-Cuenca, G.; Arquero-Avilés, R. (2021). Evaluación de la investigación con encuestas en artículos publicados en revistas del área de Biblioteconomía y Documentación. *Revista Española de Documentación Científica*, 44 (2), e295. <https://doi.org/10.3989/redc.2021.2.1774>

Resumen: Introducción y objetivo: La investigación con encuestas es un método utilizado con frecuencia en el área de Biblioteconomía y Documentación. El objetivo principal de este estudio es evaluar si los estudios realizados en el ámbito de la Biblioteconomía que utilizan encuestas como método de investigación proporcionan información completa y detallada sobre su diseño y ejecución.

Métodos: Se seleccionaron de la base de datos Web of Science los artículos publicados en 2019 en revistas del área de *Library Science* incluidas en el Journal Citation Reports que empleaban encuestas como método de investigación cuantitativa. Para valorar el grado de cobertura informativa se creó una herramienta formada por 32 elementos utilizados en diversas guías y recomendaciones.

Resultados y conclusiones: La mayoría de los artículos basados en encuestas en el área de Biblioteconomía y Documentación presentan una grave deficiencia en la información proporcionada. Es necesario mejorar y completar información sobre el procedimiento de muestreo, desarrollo y administración del cuestionario, así como sobre el análisis de sus resultados. Esta información permitirá valorar los potenciales errores cometidos y, en consecuencia, la calidad y validez de las conclusiones del estudio.

Palabras clave: investigación con encuestas; Biblioteconomía y Documentación; evaluación de la calidad de la investigación; metodología de investigación; técnicas y métodos de investigación

Evaluating survey research in articles published in Library Science journals

Abstract: Introduction and objective: Survey research is a method frequently used in Library Science. The main objective of this study is to evaluate if the studies carried out in the field of Library Science, that use surveys as a research method, provide complete and detailed information on their design and execution.

Methods: Scientific papers published in 2019 in journals in the Library Science area included in the Journal Citation Reports, that used surveys as a quantitative research method, were selected from the Web of Science database. To assess the degree of information coverage, a tool was created consisting of 32 items used in various guides and recommendations.

Results and conclusions: most of the scientific papers based on surveys in Library Science present a serious deficiency on the information provided. It is necessary to improve and complete information about the sampling procedure, development and administration of the questionnaire, as well as on the analysis of its results. This information will allow us to evaluate the potential errors made and, consequently, the quality and validity of the conclusions reached.

Keywords: survey research; Library Science; research quality assessment; research methodology; research techniques and methods

Copyright: © 2021 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1. INTRODUCCIÓN

La encuesta se ha definido como un método sistemático de obtención de datos mediante respuestas a preguntas realizadas a un grupo de entidades con el propósito de describir las características de la población a la que pertenecen (Groves y otros, 2009). Este método de investigación ha sido ampliamente utilizado en todos los campos de conocimiento, desde las Ciencias de la Salud a las Ciencias Sociales incluyendo la Biblioteconomía y Documentación.

Realizar una investigación apropiada con encuestas implica una "ciencia" que requiere la misma planificación y estructura que cualquier otro estudio que emplee el método cuantitativo, donde hay que definir principalmente (Morgan y Carcioppolo, 2014): a) la población a estudiar, número de individuos a incluir y cómo se van a seleccionar; b) qué datos se van a recoger (diseño del cuestionario); c) cómo se va a distribuir el cuestionario (método de administración); y d) cuándo iniciar la encuesta y hacer un seguimiento de los que no responden.

La replicabilidad de un estudio es una piedra angular del método científico, y para garantizar que sea posible, hay que proporcionar suficiente transparencia en la información sobre el trabajo realizado (McNutt, 2014). Es obligación del investigador describir los métodos y hallazgos de manera precisa y detallada en todas las publicaciones derivadas de la investigación, cumpliendo con los estándares para la divulgación (*American Association for Public Opinion Research* (AAPOR), 2015a).

La divulgación completa de los métodos empleados y una completa transparencia en todos los niveles sigue siendo el incentivo más poderoso para promover datos de alta calidad (Logan y otros, 2020) y requisito esencial para valorar la calidad de una encuesta. Esta calidad está relacionada con los posibles errores que se pueden cometer en cada fase del proceso de investigación (Jedinger y otros, 2018). Como método de investigación, una encuesta está sujeta a distintos tipos de errores, y cada uno de ellos reduce la posibilidad de llegar a conclusiones correctas a partir de los datos recogidos.

Estos errores se pueden clasificar en dos grandes grupos: los debidos a la observación o de medición y los errores debidos a la no observación. Los tres tipos de errores debidos a la no observación son: (1) error de cobertura, (2) error de no respuesta, y (3) error de muestreo. Y las potenciales fuentes de errores de observación: (1) el entrevistador, (2) el encuestado, (3) el cuestionario, y (4) el modo de recolección de datos (Groves, 1987; Weisberg,

2018). Los errores de observación y de no observación conforman el error total que se puede cometer en una encuesta (Groves y Lyberg, 2010).

Sin información clara y completa sobre los procedimientos en los que se basan las encuestas, es difícil poder evaluar los posibles errores y la calidad de la investigación, de ahí que se haya destacado la importancia de que los investigadores proporcionen información completa y exhaustiva sobre el diseño de la encuesta, así como sobre el proceso y análisis estadístico de los datos recogidos. La calidad de los datos de una encuesta sigue siendo una preocupación primordial y la calidad de la metodología empleada en estudios de encuestas es esencial tanto para asegurar la fiabilidad y validez de los resultados como para generalizar los resultados a las poblaciones de las que proceden las muestras analizadas.

Cada vez existe un mayor reconocimiento de la importancia de ser transparentes y proporcionar la información necesaria para permitir que la comunidad científica pueda evaluar de manera objetiva e imparcial la validez y utilidad de conclusiones de un estudio. La falta de información compromete tanto la transparencia como la fiabilidad de la investigación (Bennett y otros, 2011). Asimismo, la falta de transparencia obstaculiza el progreso científico y puede llevar a conclusiones erróneas sobre las implicaciones de los resultados de una investigación. Para evitar estos problemas, se han desarrollado diferentes iniciativas con el propósito de mejorar la calidad de la información de los estudios.

La necesidad de divulgar los métodos empleados en una encuesta ha sido promovida por la *American Association for Public Opinion Research* a través de su Iniciativa de Transparencia (AAPOR, 2015b), y de la publicación reciente de recomendaciones para diseñar y crear una encuesta lo mejor posible (AAPOR, 2020), complementado con una serie de preguntas que ayudan a evaluar la validez y calidad de los resultados de una encuesta (Baker y otros, 2016). Otras recomendaciones de interés sobre mejores prácticas en investigación de encuestas son las publicadas en un informe de la *National Science Foundation* basado en dos conferencias celebradas (Krosnick, y otros, 2015).

También se han desarrollado guías con la finalidad de ayudar a los investigadores a que no omitan ningún detalle importante y garantizar una publicación mínimamente adecuada que permita la reproducibilidad (Wharton, 2017). Son herramientas que indican los criterios que los autores deben informar para que los lectores puedan evaluar de manera crítica la calidad de la investigación (Hui y otros, 2019). En la investigación con encuestas

destacan en el campo de la Medicina las publicadas por Bennett y otros (2011), Grimshaw (2014), Draugalis y otros (2008), además de la de Rybakov y otros (2020) en el área de Farmacia, Hui y otros (2019) en el campo de Sistemas de Información, y Eysenbach (2004) en encuestas web.

Existe mucha literatura sobre cómo diseñar y ejecutar con rigor una investigación con encuestas. Sin embargo, en la práctica, incluso la encuesta de mayor calidad no llega a cumplir con todas los requisitos establecidos en los manuales de diseño de encuestas y recopilación de datos (Dale, 2006). Mucha investigación con encuestas no se hace bien y, por ello, a menudo tiene una reputación empañada (Totten y otros, 1999).

No se han realizado muchos estudios que evalúen la calidad de los informes de investigación de encuestas, pero todos ellos tienen un denominador común: proporcionan insuficiente información para evaluar la calidad de las encuestas; así se ha constatado en el área de Farmacia (Rybakov y otros, 2020), en Radiología (Shankar y Maturen, 2019) y en el campo de la Medicina (Bennett y otros, 2011; Turk y otros, 2018).

La encuesta ha sido un método de investigación empleado habitualmente en el área de Biblioteconomía y Documentación (Powell, 1999; Starr, 2012; Hernon y Schwartz, 2000), siendo más frecuente en los artículos publicados en las revistas *Journal of the Medical Library Association* y *Bulletin of the Medical Library Association* (Gore y otros, 2009), habiéndose confirmado en la actualidad como el método de investigación dominante en Biblioteconomía y Documentación (Ullah y Ameen, 2018). Su empleo en el ámbito bibliotecario permite demostrar científicamente su valor y contribución y proporcionar datos sobre el impacto en sus usuarios (Bertot y Jaeger, 2008), optando en algunos casos de ellas por incorporar encuestas en sus páginas web de inicio para averiguar las preferencias y medir la satisfacción de sus usuarios (Hernon y Schwartz, 2000).

A pesar de su amplio uso, se desconoce en qué medida los académicos en Biblioteconomía y Documentación han seguido unas buenas prácticas en sus investigaciones con encuestas al no existir estudios que evalúen su calidad y rigor metodológico.

2. OBJETIVOS

El objetivo principal consiste en evaluar la información proporcionada en los estudios realizados en el área de la Biblioteconomía y Documentación (en inglés, *Library Science*, LS) que emplean exclusivamente encuestas como método de investi-

gación cuantitativa. De manera específica, se trata de evaluar la transparencia en la información que se proporciona sobre el diseño y ejecución de la investigación necesaria para garantizar su reproducibilidad, así como para contrastar la validez de sus conclusiones.

3. MÉTODO

Estudio transversal de artículos publicados sobre investigación con encuestas entre el 1 de enero y el 31 de diciembre de 2019 en revistas del área de Biblioteconomía y Documentación incluidas en el Journal Citation Reports, edición 2018.

Fuentes de información y estrategia de búsqueda:

Para identificar los artículos se realizó una búsqueda en la colección principal de Web of Science (WoS) con los términos "survey" OR "questionnaire" en el campo *Tema* (título, abstract, palabras clave de autor y Keyword Plus) en las revistas clasificadas en la categoría *Library Science* según los estudios de Huang y otros (2019) y Abrizah y otros (2014). Además, se añadieron las revistas *Reference Services Quarterly* y *Journal of the Australian Library and Information Association*, no incluidas en los listados de estos estudios, pero específicas del área LS. Todas las revistas (Anexo 2) se buscaron con el prefijo SO= y se combinaron con el operador OR.

La decisión de utilizar solo revistas del área de Biblioteconomía y Documentación y no todas las clasificadas en la categoría JCR "*Information Science and Library Science*", se debe a que este etiquetado tradicional que une la Biblioteconomía y Documentación al campo de las Ciencia de la Información ha sido una decisión problemática, ya que aunque están relacionados no son lo mismo y no deberían ir unidos bajo la misma etiqueta descriptiva. Para Huang y otros (2019), la categoría *Information Science and Library Science* en el JCR combina dos campos diferentes, *Management Information Systems* (MIS) y *Library and Information Science* (LIS) llegando a identificar hasta cuatro subcampos diferentes: *Library Science*, *Information Science*, *Scientometrics* y *Management Information Systems*.

La fecha de ejecución de la búsqueda fue el 20 de febrero de 2020, y se actualizó el 16 de marzo para comprobar si se habían introducido nuevos registros en la base de datos que cumplieran con los criterios de búsqueda.

Criterios de inclusión/exclusión: Se definió como criterio de inclusión que las publicaciones fuesen artículos originales y usaran cuestionarios como método de recogida de la colección de datos empleando metodología cuantitativa. Se aplicaron

los siguientes criterios de exclusión: artículos con métodos de entrevista semiestructurada, *focus groups* y otros métodos de investigación cualitativos que emplearan cuestionarios exclusivamente con preguntas abiertas. También se excluyeron artículos no disponibles a texto completo.

Para comprobar que cumplían los criterios de elegibilidad, los autores examinaron el título y resumen de los artículos y, cuando no había suficiente información para tomar una decisión fundamentada, el texto completo.

Extracción de datos: Se desarrolló, por consenso entre los tres autores, un instrumento de evaluación o lista de validación (*checklist*) basado en los empleados en la literatura científica. Se eligieron únicamente ítems que definen una investigación con encuesta y están relacionados específicamente con el cuestionario. En una primera fase, se seleccionaron 24 ítems procedentes de la guía SURGE (SURvey Reporting GuidelinE) (Grimshaw, 2014) y de las empleadas en los estudios de Bennet y otros (2011) y Rybakov y otros (2020). En una segunda fase, se revisaron las 10 preguntas relacionadas con las mejores prácticas para las encuestas de la American Association for Public Opinion Research (AAPOR, 2020) y se añadieron 5 ítems. En la tercera y última fase, se revisó la Checklist for Reporting Results of Internet E-Surveys-CHERRIES (Eysenbach, 2004) de la que se seleccionaron otros 3 ítems.

Se excluyeron de estas guías todos aquellos ítems presentes en el título, resumen e introducción relacionados con los objetivos o contexto del estudio, y los que hacen referencia a la presentación completa y adecuada de los resultados, ya que estos apartados son comunes a otros tipos de estudios científicos y se presupone que en los artículos evaluados por pares deberían estar siempre presentes. También se eliminaron otros ítems como la declaración de espónsor o financiación del estudio.

El instrumento quedó integrado finalmente por 32 ítems, codificados según se muestra en el Anexo 1 y agrupados en categorías que reflejan las etapas necesarias en el diseño y ejecución de una encuesta científica: a) población diana y muestra; b) desarrollo de un cuestionario válido y fiable; c) administración y distribución del cuestionario; y d) proceso de datos y análisis de resultados. Además, contiene ítems relacionados con aspectos éticos como consentimiento informado, anonimato y confidencialidad.

Los tres autores evaluaron de manera independiente el texto completo del conjunto final de estudios incluidos (n=181); dudas sobre la codificación se resolvieron por consenso.

Análisis estadístico de los datos: Los datos se recogieron y procesaron con el paquete estadístico SPSS v.22. Los resultados se resumieron con estadísticos descriptivos, presentando el número y porcentaje de artículos que cumplían con cada ítem.

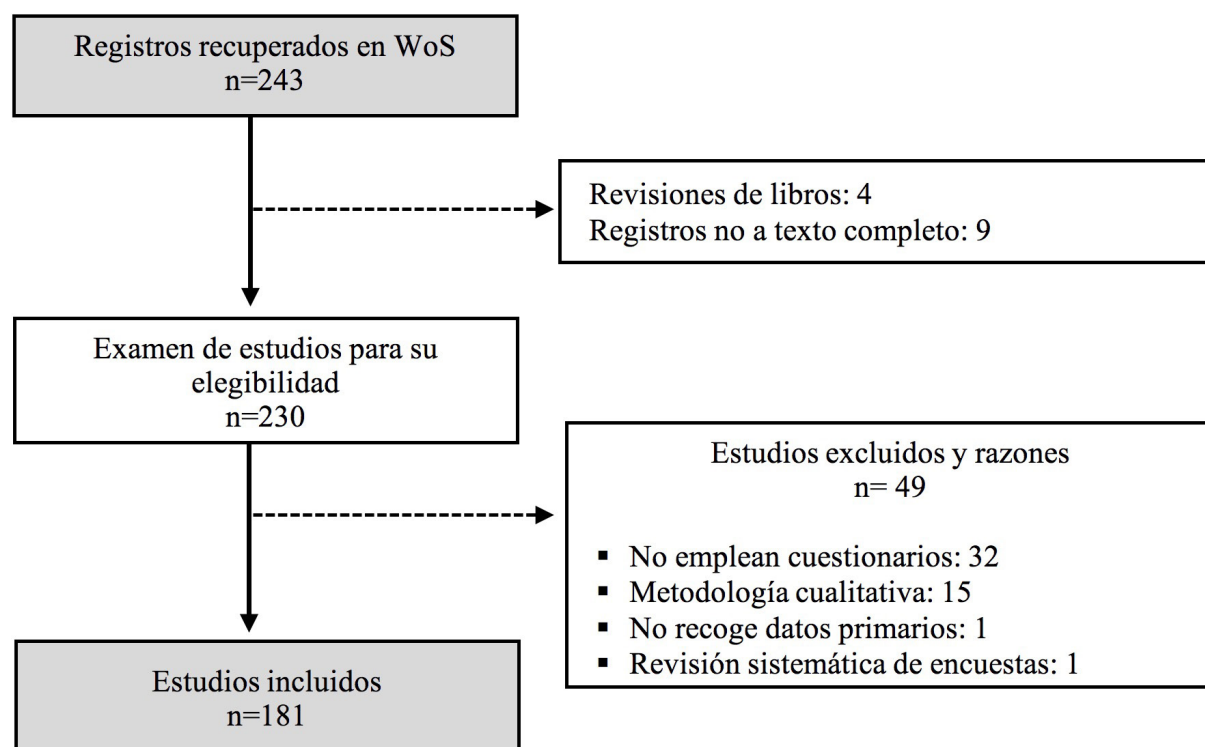
4. RESULTADOS

La estrategia de búsqueda recuperó 243 publicaciones. En un primer paso se excluyeron 4 que eran revisiones de libros y 9 artículos de los que no se pudo conseguir el texto completo. De los 230 artículos que quedaron, se excluyeron 49 que no cumplían los criterios de inclusión, bien porque no empleaban el cuestionario como técnica de recolección de datos (en estos artículos se empleaba el término de búsqueda "survey" entendido como estudio, examen, revisión o análisis) o utilizaban técnicas de metodología cualitativa (entrevistas, análisis de contenido en cuestionarios que contenían exclusivamente preguntas abiertas); se excluyó también un artículo que no empleaba datos primarios sino procedentes de encuestas nacionales, y otro que era una revisión sistemática de estudios de encuestas. Finalmente, quedaron 181 artículos que cumplían los criterios de inclusión (Figura 1).

La distribución del número de artículos evaluados en cada revista se muestra en el Anexo 2. De las 37 revistas que pertenecen a LS, 22 revistas contienen artículos evaluados, destacando *Journal of Librarianship and Information Science* con 45 artículos; 3 revistas han publicado artículos que emplean metodología de encuestas, pero se han excluido por no poder acceder al texto completo: *Libri*, *Serials Review* y *Journal of the Australian Library and Information Association*. En 2 revistas los artículos publicados no emplean el término "survey" entendido como encuesta, y en 12 revistas no se han publicado artículos sobre cuestionarios o encuestas en el año 2019.

En la tabla I se presenta el porcentaje de artículos que cumple con cada uno de los elementos listados en el instrumento de evaluación. De aquellos en los que se evalúa su presencia o ausencia, el mayor grado de cumplimiento se da en los siguientes: definición de la población diana y, en los artículos que utilizan un cuestionario existente, todos hacen referencia al trabajo original.

Cabe destacar que existe una deficiente información en muchos artículos. Así, más del 75% de los artículos carecen de información en 17 de los elementos evaluados. Aquellos que faltan en más del 95% de los artículos son los siguientes: no se describe el proceso de cálculo o se justifica el tamaño de la muestra, no se proporciona un libro de códigos, no se describe el procedimiento empleado

Figura 1. Diagrama de flujo de los estudios seleccionados

para verificar la calidad de los datos, no se indica el método para tratar los datos perdidos, o no se describe cómo se ha garantizado la confidencialidad o privacidad de los datos. En ninguno de los artículos se proporciona información sobre el método para el análisis del error de no respuesta, o cómo difieren los que han contestado de los que no lo han hecho.

En cuanto al tipo de muestreo utilizado, predomina el no probabilístico frente al probabilístico, si bien no consta en más de la mitad de los artículos. En 15 artículos (8,9%) se emplea una muestra probabilística y en 64 (35,3%) una muestra no probabilística, de los que 17 (9,4%) usan una muestra de conveniencia, en 13 (7,2%) utilizan una muestra de voluntarios, en 10 (5,5%) una muestra intencional, en 2 artículos emplean el muestreo de bola de nieve, en otros 2 el de cuotas, y en 1 combinan la muestra de conveniencia y de bola de nieve. En 1 artículo se utilizan dos poblaciones, seleccionando una muestra probabilística en una población y una muestra no probabilística en la otra población. En 19 artículos solo se indica que es no probabilístico.

En los artículos que emplean cuestionarios creados *de novo* o adaptados, solo en 36 (20,8%) consta que se ha evaluado la fiabilidad, la mayoría de

ellos (33) con el test de consistencia interna (alfa de Cronbach), en 1 artículo con el test-retest, y en 2 artículos se ha hecho tanto el test-retest como el test de consistencia interna. Solo 26 artículos han validado el cuestionario, en 14 de ellos la validez de constructo, en 6 la validez de contenido, en 2 artículos la validez aparente, y en 4 tanto la validez de contenido como la de constructo.

El modo de administración más frecuente es directamente a través de la web o de correo electrónico (invitación enviada a servidores de listas) con un enlace al cuestionario.

Aunque muchos artículos no declaraban explícitamente el método o programa empleado para introducir los datos; se contabilizaron todos aquellos que indicaban que el cuestionario se distribuía a través de Qualtrics o Google Forms.

De los 4 artículos que indicaban que se tenían en cuenta los datos perdidos en el análisis de datos, en uno de ellos se calculaba la tasa de respuesta en cada ítem y en otro se empleaba el método de regresión de mínimos cuadrados con 5 imputaciones; en los otros dos, solo se indicaba que se tenían en cuenta los datos perdidos en los porcentajes calculados.

Tabla I. Cumplimiento de los ítems de la checklist en la muestra total de artículos

		Si n (%)	No n (%)
Población y muestra			
1	Se define con precisión qué elementos forman parte de la población diana	153 (84,5%)	28 (15,5%)
2	Se describe el marco muestral	72 (39,8%)	109 (60,2%)
3	Método de selección de la muestra (tipo de muestreo) Probabilístico No probabilístico Probabilístico y no probabilístico No consta	15 (8,3%) 64 (35,3%) 1 (0,6%) 101 (55,8%)	
4	Se describe el proceso de cálculo o justificación del tamaño muestral (generalización de resultados o comparación de grupos)	9 (5,0%)	172 (95,0%)
Diseño y desarrollo del cuestionario			
5	Tipo de herramienta Nueva Adaptada Existente	156 (86,2%) 17 (9,4%) 8 (4,4%)	
6	Se presenta el cuestionario completo o las principales preguntas	61 (33,7%)	120 (66,3%)
7	Para herramientas existentes, se presentan las propiedades psicométricas	4 (50%)	4 (50%)
8	Para herramientas existentes, hay referencias al trabajo original	8 (100%)	
9	Para herramientas nuevas o adaptadas, se informa de su desarrollo o que se ha hecho prueba piloto o pre-test	39 (22,5%)	134 (77,5%)
10	Para herramientas nuevas o adaptadas, se ha valorado la fiabilidad	36 (20,8%)	137 (79,2%)
11	Para herramientas nuevas o adaptadas, se ha valorado la validez	26 (15,0%)	147 (85,0%)
12	Se proporciona un libro de códigos	2 (1,1%)	179 (98,9%)
Administración del cuestionario			
13	Modo de administración Autoadministrado por email/web Autoadministrado en persona en papel Entrevista personal (cara a cara) Entrevista asistida por ordenador Entrevista asistida por teléfono Modo mixto No declarado Autoadministrado por web/email y en papel Entrevista personal, asistida por ordenador y por teléfono	120 (66,3%) 15 (8,3%) 2 (1,1%) 1 (0,6%) 0 (0,0%) 2 (1,1%) 31 (17,1%) 9 (5,0%) 1 (0,6%)	
14	Se indica el lugar y/o fecha de administración	108 (59,7%)	73 (40,3%)
15	Información sobre método de contacto (email, carta) y nº de contactos (recordatorios) Tipo y número Solo tipo Ninguna información	23 (12,7%) 51 (28,2%) 107 (59,1%)	
16	Se proporciona información sobre el estudio, instrucciones u otro material a los encuestados	14 (7,7%)	167 (92,3%)
17	Información sobre si había o no incentivos para aumentar la tasa de respuesta	17 (9,4%)	164 (90,6%)
Proceso de datos y resultados			
18	Se indica el programa o herramienta empleada para introducir los datos	65 (35,9%)	116 (64,1%)
19	Se describe el procedimiento empleado para verificar la calidad de los datos	5 (2,8%)	176 (97,2%)
20	Se indica el método para tratar los datos perdidos	4 (2,2%)	177 (97,8%)
21	Se describe el método y tests estadísticos utilizados para el análisis de datos	90 (49,7%)	91 (50,3%)
22	Se indica la tasa de respuesta	83 (45,9%)	98 (54,1%)
23	Se proporciona el método para calcular la tasa de respuesta	28 (15,5%)	153 (84,5%)
24	Se proporciona el método para el análisis del error de no respuesta		181 (100%)
25	Hay información sobre cómo difieren los que han respondido de los que no		181 (100%)

		Si n (%)	No n (%)
26	Se informa del número de cuestionarios incompletos	24 (13,3%)	157 (86,7%)
27	Se exponen las limitaciones del estudio	94 (51,9%)	87 (48,1%)
28	Se discute la generalización de los resultados	53 (29,3%)	128 (70,7%)
Aspectos éticos			
29	Se ha aprobado por un comité de ética o de revisión institucional	20 (11,0%)	161 (89,0%)
30	Existe un procedimiento de consentimiento informado	14 (7,7%)	167 (92,3%)
31	Se informa de que el cuestionario es anónimo y/o confidencial	43 (23,8%)	138 (76,2%)
32	Se describe cómo se ha garantizado la confidencialidad/privacidad de los datos	9 (5,0%)	172 (95,0%)

Poco más de la mitad de los artículos describen las limitaciones del estudio, y la mayoría (70,7%) no incluye ninguna discusión sobre la generalización de los resultados. En cuanto a los indicadores relacionados con la ética, en todos ellos hay un muy deficiente registro siendo el que mayor porcentaje obtiene la información sobre el anonimato o confidencialidad de las encuestas.

5. DISCUSIÓN

Las encuestas son un método de investigación frecuente en las revistas especializadas en LS en el *Journal Citation Reports*; durante el año 2019, las dos terceras partes de las publicaciones del alcance de nuestro objeto de estudio han publicado artículos basados en investigaciones con encuestas. Sin embargo, en este estudio se han detectado deficiencias importantes en muchos de los aspectos evaluados y recomendados por la AAPOR y en las guías de información SURGE, CHERRIES y de los autores en las que se basa nuestro instrumento de evaluación.

En cualquier estudio, los investigadores deben cumplir con la norma científica de reproducibilidad, y ello exige que se detalle de manera completa y precisa el método empleado en una encuesta, lo que aumenta la credibilidad del estudio. Por el contrario, la falta de información implica desconocimiento sobre la ejecución de cada una de las etapas necesarias y no permite evaluar la precisión o sesgo de los resultados y, en consecuencia, la validez de sus conclusiones.

A continuación, se analizan y discuten los elementos de información evaluados en este estudio, haciendo hincapié en su utilidad para la replicabilidad y en su influencia a la hora de poder identificar y valorar los distintos tipos de errores que se pueden cometer en los estudios con encuestas, así como su repercusión en los resultados.

También se comparan nuestros resultados con los obtenidos por Bennet y otros (2011), Turk y otros (2018), y Rybakov y otros (2020), los únicos encontrados en la literatura que han aplicado instrumentos de evaluación similares, los dos primeros en el ámbito de la Medicina y el tercero en Farmacia. Hay que resaltar que en muchos ítems hay una amplia variación entre los resultados obtenidos en nuestro estudio y en los tres restantes, principalmente con los de Turk y Rybakov. Posibles razones que puedan justificar estas diferencias son que en el área de Medicina y de Farmacia hay publicadas guías específicas sobre la información que debe describirse en estudios realizados con encuestas; también puede ser debido a que la política de las revistas en estos campos de conocimiento puede ser diferente en el proceso de evaluación a la hora de exigir determinada información; y por último, el estudio de Bennet se realizó en 2011, lapso de tiempo suficientemente largo para que los autores hayan mejorado la información en sus estudios.

Población y muestra: En cualquier investigación mediante encuestas lo primero que hay que hacer es definir de manera clara la población objeto de estudio y, posteriormente, describir el marco muestral y el procedimiento de muestreo, si es probabilístico o no probabilístico. Además, hay que explicar cómo se ha calculado el tamaño de la muestra, especialmente importante cuando el propósito del estudio es describir las características de la población o contrastar diferencias en determinadas variables de interés entre subgrupos de la población. Todos estos datos pueden ayudar a determinar si la muestra es representativa o no, y si los resultados se pueden generalizar a toda la población.

Información sobre los ítems de este apartado permitirá conocer el error de cobertura y el error de muestreo. El error de cobertura ocurre cuando el marco muestral no incluye todos los elementos de la

Tabla II. Comparación entre diversos estudios de los resultados obtenidos en los ítems relacionados con la población, diseño y desarrollo del cuestionario

	Bennet (2011) n=117	Turk (2018) n=100	Rybakov (2020) n=105	Propio (2020) N=181
Población y muestra				
Se define con precisión la población diana	40,2%	88%	92%	84,5%
Se describe el marco muestral	57,2%	85%	97%	39,8%
Método de selección de la muestra (tipo de muestreo)				
Probabilístico	11%	43%	43%	8,3%
No probabilístico				35,3%
No consta				55,8%
Se describe el proceso de cálculo o justificación del tamaño muestral	6%	37%	53%	5,0%
Diseño y desarrollo del cuestionario				
Tipo de herramienta				
Nueva	94%		93%	95,6%
Existente	6%		7%	4,4%
Se presenta el cuestionario completo o las principales preguntas	35%	49%	26%	33,7%
En herramientas existentes, se presentan las propiedades psicométricas	23%	64%	14%	50%
En herramientas existentes, hay referencias al trabajo original	96%	73%	100%	100%
En herramientas nuevas, se informa de su desarrollo o pre-test	18%	42,8%	44%	22,5%
En herramientas nuevas, se ha valorado la fiabilidad	12,7%	8,5%	20%	20,8%
En herramientas nuevas, se ha valorado la validez	10%	25,7%	39%	15,0%
Se proporciona un libro de códigos				1,1%

población; si falta esta información, se desconoce si la muestra es representativa o no de la población. El error de muestreo o aleatorio es la discrepancia entre el verdadero valor del parámetro poblacional y el valor del estadístico en la muestra; ocurre siempre que se trabaja con muestras seleccionadas al azar y es un error que establece el investigador a priori cuando estima el número de individuos necesario para realizar el estudio; siempre que se emplean muestras probabilísticas se debería informar sobre el procedimiento de cálculo del tamaño muestral y el error de muestreo prefijado.

Nuestros resultados se aproximan más a los obtenidos por Bennet y son más deficientes que los de Turk y Rybakov. Revelan que en la mayoría de los estudios evaluados se describen las características sociodemográficas de los individuos que conforman la población diana, pero en más de la mitad de los estudios no se describe el marco muestral y en un 43,6% no consta la modalidad de muestreo empleada; solo en un 5% se describe el proceso de cálculo o se justifica el tamaño de la muestra, porcentaje menor que el de muestras probabilísticas, lo que evidencia que en varias de ellas no se indica el error aleatorio. Las deficiencias de información en estos ítems son frecuentes en encuestas web que habitualmente usan muestras no probabilísticas, que típicamente carecen de una definición de la población diana y de un marco

muestral (Jedinger y otros, 2018), salvo que se utilicen paneles de Internet.

En los otros estudios no se recoge si el tipo de muestreo es no probabilístico, tan solo si la muestra es representativa. Hemos asumido en tales casos que la muestra es probabilística.

Diseño y desarrollo del cuestionario: El cuestionario es el instrumento formado por las preguntas mediante las que se van a obtener los datos necesarios para llevar a cabo el estudio, por lo que todos los aspectos relacionados con su diseño y desarrollo son esenciales. Aunque se considera importante la evidencia que sustenta el contenido de un cuestionario, particularmente durante la fase de desarrollo, no se suele proporcionar tal información y lo que es peor, las respuestas y sus consecuencias en las encuestas son omitidas en la práctica de validación (Menold y otros, 2018).

No siempre es necesario desarrollar un instrumento nuevo. Existen instrumentos ya validados que pueden ser utilizados directamente, pero la gran mayoría de las ocasiones, como así se revela en la tabla II, se emplea un cuestionario nuevo o adaptado.

La descripción del contenido del cuestionario o su presentación completa como anexo al final de un estudio, junto con el esquema de codificación em-

pleado, facilita la evaluación y reproducción de un estudio, además de ayudar a otros investigadores a utilizarlo, adaptarlo u orientarles en el desarrollo de un nuevo cuestionario. A pesar de su importancia, los resultados de todos los estudios coinciden en que se incluye solo en menos de la mitad de los artículos.

El diseño del cuestionario es clave para el éxito de la investigación. Un cuestionario mal diseñado con preguntas que no reflejan con precisión el tema de interés, mal redactadas, ambiguas, no se entiende bien el significado o las opciones de respuestas no están claras, puede producir errores de medición. Este tipo de errores habitualmente permanecen ocultos en los datos y no se identifican salvo que se repita el proceso de medición o se comparen las respuestas con un patrón de referencia o *gold-estándar* (Biemer, 2009).

Si se ha realizado una prueba piloto o pre-test, hay que documentar en qué ha consistido y cómo ha afectado a posibles cambios en el cuestionario. La prueba piloto tiene gran valor al permitir identificar defectos en la preguntas, en definitiva, un mal diseño del cuestionario. Un cuestionario bien diseñado es más probable que sea completado por los encuestados, con lo que se reduce el riesgo de error de no respuesta y se mejora la tasa de respuesta.

Aunque no todas las encuestas requieren que se compruebe su validez y fiabilidad, sí es importante en aquellas diseñadas para medir o describir constructos (por ejemplo, satisfacción, calidad, etc.) y asegurar que las preguntas realmente miden aquello para lo que han sido diseñadas (validez interna) y que se obtienen los mismos resultados en los individuos en repetidas ocasiones o en muestras diferentes de la misma población (fiabilidad) (Heron y Schwartz, 2009). En estos casos se debería incluir una descripción completa de su desarrollo y test (cómo se han seleccionado los ítems, elección de las opciones de respuesta o escalas, propiedades psicométricas y prueba piloto), lo que contribuirá a mejorar la credibilidad de las encuestas.

Los resultados obtenidos en todos los estudios coinciden en que son áreas de mejora las relacionadas con la validez y fiabilidad del cuestionario, así como la descripción de los procedimientos utilizados en el pre-test, prueba que se ha realizado en casi la cuarta parte de los artículos. Un instrumento bien diseñado, válido, fiable y en el que se ha realizado una prueba piloto, ayuda a reducir el error de medición y el error de no respuesta.

Administración del cuestionario: Uno de los aspectos metodológicos más importantes antes de comenzar la colección de datos es el método de administración del cuestionario (Turk y otros,

2018). Consideramos que es una deficiencia grave no indicar algo tan sencillo como la forma de administración del cuestionario, y en el ámbito de nuestra investigación, no se declara en el 17,1% de los estudios.

Existen diferentes formas de administrar la encuesta, siendo el tipo más común en el área de Biblioteconomía y Documentación la encuesta auto-administrada debido a su economía y eficiencia (Janes, 2001). Así, el modo de administración más frecuente es a través de la web o por correo electrónico (66,3%), siendo el correo electrónico, con un enlace a la encuesta web, el medio de contacto con los individuos. Este tipo de encuestas llevadas a cabo en muestras no probabilísticas son las más coste-eficaces, ya que permiten alcanzar a un mayor número de participantes de manera simple, rápida y barata (Kalton, 2019), pero tienen tendencia a introducir un sesgo en la muestra (error de cobertura) ya que los encuestados suelen ser voluntarios o autoseleccionados y su opinión puede ser diferente de aquellos que rehúsan, no tienen interés o no tienen acceso a Internet (Bethlehem, 2010).

Información sobre cómo fueron invitados los encuestados y número de recordatorios puede ayudar a saber si se han hecho suficientes esfuerzos para maximizar la tasa de respuesta y, por tanto, si ha influido en la reducción del error de cobertura al aumentar el tamaño de la muestra, al mismo tiempo que puede alertar de un probable sesgo de no respuesta en los resultados de la encuesta. En el 60% de los artículos no se proporciona ninguna información sobre el tipo y/o número de contactos con los individuos, porcentaje algo más elevado que en el resto de los estudios.

Una breve introducción o protocolo de presentación escrito de manera profesional y explicando los objetivos, enfoque, utilidad y beneficios del estudio es también muy importante para animar a participar en la encuesta y, en consecuencia, aumentar la tasa de respuesta (Story y Tait, 2019). También es importante incluir información sobre instrucciones, que deberán ser claras y simples, tanto a entrevistadores como a encuestados, sobre todo cuando los cuestionarios son autoadministrados (Dillman y otros, 2007). Instrucciones poco claras, así como formación deficiente de los entrevistadores a la hora de formular las preguntas, pueden sesgar sistemáticamente las respuestas del encuestado y afectar al error de medición.

Siempre que se trabaja con muestras es importante conseguir unas tasas de respuesta altas, y la existencia de incentivos para responder puede ayudar a maximizar la tasa de respuesta y minimizar el error de no respuesta. En este caso, la

no respuesta se produce porque el individuo (elemento de la muestra) no es localizado o no tiene suficiente interés en contestar la encuesta, y ello provoca que los encuestados no representen a la población y no se puedan generalizar los resultados, pudiendo afectar seriamente a la validez externa de las conclusiones. El porcentaje de artículos que dan información sobre el uso de incentivos es muy bajo en todos los estudios.

Proceso de datos y resultados: La investigación realizada con encuestas debe cumplir con unos niveles mínimos de rigor científico, especialmente en lo concerniente a la transparencia de la tasa de respuesta y la representatividad y generalización de los resultados del estudio. Tres aspectos son fundamentales para conseguir una mejora en la calidad de los datos recogidos en una encuesta: la tasa de respuesta, la representatividad de la muestra y la tasa de preguntas completadas (Olson y otros, 2012).

La tasa de respuesta suele ser lo primero que los revisores y lectores buscan en el *abstract* de un artículo (Chung, 2014), por lo que siempre se debería proporcionar este dato para conocer el alcance o extensión del marco muestral encuestado. Debido a que hay varios métodos para calcular la tasa de respuesta, hay que informar siempre cómo se ha calculado. En nuestro estudio, la tasa de respuesta solo se indica en el 45,9% de los artículos evaluados, porcentaje algo más elevado que el 40% registrado en revistas de Sociología y Ciencia Política (Smith, 2002) pero más bajo que en los estudios contemplados en la tabla III; la amplia variación observada en los diferentes estudios sobre la tasa de respuesta y la descripción del método para calcularla puede ser debida a un reflejo histórico sobre la falta de consenso y la existencia de diferentes fórmulas para calcularla, y/o a que las revistas carecen de una política editorial explícita para exigir que se informe sobre ella y se basan en la experiencia de los evaluadores para determinar si esa información es necesaria (Johnson y Owens, 2004).

La tasa de respuesta puede afectar al error de no respuesta; cuanto más baja sea la tasa de respuesta, más alto será el sesgo de respuesta o el error de no-respuesta. Este puede producirse a nivel de unidad o a nivel de ítem. El error de no respuesta a nivel de unidad se produce cuando no se recogen datos de todos los individuos que conforman la muestra y cuando los que responden son diferentes de los que no lo hacen y, por tanto, no representan la opinión de toda la población. El error de no respuesta a nivel de ítem se produce cuando los individuos no contestan alguna(s) de

las preguntas. El aumento en errores de cobertura y de no respuesta, tanto a nivel individual como de ítem, puede deberse a la presencia de temas sensibles en las preguntas del cuestionario (Plutzer, 2019).

La tasa de respuesta es un indicador importante de la calidad de una encuesta, pero ha ido disminuyendo de manera constante a lo largo del tiempo conforme las encuestas se han enviado a través de la web o del correo electrónico. Conseguir una tasa alta de respuesta es necesario, pero no suficiente; es más importante que la muestra represente a la población que la tasa de respuesta absoluta; es decir, el objetivo es minimizar el error de no respuesta, aunque tasas altas de respuesta reducen el riesgo de error de no-respuesta (Marcopulos y otros, 2020).

No siempre es posible calcular la tasa de respuesta; en muestras no probabilísticas no se puede conocer ya que se desconoce el denominador (todas las unidades de la muestra). Siguiendo a Baker y otros (2016) y la norma ISO 20252:2019 (ISO, 2019), se recomienda llamar a esta tasa "tasa de participación", que se define como el número de encuestados que han contestado dividido por el número total de personas invitadas inicialmente a participar.

El análisis de no respuesta puede ser complicado de hacer por la falta de información de los que no responden, siendo más fácil de hacer cuando la muestra es finita y fácilmente accesible. Además de informar sobre el marco muestral, el método de selección de los participantes y la tasa de respuesta, los investigadores deberían incluir información sobre análisis de la no respuesta, sobre todo de la diferencia que existe en variables clave entre los que han respondido y los que no (Chung, 2014). Esta posible diferencia puede estimarse seleccionando una muestra de no encuestados – que no han contestado, (u obteniendo datos de ellos a través de fuentes secundarias) y comparar las características importantes con los que han contestado.

En toda investigación con encuestas se debe informar sobre el análisis de no respuesta (Werner y otros, 2007), y en ninguno de los estudios evaluados se hace referencia a ello ni a cómo difieren los que han contestado de los que no. Esta falta de información no permite valorar la representatividad de la muestra.

Cuestionarios incompletos se producen cuando los encuestados evitan responder a ciertas preguntas (por ejemplo, temas sensibles como el sexo o comportamientos ilegales), afectando en este caso

al error de no respuesta a nivel de ítem (Rao y Fuller, 2017). Es un dato íntimamente relacionado con el tratamiento de los datos perdidos, información que puede ayudar a los lectores a comprender mejor los datos presentados y su análisis. Los datos perdidos son de capital importancia en todas las encuestas, ya que ignorarlos puede llevar a sesgos importantes en las estimaciones de los resultados (Brick y Kalton, 1996). Debido a los diferentes métodos que existen para corregir los datos perdidos, es necesario informar siempre cómo se han tratado (Dale, 2006).

Áreas de mejora son informar y describir cómo se tratan los valores perdidos y el número de cuestionarios incompletos que, a pesar de ser bajo en nuestro estudio (13,3%), es mayor que en el resto de los estudios comparados.

Hay que describir también cómo se ha llevado a cabo la depuración y edición de datos, así como las medidas para garantizar su calidad antes de comenzar el análisis de los datos (Jedinger y otros, 2018). Hay que documentar cómo se identificaron inconsistencias o valores fuera de rango y cómo se corrigieron, por ejemplo, eliminando del conjunto de datos a los encuestados con dichos valores. Consideramos información importante describir cómo se ha comprobado la calidad de los datos y el programa empleado para la entrada de datos, dos elementos relacionados, ya que, si se introducen de manera automática, se debería describir cómo se ha garantizado la consistencia y validez de los datos.

Por último, resulta de gran utilidad tratar en el apartado de limitaciones del estudio todos los potenciales de error que se han podido cometer y cómo afectan a las conclusiones del estudio, así como la posible generalización de los resultados a la población.

Aspectos éticos y legales: Los códigos de ética establecen que debe respetarse siempre el derecho de los participantes a su privacidad (Nardi, 2006). El anonimato y la confidencialidad de los datos son aspectos que deberían respetarse siempre en cualquier estudio donde se recogen datos de individuos (American Association for Public Opinion Research (AAPOR), 2020). Los investigadores deberían informar siempre a los participantes de los procedimientos que se van a utilizar para asegurar la confidencialidad de la información, habiéndose constatado que la promesa a los encuestados de que sus respuestas son confidenciales aumenta la tasa de respuesta (Guterbock y Marcopulos, 2020).

Generalmente, la participación en las encuestas suele ser voluntaria, por lo que los individuos tie-

nen derecho a ser informados de los cuatro elementos que conforman el consentimiento informado: objetivos del estudio, cómo se va a garantizar la confidencialidad o anonimato, consentimiento activo o pasivo, y cualquier información relacionada con su participación voluntaria (Singer, 1993). Al mismo tiempo, los metodólogos de encuestas tienen la obligación ética y profesional de proteger la privacidad de los encuestados y de asegurar que se proporciona el consentimiento informado para su participación, aunque se ha señalado que este proceso puede contribuir al error total de la encuesta, concretamente debido a errores de representación y a errores de medición (Plutzer, 2019). En general, proporcionar información sobre cualquier aspecto relacionado con la ética contribuye a aumentar su transparencia y credibilidad (Oldendick, 2012).

El porcentaje de estudios en los que se informa de que se ha aprobado o revisado por un comité de ética institucional y hay un procedimiento de consentimiento informado es mucho más bajo en nuestro estudio. Esto puede ser debido a que en el ámbito de las Ciencias de la Salud los datos recogidos suelen ser más sensibles y hay mayor costumbre de cuidar los aspectos éticos, sobre todo los relacionados con la exigencia de consentimiento informado.

Fortalezas: Se ha utilizado un instrumento de evaluación empleado en otras áreas de conocimiento complementado con las recomendaciones de AAPOR, se ha seguido un método sistemático para la identificación e inclusión de los artículos evaluados, y se ha revisado por tres investigadores de manera independiente.

Limitaciones

El estudio es transversal y se ha realizado en artículos publicados en un solo año, por lo que sus resultados no pueden ser generalizados a toda la investigación realizada con encuestas en el área de Biblioteconomía y Documentación. Los resultados de este estudio se contemplan para todo el campo de conocimiento de Biblioteconomía y Documentación, pero algunas revistas están más representadas que otras; puede ser una situación puntual ocurrida en un año o puede ser debido a que habitualmente las revistas más representadas publiquen más investigaciones de este tipo.

En algunos ítems se ha evaluado la presencia o ausencia con cierta flexibilidad y subjetividad, por ejemplo, la definición de la población, que en algunos casos no se declaraba de manera explícita como población diana, pero sí se describían las características que definían al conjunto de individuos encues-

Tabla III. Comparación entre diversos estudios de los resultados obtenidos en los ítems relacionados con la administración y resultados del cuestionario

	Bennet (2011) n=117	Turk (2018) n=100	Rybakov (2020) n=105	Propio (2020) n=181
Administración del cuestionario				
Modo de administración				
Autoadministrado por email/web	57%	17%	51%	66,3%
Autoadministrado en papel	11%	43%	30%	8,3%
Modo mixto	12%	15%	9%	1,1%
No declarado	20%	19%	11%	17,1%
Se indica el lugar y/o fecha de administración				59,7%
Información sobre tipo de contacto y número de contactos				
Tipo y número	52%	18%	33%	12,7%
Solo tipo	13%	33%	28%	28,2%
Ninguna información	35%	49%	39%	59,1%
Se proporciona información sobre el objetivo del estudio, instrucciones u otro material a los encuestados				7,7%
Información sobre si había o no incentivos para aumentar la tasa de respuesta	23%	21%	30%	9,4%
Proceso de datos y resultados				
Se indica el programa o herramienta empleada para introducir los datos				35,9%
Se describe el procedimiento empleado para verificar la calidad de los datos				2,8%
Se indica el método para tratar los datos perdidos	11%	21%	19%	2,2%
Se describe el método y tests estadísticos utilizados para el análisis de datos	90%	94%	94%	49,7%
Se indica la tasa de respuesta	76%	60%	90%	45,9%
Se proporciona el método para calcular la tasa de respuesta	4%	32%	88%	15,5%
Se proporciona el método para el análisis del error de no respuesta	13%	13%	8%	0%
Hay información sobre cómo difieren los que han respondido de los que no	31%	19%	14%	0%
Se informa del número de cuestionarios incompletos	4%	11%	8%	13,3%
Se exponen las limitaciones del estudio	94%	75%	92%	51,9%
Se discute la generalización de los resultados	40%	35%	68%	29,3%
Ética				
Se ha aprobado por un comité de ética o de revisión institucional	59%	67%	89%	11,0%
Existe un procedimiento de consentimiento informado	25%	49%	25%	7,7%
Se informa de que el cuestionario es anónimo y/o confidencial				23,8%
Se describe cómo se ha garantizado la confidencialidad/privacidad de datos				5,0%

tados; o la tasa de respuesta, que no se indicaba la cifra de manera explícita, pero se podía obtener a partir de los datos proporcionados en los artículos.

6. CONCLUSIONES

La falta de información sobre los componentes fundamentales que definen una investigación con encuestas compromete la transparencia y la reproducibilidad, principios en los que se basa la investigación científica, e impide una evaluación rigurosa al ocultar posibles errores que afectarían a la validez de las conclusiones de los estudios.

La falta de información o documentación no significa que la calidad de los datos sea pobre, sino

que no es posible evaluar su calidad. Los autores deben especificar y describir de manera completa y detallada el método empleado para que cualquier investigador pueda replicar el estudio y los lectores, editores y revisores, evaluar su calidad. La investigación con encuestas es una potente herramienta para obtener información de poblaciones. Una buena planificación, diseño y ejecución del estudio debe llevar aparejado un registro de información suficiente para poder evaluar si los resultados son válidos y representativos de la población objeto de estudio.

Nuestro estudio revela falta de información en muchas de las investigaciones realizadas con encuestas en el área de Biblioteconomía y Documen-

tación. De manera general, se ha de poner más atención en informar con detalle qué datos se han recogido, cómo se han recogido, a quiénes, y qué resultados se han obtenido.

De manera más específica, es necesario mejorar de manera notable y proporcionar información completa y detallada en los siguientes apartados:

- Población y muestra: Descripción del marco muestral y método de selección de los individuos.
- Diseño y desarrollo del cuestionario: si se ha realizado prueba piloto, y presentar siempre el cuestionario completo con su esquema de codificación.
- Administración del cuestionario: tipo y número de contactos realizados, incentivos para aumentar la tasa de respuesta, e información sobre objetivos del estudio e instrucciones para cumplimentar el cuestionario.
- Proceso de datos y resultados: tasa de respuesta y cómo se ha calculado, método para analizar el error de no respuesta, cómo difieren los que han contestado de los que no, número de cuestionarios incompletos, programa informático utilizado para introducir los datos y cómo se ha comprobado la calidad de los datos. También

se deben de incluir siempre las limitaciones del estudio analizando todos los posibles errores y la generalización de los resultados.

- Ética: Informar siempre si los cuestionarios son anónimos y cómo se ha garantizado la confidencialidad de los datos.

Los investigadores deben aportar información completa de estas características de la encuesta. Ello mejorará la transparencia del estudio y resultará de utilidad a los lectores y evaluadores para conocer cómo se han minimizado las cuatro fuentes principales de error (error de cobertura, error de muestreo, error de no-respuesta y error de medición), valorar la calidad de las encuestas y determinar la validez de sus conclusiones.

El uso de guías existentes sobre mejores prácticas en estudios con metodología de encuestas puede ayudar a investigadores, editores y revisores a evaluar de manera crítica la información presentada en un artículo, y pueden servir de base para la creación y validación de nuevas *checklists* en distintos campos del conocimiento. Los editores de revistas deberían ofrecer a los autores recomendaciones de calidad más rigurosas sobre la información a incluir en estudios de encuestas.

7. BIBLIOGRAFÍA

- Abrihah, A.; Noorhidawati, A.; Zainab, A. N. (2014). LIS journals categorization in the Journal Citation Report: a stated preference study. *Scientometrics*, 102 (2), 1083–1099. <https://doi.org/10.1007/s11192-014-1492-3>
- American Association for Public Opinion Research (AAPOR). (2015a). The Code of Professional Ethics and Practices (Revised 11/30/2015). Disponible en: https://www.aapor.org/Standards-Ethics/AAPOR-Code-of-Ethics/AAPOR_Code_Accepted_Version_11302015.aspx [Fecha de consulta: 10/03/2020]
- American Association for Public Opinion Research (AAPOR). (2015b). Transparency Initiative Disclosure Elements. Disponible en: https://www.aapor.org/AAPOR_Main/media/transparency-initiative/Transparency_Initiative_Disclosure_Elements_050115.pdf [Fecha de consulta: 10/03/2020]
- American Association for Public Opinion Research (AAPOR). (2020). Best Practices for Survey Research. Disponible en: <https://www.aapor.org/Standards-Ethics/Best-Practices.aspx> [Fecha de consulta: 10/03/2020]
- Baker, R.; Brick, M.; Keeter, S.; Biemer, P.; Kennedy, C.; Kreuter, F.; Mercer, A.; Terhanian, G. (2016). AAPOR report: Evaluating Survey Quality in Today's Complex Environment. Disponible en: http://www.aapor.org/AAPOR_Main/media/MainSiteFiles/AAPOR_Reassessing_Survey_Methods_Report_Final.pdf [Fecha de consulta: 10/03/2020]
- Bennett, C.; Khangura, S.; Brehaut, J. C.; Graham, I. D.; Moher, D.; Potter, B. K.; Grimshaw, J. (2011). Reporting guidelines for survey research: An analysis of published guidance and reporting practices. *PLoS Medicine*, 8 (8), 1–11. <https://doi.org/10.1371/journal.pmed.1001069>
- Bertot, J. C.; Jaeger, P. T. (2008). Survey research and libraries: Not necessarily like in the textbooks. *Library Quarterly*, 78 (1), 99–105. <https://doi.org/10.1086/523911>
- Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, 78 (2), 161–188. <https://doi.org/10.1111/j.1751-5823.2010.00112.x>
- Biemer, P. (2009). Measurement errors in sample surveys. En: Rao, C. R. (ed). *Handbook of Statistics. Sample Surveys: Design, Methods and Applications. Volume 29, Part A*, pp. 281–315. Amsterdam: Elsevier. [https://doi.org/10.1016/S0169-7161\(08\)00012-6](https://doi.org/10.1016/S0169-7161(08)00012-6)
- Brick, J. M.; Kalton, G. (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, 5 (3), 215–238. <https://doi.org/10.1177/096228029600500302>
- Chung, K. C. (2014). Survey response rate, a guide for readers and authors. *Journal of Hand Surgery*, 39 (3), 421–422. <https://doi.org/10.1016/j.jhsa.2013.11.007>
- Dale, A. (2006). Quality issues with survey research. *International Journal of Social Research Methodology: Theory and Practice*, 9 (2), 143–158. <https://doi.org/10.1080/13645570600595330>
- Dillman, D. A.; Smyth, J. D.; Christian, L. M. (2007). *Internet, Phone, Mail and Mixed-Mode Surveys: The tailored design method (4th edition)*. New Jersey: John Wiley & Sons.

- Draugalis, J. L. R.; Coons, S. J.; Plaza, C. M. (2008). Best practices for survey research reports: A synopsis for authors and reviewers. *American Journal of Pharmaceutical Education*, 72 (1). <https://doi.org/10.5688/aj720111>
- Eysenbach, G. (2004). Improving the quality of web surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *Journal of Medical Internet Research*, 6 (3), 1–6. <https://doi.org/10.2196/jmir.6.3.e34>
- Gore, S. A.; Nordberg, J. M.; Palmer, L. A.; Piorun, M. E. (2009). Trends in health sciences library and information science research: An analysis of research publications in the Bulletin of the Medical Library Association and Journal of the Medical Library Association from 1991 to 2007. *Journal of the Medical Library Association*, 97 (3), 203–211. <https://doi.org/10.3163/1536-5050.97.3.009>
- Grimshaw, J. (2014). SURGE (The SURvey Reporting GuidelinE). En: Moher, D.; Altman, D. G.; Schulz, K. F.; Simera, I.; Wager, E.(eds.). *Guidelines for Reporting Health Research: A User's Manual*, pp. 206–213. New York: John Wiley & Sons. <https://doi.org/10.1002/9781118715598.ch20>
- Groves, R. M. (1987). Research on Survey Data Quality. *The Public Opinion Quarterly*, 51 (Part 2: Supplement: 50th Anniversary Issue), S156–S172. <https://doi.org/10.1086/269077>
- Groves, R. M.; Fowler, F. J.; Couper, M. P.; Lepkowski, J. M., Singer, E.; Roger, T. (2009). *Survey Methodology (2nd edition)*. Hoboken (New Jersey): John Wiley & Sons.
- Groves, R. M.; Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, 74 (5), 849–879. <https://doi.org/10.1093/poq/nfq065>
- Guterbock, T. M.; Marcpulos, B. A. (2020). Survey methods for neuropsychologists: A review of typical methodological pitfalls and suggested solutions. *Clinical Neuropsychologist*, 34 (1), 13–31. <https://doi.org/10.1080/13854046.2019.1590642>
- Hernon, P.; Schwartz, C. (2000). Survey research. A time for introspection. *Library and Information Science Research*, 22 (2), 117–121. [https://doi.org/10.1016/S0740-8188\(99\)00049-3](https://doi.org/10.1016/S0740-8188(99)00049-3)
- Hernon, P.; Schwartz, C. (2009). Reliability and validity. *Library and Information Science Research*, 31 (2), 73–74. <https://doi.org/10.1016/j.lisr.2009.03.001>
- Huang, M-H; Shaw, W. C.; Lin, C. S. (2019). One category, two communities: subfield differences in "Information Science and Library Science" in Journal Citation Reports. *Scientometrics*, 119 (2), 1059–1079. <https://doi.org/10.1007/s11192-019-03074-3>
- Hui, W.; Lui, S. M.; Lau, W. K. (2019). A reporting guideline for IS survey research. *Decision Support Systems*, 126 (May), 113136. <https://doi.org/10.1016/j.dss.2019.113136>
- ISO. (2019). ISO 20252:2019 Investigación de mercado, de opinión y social, incluidos los conocimientos y el análisis de datos - Vocabulario y requisitos de servicio. Madrid: Asociación Española de Normalización.
- Janes, J. (2001). Survey research design. *Library Hi Tech*, 19 (4), 419–421. <https://doi.org/10.1108/EUM000000006543>
- Jedinger, A.; Watteler, O.; Förster, A. (2018). Improving the quality of survey data documentation: A total survey error perspective. *Data*, 3(4), 1–10. <https://doi.org/10.3390/data3040045>
- Johnson, T.; Owens, L. (2004). Survey response rate reporting in the professional literature. En: *2003 Proceedings of the Section on Survey Research Methods*, pp. 127–133. Alexandria: American Statistical Association.
- Kalton, G. (2019). Developments in Survey Research over the Past 60 Years: A Personal Perspective. *International Statistical Review*, 87 (S1), S10–S30. <https://doi.org/10.1111/insr.12287>
- Krosnick, J. A; Presser, S.; Fealing, K. H.; Ruggles, S.; Vannette, D. L. (2015). *The Future of Survey Research: Challenges and Opportunities*. Arlington, VA: National Science Foundation.
- Logan, C.; Parás, P.; Robbins, M.; Zechmeister, E. J. (2020). Improving Data Quality in Face-to-Face Survey Research. *Political Science and Politics*, 53 (1), 46–50. <https://doi.org/10.1017/S1049096519001161>
- Marcopulos, B. A.; Guterbock, T. M.; Matusz, E. F. (2020). Survey research in neuropsychology: A systematic review. *Clinical Neuropsychologist*, 34 (1), 32–55. <https://doi.org/10.1080/13854046.2019.1590643>
- McNutt, M. (2014). Journals unite for reproducibility. *Science*, 346 (6210), 679. <https://doi.org/10.1126/science.aal724>
- Menold, N.; Bluemke, M.; Hubley, A. M. (2018). Validity: Challenges in Conception, Methods, and Interpretation in Survey Research. *Methodology*, 14 (4), 143–145. <https://doi.org/10.1027/1614-2241/a000159>
- Morgan, S. E.; Carcioppolo, N. (2014). Survey research methodology in health communication. En: Whaley B. B. (ed.). *Research Methods in Health Communication. Principles and Application*; pp. 78–96. New York: Routledge.
- Nardi, P. M. (2006). *Doing Survey Research. A Guide to Quantitative Methods (Second edition)*. Boston: Pearson Education.
- Oldendick, R. W. (2012). Survey Research Ethics. En: Gideon, L. (ed.). *Handbook of Survey Methodology for the Social Science*, pp. 23–35. New York: Springer. https://doi.org/10.1007/978-1-4614-3876-2_3
- Olsen, F.; Abelsen, B.; Olsen, J. A. (2012). Improving response rate and quality of survey data with a scratch lottery ticket incentive. *BMC Medical Research Methodology*, 12, 2–7. <https://doi.org/10.1186/1471-2288-12-52>
- Plutzer, E. (2019). Privacy, Sensitive Questions, and Informed Consent. *Public Opinion Quarterly*, 83 (S1), 169–184. <https://doi.org/10.1093/poq/nfz017>
- Powell, R. R. (1999). Recent trends in research: A methodological essay. *Library and Information Science Research*, 21 (1), 91–119. [https://doi.org/10.1016/S0740-8188\(99\)80007-3](https://doi.org/10.1016/S0740-8188(99)80007-3)
- Rao, J. N. K.; Fuller, W. A. (2017). Sample survey theory and methods: Past, present, and future directions. *Survey Methodology*, 43 (2), 145–160.
- Rybakov, K. N.; Beckett, R.; Dilley, I.; Sheehan, A. H. (2020). Reporting quality of survey research articles published in the pharmacy literature. *Research in Social and Administrative Pharmacy*. <https://doi.org/10.1016/j.sapharm.2020.01.005>

- Shankar, P. R.; Maturen, K. E. (2019). Survey Research Reporting in Radiology Publications: A Review of 2017 to 2018. *Journal of the American College of Radiology*, 16 (10), 1378–1384. <https://doi.org/10.1016/j.jacr.2019.07.012>
- Singer, E. (1993). Informed consent and survey response: A summary of the empirical literature. *Journal of Official Statistics*, 9 (2), 361–375.
- Smith, T. W. (2002). Reporting Survey Nonresponse in Academic Journals. *International Journal of Public Opinion Research*, 14 (4), 469–474. <https://doi.org/10.1093/ijpor/14.4.469>
- Starr, S. (2012). Survey research: We can do better. *Journal of the Medical Library Association*, 100 (1), 1–2. <https://doi.org/10.3163/1536-5050.100.1.001>
- Story, D. A.; Tait, A. R. (2019). Survey Research. *Anesthesiology*, 130 (2), 192–202. <https://doi.org/10.1097/ALN.0000000000002989>
- Totten, V. Y.; Panacek, E. A.; Price, D. (1999). Basics of research (Part 14) Survey Research Methodology: Designing the survey instrument. *Air Medical Journal*, 18 (1), 26–34. [https://doi.org/10.1016/S1067-991X\(99\)90006-8](https://doi.org/10.1016/S1067-991X(99)90006-8)
- Turk, T.; Elhady, M. T.; Rashed, S.; Abdelkhalek, M.; Nasef, S. A.; Khallaf, A. M; Huy, N. T. (2018). Quality of reporting web-based and non-web-based survey studies: What authors, reviewers and consumers should consider. *PLoS One*, 13 (6), 1–15. <https://doi.org/10.1371/journal.pone.0194239>
- Ullah, A.; Ameen, K. (2018). Account of methodologies and methods applied in LIS research: A systematic review. *Library and Information Science Research*, 40 (1), 53–60. <https://doi.org/10.1016/j.lisr.2018.03.002>
- Weisberg, H. F. (2018). Total Survey Error. En: Atkeson, L. R.; Michael, A. R. (eds.). *The Oxford Handbook of Polling and Survey Methods*; pp. 13–27. New York: Oxford University Press.
- Werner, S.; Praxedes, M.; Kim, H. G. (2007). The reporting of nonresponse analyses in survey research. *Organizational Research Methods*, 10 (2), 287–295. <https://doi.org/10.1177/1094428106292892>
- Wharton, T. (2017). Rigor, Transparency, and Reporting Social Science Research: Why Guidelines Don't Have to Kill Your Story. *Research on Social Work Practice*, 27 (4), 487–493. <https://doi.org/10.1177/1049731515622264>

ANEXO 1

Instrumento utilizado para evaluar la calidad de las encuestas

Población y muestra		
1	Se define con precisión qué elementos forman parte de la población diana	(Sí/No)
2	Se describe el marco muestral	(Sí/No)
3 ^a	Método de selección de la muestra (tipo de muestreo)	Probabilístico (1) No probabilístico (2) No consta (3)
4	Se describe el proceso de cálculo o justificación del tamaño muestral (generalización de resultados o comparación de grupos)	(Sí/No)
Diseño y desarrollo del cuestionario		
5	Tipo de herramienta	Nueva (1) Adaptada (2) Existente (3)
6	Descripción del cuestionario	Cuestionario completo o gran parte (1) No se proporciona el cuestionario (2)
7	Para herramientas existentes, se proporcionan las propiedades psicométricas	(Sí/No)
8	Para herramientas existentes, hay referencias al trabajo original	(Sí/No)
9	Herramientas nuevas y adaptadas, se informa de su desarrollo o de que se ha hecho prueba piloto o pre-test	(Sí/No)
10	Herramientas nuevas y adaptadas: Se ha valorado la fiabilidad	Test-retest (1) Intraobservador (2) Interobservador (3) Consistencia interna (4) No consta (0)
11	Herramientas nuevas y adaptadas: Se ha valorado la validez	Validez aparente (1) Validez de contenido (2) Validez de criterio (3) Validez de constructo (4) No consta (0)
12 ^a	Se proporciona un libro de códigos	(Sí/No)
Administración del cuestionario		
13	Modo de administración	Autoadministrado por mail (1) Autoadministrado por web (2) Autoadministrado en papel (3) Entrevista personal (cara a cara) (4) Asistido por ordenador (5) Asistido por teléfono (6) Modo mixto (7) No declarado (8)
14 ^a	Lugar y fecha de administración	(Sí/No)
15	Información sobre método de contacto (por email, carta, teléfono...) y número de contactos que se hicieron (seguimiento, recordatorios)	Tipo y número (2) Solo tipo (1) No información (0)
16 ^a	Se proporciona información sobre el estudio, instrucciones u otro material a encuestados	(Sí/No)
17	Información sobre si había o no incentivos para aumentar la tasa de respuesta	(Sí/No)

Proceso de datos y resultados		
18 ^b	Se describe el programa o herramienta empleada para introducir los datos	(Sí/No)
19 ^a	Se describe el procedimiento empleado para verificar la entrada y calidad de los datos	(Sí/No)
20	Se proporcionan los métodos para tratar los datos perdidos	(Sí/No)
21	Se describe el método y tests estadísticos utilizados para el análisis de datos	(Sí/No)
22	Se indica la tasa de respuesta	(Sí/No)
23	Se proporciona el método para calcular la tasa de respuesta	(Sí/No)
24	Se proporciona el método para el análisis del error de no respuesta	(Sí/No)
25	Se proporciona información sobre cómo difieren los que han respondido de los que no	(Sí/No)
26	Se informa del número de cuestionarios incompletos	(Sí/No)
27	Se exponen las limitaciones del estudio	(Sí/No)
28	Se discute la generalización de los resultados	(Sí/No)
Aspectos éticos		
29	Se ha aprobado por un comité de ética o de revisión institucional	(Sí/No)
30	Se describe que hay un procedimiento de consentimiento informado (Cómo se les informó sobre la duración de la encuesta, sobre qué datos se almacenaría, durante cuánto tiempo, y quiénes son los promotores y el objetivo del estudio)	Sí (1) Descripción completa (2) No declarado (0)
31 ^b	Se informa de que el cuestionario es anónimo y/o confidencial	(Sí/No)
32 ^b	Se describe cómo se ha garantizado la confidencialidad/privacidad de los datos	(Sí/No)

^a Ítems seleccionados de AAPOR (2020)

^b Ítems seleccionados de CHERRIES (Eysenbach, 2004)

Resto de ítems, seleccionados de SURGE (Grimshaw, 2014), Bennett et al. (2011) y Rybakov et al. (2020)

ANEXO 2.

Distribución de artículos evaluados en revistas de la categoría Library Science

Revista	Incluidos	Excluidos	Total
Journal of Librarianship and Information Science	45	9	54
Library Hi Tech	18	8	26
College & Research Libraries	14	3	17
Profesional de la Información	14	3	17
Journal of Documentation	12	4	16
Journal of the Medical Library Association	11	3	14
Aslib Journal of Information Management	11	1	12
Reference Services Review	7	3	10
Library & Information Science Research	7	1	8
Library Quarterly	6		6
Learned Publishing	5	2	7
Library Trends	5		5
African Journal of Library Archives and Information Science	4	1	5
Portal-Libraries and the Academy	4	1	5
Malaysian Journal of Library & Information Science	4		4
Journal of Scholarly Publishing	3	1	4
Library Resources & Technical Services	3	1	4
Revista Española de Documentación Científica	3	1	4
Law Library Journal	2		2
Library Journal	1	4	5
Information Technology and Libraries	1	1	2
Reference & User Services Quarterly	1		1
Libri		7	7
Serials Review		3	3
Journal of the Australian Library and Information Association		2	2
Library and Information Science		2	2
Information & Culture		1	1
Australian Library Journal			
Canadian Journal of Information and Library Science			
Econtent			
Electronic Library			
Health Information and Libraries Journal			
Interlending & Document Supply			
Library Collections Acquisitions & Technical Services			
Program-Electronic Library and Information Systems			
Restaurator-International Journal for the Preservation of Library and Archival Material			
Zeitschrift für Bibliothekswesen und Bibliographie			
TOTAL	181	62	243

ESTUDIOS / RESEARCH STUDIES

Curación de contenidos en periodismo. Indicadores y buenas prácticas

Javier Guallar*, Rafael Pedraza-Jiménez**, Mario Pérez-Montoro*, Laura Anton****

*Departamento de Biblioteconomía, Documentación y Comunicación Audiovisual de la Universidad de Barcelona. Centro de Investigación en Información, Comunicación y Cultura (CRICC).

Correo-e: jguallar@ub.edu ORCID iD: <https://orcid.org/0000-0002-8601-3990>

Correo-e: perez-montoro@ub.edu ORCID iD: <http://orcid.org/0000-0003-2426-8119>

**Departamento de Comunicación de la Universidad Pompeu Fabra. Grupo de investigación en Documentación Digital y Comunicación Interactiva (DigiDoc).

Correo-e: rafael.pedraza@upf.edu ORCID iD: <https://orcid.org/0000-0002-6918-6910>

****Barcelona Televisió (Betevé)

Correo-e: lauraansa@gmail.com ORCID iD: <https://orcid.org/0000-0001-8615-5932>

Recibido: 15-12-19; 2ª versión: 18-05-20; Aceptado: 23-05-20; Publicado: 27-04-2021

Cómo citar este artículo/Citation: Guallar, J.; Pedraza-Jiménez, R.; Pérez-Montoro, M.; Anton, L. (2021) Curación de contenidos en periodismo. Indicadores y buenas prácticas. *Revista Española de Documentación Científica*, 44 (2), e296. <https://doi.org/10.3989/redc.2021.2.1742>

Resumen: El artículo presenta un método o sistema para el análisis y la evaluación de la curación de contenidos en medios de comunicación digitales. Este sistema se basa en dos dimensiones (Curación y Contenido) que agrupan una serie de parámetros e indicadores a los cuales se les ha asignado una puntuación. Para cada parámetro se señalan los siguientes elementos: definición, explicación, procedimiento y ejemplos. El elemento "procedimiento" se defiende como una metodología para valorar el uso de la curación de contenidos en un sitio web periodístico, lo que incluye una relación de indicadores, y para cada uno de ellos una pregunta de análisis y una puntuación. El elemento "ejemplos" presenta un conjunto de buenas prácticas de curación de contenidos por parte de medios de comunicación digitales.

Palabras clave: curación de contenidos; curación periodística; curación de noticias; documentación periodística; periodismo digital; fuentes de información; sistemas de análisis; evaluación; indicadores.

Content curation in journalism. Indicators and good practices

Abstract. This article presents a method or system for the analysis and evaluation of content curation in digital media. The system is based on two dimensions (Content and Curation) that group a series of parameters and indicators to which a score has been assigned. The following elements are indicated for each parameter: definition, explanation, procedure and examples. The "procedure" element is defended as a methodology to assess the use of content curation on a journalistic website, which includes a list of indicators, and an analysis question and a score for each of them. The "examples" element presents a set of good practices for content curation in digital media.

Keywords: content curation; journalistic curation; news curation; news librarianship; digital journalism; information sources; analyses systems; evaluation; indicators.

Copyright: © 2021 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1. INTRODUCCIÓN

La curación de contenidos en periodismo se puede definir como: "Un complejo de actividades que incluye: 1) búsqueda y monitorización, 2) selección, 3) análisis y verificación, 4) gestión y edición, y 5) caracterización o *sense making* de informaciones publicadas en la web, con el objetivo de producir o mejorar productos periodísticos, lo que implica 6) la difusión de tales productos a través de plataformas digitales como sitios web de medios de comunicación, blogs y otros medios sociales" (Guallar y Codina, 2018). Las actividades anteriores guardan relación con algunas de las funciones que se han considerado tradicionalmente como propias de la documentación, si bien en el contexto actual del periodismo del siglo XXI, se encuadran cada vez más bajo la denominación de curación de contenidos, como señalan los autores anteriores.

A su vez, se debe señalar que el término "curación de contenidos", o "curaduría de contenidos", es de aparición reciente, situándose su origen en 2009 en el artículo del profesor de marketing Rohit Bhargava "Manifiesto for the content curator" (Bhargava, 2009). En dicho trabajo, el autor defendía, en una época caracterizada por el crecimiento exponencial de los contenidos digitales, la necesidad de una actividad profesional especializada en seleccionar el contenido más relevante para una audiencia determinada y en ofrecerlo de la manera más adecuada posible a esa audiencia, aportando un valor añadido.

Desde su aparición en el sector del marketing digital, ese concepto se expande hacia otros campos que tratan con información digital y de manera más destacada, a la biblioteconomía y documentación, el periodismo, la comunicación corporativa, la educación, y la ingeniería y las ciencias informáticas. Diez años después de su aparición, el concepto de curación de contenidos está plenamente presente en la literatura profesional y académica de diversas disciplinas y especialidades. Sin ánimo de exhaustividad, se pueden destacar en la bibliografía especializada, desde trabajos que abarcan la curación en un sentido amplio y transversal, como por ejemplo monografías con un enfoque profesional (Rosenbaum, 2011; Guallar y Leiva-Aguilera, 2013) o ensayístico (Bhaskar, 2016), a trabajos que abordan la aplicación de la curación de contenidos en diferentes sectores, como por ejemplo en bibliotecas (Parra-Valero, 2017; Martínez-Cañadas, 2017). En el ámbito concreto que nos ocupa en esta investigación, la curación de contenidos en periodismo, se pueden destacar entre otros, estudios como los de Bradshaw (2013), Bruns (2018),

Cappelletti Júnior y Domínguez Quintas (2014), Codina (2018), Cui y Liu (2017), Díaz Arias (2015), Guerrini (2013), Guallar (2017 a,b), Guallar y Codina (2018).

El enfoque diferencial de esta investigación respecto a otros trabajos sobre el tema es el de proponer un nuevo sistema para el análisis y la evaluación de la curación de contenidos en productos periodísticos, que se pueda aplicar y sea de utilidad en diversos contextos o escenarios, tanto a nivel académico como profesional.

2. OBJETIVO, MÉTODO Y PROCEDIMIENTO

2.1. Objetivo y justificación

El objetivo de este artículo es presentar un sistema de análisis basado en un conjunto de parámetros e indicadores que pueda ser útil para la evaluación de la calidad de los productos periodísticos basados en curación de contenidos.

Se justifica su necesidad porque no existe en la bibliografía especializada una propuesta como la que aquí se presenta para la evaluación y el análisis de la curación de contenidos periodísticos, ya que la mayoría de trabajos sobre la temática son descriptivos, teóricos o estudios de caso sobre medios o productos. Excepciones a ello, y como aproximaciones más cercanas a este trabajo, son la propuesta de taxonomía de Guallar (2017a) y la investigación sobre la curación en tres diarios norteamericanos utilizando tres indicadores de Cui y Liu (2017).

En este sentido, este artículo pretende ser una aportación original a la investigación sobre curación de contenidos, que pueda ser de utilidad en diversos escenarios (Rodríguez-Martínez y otros, 2012), como por ejemplo: 1) evaluación de la calidad de la curación de contenidos en sitios web periodísticos, es decir, que permita apreciar de qué manera se ha implementado la curación, con sus puntos fuertes o débiles; 2) análisis académicos de medios de comunicación, que permita establecer rankings, crear pautas o detectar tendencias; 3) ámbito profesional: que permita crear, mejorar o rediseñar productos o servicios de curación de contenidos a partir de los indicadores propuestos.

2.2. Metodología de investigación

Para conseguir los fines señalados, en primer lugar, se realizó una extensa revisión bibliográfica tanto sobre curación de contenidos en periodismo como sobre métodos de análisis y evaluación de sistemas y productos de información, en la que se constató la inexistencia de sistemas de

análisis sobre curación periodística con indicadores de evaluación. En segundo lugar, y a partir de dicha constatación, se decidió utilizar el análisis de sistemas y productos de información (Glaser y Strauss, 1967; Creswell, 2009; Ferran-Ferrer y otros, 2017; Guallar y otros, 2017) y la observación experta (Denzin y Lincoln, 2011) para la identificación y el análisis de una serie de elementos que permitan caracterizar, analizar y evaluar el objeto de estudio. A partir de este análisis, se ha concretado el nuevo sistema para la evaluación de curación de contenidos en medios periodísticos que presenta y describe este trabajo, el sistema CAS (Curation Analysis System).

La investigación sobre análisis de productos y servicios de información cuenta con varios precedentes, aplicados en diversos contextos y temáticas, como por ejemplo: a) sitios web en general (Codina, 2000; Pedraza-Jiménez y otros, 2016; Sanabre y otros, 2018); b) portales de archivos (Perpinyà-Morera y Cid-Leal, 2018); c) sitios web de comercio electrónico (Roig y Pedraza-Jiménez, 2016); d) cibermedios (Rodríguez-Martínez y otros, 2010, 2012); e) hemerotecas de prensa (Guallar y Abadal, 2009; Guallar y otros, 2012); f) bases de datos de prensa (Linares y otros, 2016); g) imágenes para prensa (Carbonell y Pedraza-Jiménez, 2016); h) webs de televisiones (Díaz Campo, 2014); i) archivos de televisiones en internet (Anton y Guallar, 2014); j) software de archivos audiovisuales (Cascón-Katchadourian y otros, 2018).

Nuestra propuesta metodológica ha tenido en cuenta todas las aportaciones señaladas, y de ellas ha prestado una atención especial a aquellos sistemas de análisis y evaluación basados en parámetros e indicadores, como el que se expone en detalle en Pedraza-Jiménez y otros (2016). En un sistema de análisis y evaluación de este tipo, los parámetros hacen referencia a los aspectos generales de los productos periodísticos susceptibles de estudio. Cada parámetro consta de un conjunto de indicadores, que son los aspectos concretos cuyo análisis nos permitirá emitir un juicio de valor sobre el parámetro en el que se integran.

En una metodología de este tipo, otra cuestión relevante a considerar es el sistema de puntuación que se asigna a los indicadores. Puede ser binario (0-1), cuando se evalúa la presencia o ausencia de una característica, o de puntuación múltiple (0-3), cuando se desea realizar una evaluación exhaustiva de las características estudiadas, valorándolas: mal (0); regular (1); bien (2); o muy bien (3). En nuestra propuesta se combinan ambas opciones, como se detallará más adelante.

Para obtener los indicadores que se presentarán a continuación se desarrollaron dos tareas. Por un lado, se realizó una observación de los productos de curación de contenidos en los principales medios digitales españoles. Para la selección de medios se consultaron los rankings de audiencia de la prensa digital española de ComScore (DIR Confidencial, 2019) y de OJD Interactiva (OJD Interactiva 2019), obteniendo una lista de siete medios tradicionales o *legacy media* (ABC, El Mundo, El País, El Periódico, La Razón, La Vanguardia y 20 Minutos) y siete nativos digitales (El Confidencial, El Diario.es, El Español, El Nacional, Nació Digital, OK Diario y Público). La elección de esta muestra de análisis se hizo con la intención de realizar la observación en un conjunto importante, y a la vez acotado, de medios periodísticos. En cuanto a la identificación de parámetros e indicadores de calidad, se realizó a partir de un análisis en profundidad de las características de las noticias de estos diarios, lo que permitió la identificación de 2 dimensiones de análisis (curación y contenido), 9 parámetros (o aspectos generales de evaluación), y 33 indicadores (que permiten valorar la calidad de los parámetros identificados). La monitorización de dichos cibermedios se realizó entre marzo y octubre de 2019.

Por otra parte, se analizaron en profundidad las propuestas existentes en la literatura especializada y que, con diferentes enfoques, han empleado diversos elementos para analizar el uso de la curación de contenidos. En concreto los trabajos que se han estudiado como principales referentes son los de Deshpande (2013), Barnhurst (2013), Cui y Liu (2017), Guallar (2017a), Martínez-Cañadas (2017), Orero y Cebrián-Enrique (2019). A partir de esos trabajos, se han adaptado, modificado y refinado algunos de los indicadores que ahí se utilizan, y se han propuesto otros nuevos, agrupándolos en nuestro sistema. En el apartado siguiente en el que se presenta el sistema de análisis se detallan, para cada parámetro, las referencias bibliográficas consideradas y las aportaciones realizadas.

Por último, con respecto a la bibliografía señalada acerca de métodos de análisis y evaluación de productos y servicios de información, el presente artículo se puede situar en línea, en cuanto a tipo de propuesta y de estructura, con trabajos anteriores, como los de Guallar y Abadal (2009), Rodríguez-Martínez y otros (2012), o Linares y otros (2016). Como en los casos citados, se presentará a continuación un sistema de análisis con ejemplos de buenas prácticas.

3. SISTEMA DE ANÁLISIS CAS Y EJEMPLOS DE BUENAS PRÁCTICAS

El sistema de análisis que presentamos, al que hemos denominado CAS (Curation Analysis System), se basa en parámetros e indicadores, que se han agrupado a su vez en las dimensiones *Contenido* y *Curación*, tal y como muestra la tabla 1.

Tabla 1. Sistema CAS (Curation Analysis System). Tabla resumen de dimensiones, parámetros e indicadores

A. Contenido	
A1. Número de contenidos	A1.1. Cantidad
A2. Rango temporal	A2.1. Información retrospectiva o atemporal A2.2. Información reciente A2.3. Información actual A2.4. Información en tiempo real
A3. Procedencia	A3.1. Contenido propio A3.2. Contenido externo
A4. Fuente según tipo de organización	A4.1. Fuentes oficiales A4.2. Fuentes corporativas A4.3. Medios de comunicación A4.4. Ciudadanos A4.5. Comunidades online
A5. Fuente según morfología	A5.1. Sitios web A5.2. Blogs A5.3. Redes sociales A5.4. Fuentes secundarias
B. Curación	
B1. Autoría	B1.1. Autoría
B2. Técnica de caracterización	B2.1. Retitular B2.2. Resumir B2.3. Comentar B2.4. Citar B2.5. <i>Storyboarding</i> B2.6. Paralelizar
B3. Integración	B3.1. Hipervínculo en el texto B3.2. Hipervínculo en listas o etiquetas B3.3. Contenido incrustado
B4. Función periodística	B4.1. Sin modificar B4.2. Describir B4.3. Contextualizar B4.4. Interpretar B4.5. Citar fuente B4.6. Citar autor B4.7. Llamar a la acción

A continuación, se detallará el sistema. A tal fin, para cada parámetro se facilitan las siguientes informaciones:

- Definición: breve definición del parámetro
- Explicación: descripción detallada, así como indicaciones y discusión acerca de su origen o presencia en la literatura.
- Procedimiento: metodología a seguir para valorar su uso en un sitio web periodístico, que

incluye la relación de indicadores, y para cada uno de ellos la pregunta de análisis y la puntuación.

- Ejemplos: se muestran ejemplos reales de buenas prácticas por parte de medios de comunicación digitales

Por último, conviene hacer una precisión en cuanto al uso de este sistema. Este nuevo sistema de análisis pretende ser exhaustivo, es decir, pretende cubrir todos los aspectos que afectan a la calidad de la curación en un medio periodístico y en él se propone un método de puntuación que no prioriza ninguno de los aspectos de calidad identificados. No obstante, quedará a criterio del analista o evaluador que lo emplee, su utilización total o parcial (si sólo está interesado en una parte de los parámetros identificados), así como la modificación de las puntuaciones propuestas si desea dar más importancia a unos indicadores que a otros. Por ejemplo, en el parámetro de Rango temporal, que hace referencia, como se verá, al marco temporal de los contenidos curados; un evaluador podría dar más importancia al contenido retrospectivo que al contenido en tiempo real y puntuarlos de manera diferente, al considerar que solo los mejores medios utilizan sistemas de información retrospectiva.

A. Dimensión Contenido

A1. Parámetro: Número de contenidos curados

Definición:

Cantidad de contenidos curados existentes en un producto periodístico.

Explicación:

En este trabajo denominamos productos periodísticos a aquellas unidades de análisis que se pretendan estudiar con esta metodología en los medios de comunicación; por ejemplo, pueden ser las noticias de la página de inicio, las noticias de una determinada sección, las *newsletters*, etc. Este parámetro se debe aplicar siempre en el contexto de comparar productos homogéneos entre sí; por ejemplo, si queremos analizar las *newsletters* de varios medios, las noticias de la página de inicio de varios medios, etc. Para identificar cuáles son los contenidos curados dentro de un determinado producto periodístico, se debe tener en cuenta si existe un enlace, ya sea en el texto, en una imagen o en forma de contenido incrustado, hacia un contenido que es independiente y externo al que se está analizando. Si no existe ese enlace que da acceso a otro contenido original, aun cuando se le mencione, no se puede considerar que existe contenido curado.

Una de las distinciones más claras que se pueden establecer entre unos tipos de productos periodísticos de curación de otros es la cantidad de contenidos curados que ofrecen. En este sentido, la cantidad se puede considerar un indicador de calidad. Este parámetro estaba presente en la taxonomía de Guallar (2017a), como distinción entre los productos que curan un solo contenido de los que curan varios, y sin puntuación. En nuestro sistema de evaluación se le asigna una puntuación en varios niveles al considerar la cantidad de contenidos curados disponibles dentro de un mismo producto como un criterio de calidad del mismo.

Procedimiento:

En este parámetro solo se considera un indicador, del mismo nombre.

- A1.1. Cantidad de contenidos curados.

Pregunta de análisis: ¿Cuántos contenidos curados tiene el producto analizado?

Puntuación: 0-inexistente; 1-tercil inferior; 2-tercil medio; 3-tercil superior

Se debe poner en primer lugar en contexto el producto analizado con otros del mismo tipo: por ejemplo, la página de inicio de un medio digital, una determinada sección del medio, una *newsletter*, etc. A partir de ahí se establece una puntuación en tres niveles (tres terciles), considerando el número más alto de fuentes curadas que se pueda tomar como referencia de buenas prácticas como el nivel máximo y dividiéndolo en tres partes: obtienen 3 puntos los productos con un número de contenidos situado en el tercil superior, 2 para el medio, uno para el bajo.

Ejemplos:

- "Mientras Dormías", *newsletter* de *El Español*, 27 junio 2019.

A1: 3 puntos

- "Las 7 y media", *newsletter* de *ABC*, 27 junio 2019

A1: 1 punto

Para una *newsletter* periodística se podría establecer la siguiente propuesta a partir de la observación de la prensa digital: para un rango de 1-10 contenidos curados por *newsletter*, 1 punto; para 11-20 contenidos 2 puntos; para más de 21 contenidos, 3 puntos.

Aplicando este sistema de puntuación a la *newsletter* "Mientras dormías" del diario *El Español*, que el día 27 de junio de 2019 presentaba 30 contenidos curados, se evaluaría asignándole 3 puntos en

este indicador. Siguiendo el mismo procedimiento, "Las 7 y media" del diario *ABC* recibiría 1 punto, ya que en su edición de ese mismo día solo aparecían 8 contenidos curados.

A2. Parámetro: Rango temporal del contenido curado

Definición:

Rango temporal en el que se enmarca el contenido curado atendiendo a la fecha en la que fue publicado originalmente.

Explicación:

Se distinguen cuatro rangos temporales, a partir de las investigaciones previas de Guallar, 2017a y b: información retrospectiva o atemporal (de meses o de años anteriores); información reciente (de los últimos días o semanas), información actual (de las últimas horas), e información en tiempo real.

En el caso de la *curación periodística* es fácil y útil diferenciarla en cada uno de estos rangos temporales: el primero de ellos (información retrospectiva) está muy estrechamente relacionado con la práctica tradicional de la documentación periodística, mientras que el resto ha crecido exponencialmente en los últimos años a medida que los medios sociales y la información en tiempo real han ido ganando protagonismo en los hábitos informativos de las audiencias.

Procedimiento:

Se consideran cuatro indicadores para cada uno de los rangos temporales establecidos.

- A2.1. Información retrospectiva o atemporal.

Pregunta de análisis: ¿El producto tiene contenidos curados retrospectivos, de meses o años anteriores, o son contenidos atemporales?

Puntuación: 0-No; 1-Sí

- A2.2. Información reciente.

Pregunta de análisis: ¿El producto tiene contenidos curados recientes, de los últimos días o semanas?

Puntuación: 0-No; 1-Sí

- A2.3. Información actual.

Pregunta de análisis: ¿El producto tiene contenidos curados actuales, de las últimas horas?

Puntuación: 0-No; 1-Sí

- A2.4. Información en tiempo real.

Pregunta de análisis: ¿El producto tiene contenidos curados en tiempo real?

Puntuación: 0-No; 1-Sí

Los indicadores resultantes de la aplicación de este parámetro no son excluyentes, sino compatibles entre sí y será un elemento de calidad de un producto periodístico su capacidad de ofrecer el máximo posible de los mismos. En ese caso, el mínimo de puntuación que puede obtener un producto de curación en este parámetro es 1 y el máximo, 4.

Ejemplos:

- "Planes para el Valle de los Caídos después de Franco" (Natalia Junquera, El País, 12 octubre 2019) <https://elpais.com/politica/2019/10/12/actualidad/1570901355_289484.html>

A2: 3 puntos

En esta noticia predomina la curación de contenidos publicados en meses y años anteriores, concretamente hasta 2011, lo que supone valorar con 1 punto la presencia de contenido retrospectivo. Además, se ha curado información de los últimos días, por lo que se valora con 1 punto la presencia de información reciente. Por último, también se ha curado contenido publicado durante las últimas 24 horas, por lo que se valora con 1 punto la presencia de contenido actual. De este modo, este parámetro obtiene un total de 3 puntos que, como acabamos de ver, procede de los indicadores: Información retrospectiva o atemporal, Información reciente, e Información actual.

- "Los independentistas calientan motores con una protesta en Sants" (Guillem Sànchez, El Periódico, 13 octubre 2019) <<https://www.elperiodico.com/es/politica/20191013/protesta-independentista-estacion-sants-7679413>>

A2: 3 puntos

Véase en este caso una noticia que se sigue en directo, que incluye la curación de las publicaciones en Twitter del autor de la información y de otros usuarios junto a contenidos publicados en las últimas horas y un contenido atemporal. En este parámetro se obtiene un total de 3 puntos, correspondientes a los indicadores: Información en tiempo real, Información actual, e Información retrospectiva o atemporal.

A3. Parámetro: Procedencia del contenido curado

Definición:

Procedencia del contenido curado desde el punto de vista del medio u organización que lo publica.

Explicación:

El contenido curado puede tener dos procedencias desde el punto de vista de la organización que produ-

ce el producto noticioso: puede ser contenido externo al propio medio, o bien contenido propio, es decir, publicado previamente por el propio medio digital. Tradicionalmente, la prensa digital ha privilegiado los enlaces hacia el contenido propio muy por encima de los enlaces al contenido externo y, si bien algunas investigaciones señalan una ligera corrección (Karls-son y otros, 2015), el contenido curado propio suele ser muy mayoritario en los enlaces de los productos periodísticos (Fondevila-Gascón y Segura-Jiménez, 2012; Orero y Cebrian-Enrique, 2019).

Procedimiento:

La valoración de este parámetro se realiza en función de dos indicadores, uno para cada una de las categorías establecidas.

- A3.1. *Contenido propio.*

Pregunta de análisis: ¿El producto tiene contenidos curados propios?

Puntuación: 0-No; 1-Sí

- A3.2. *Contenido externo.*

Pregunta de análisis: ¿El producto tiene contenidos curados externos?

Puntuación: 0-No; 1-Sí

Los indicadores de este parámetro no son excluyentes, sino compatibles entre sí y, considerando el contexto señalado antes, será un elemento de calidad para un producto periodístico ofrecer los dos tipos de contenido. El mínimo de puntuación que se puede obtener en este parámetro es 1 y el máximo, 2.

Ejemplo:

"Sánchez asume el peso del 10N: se multiplica en medios y mítines para activar al votante" (Juanma Romero, El Confidencial, 5 octubre 2019)

<https://www.elconfidencial.com/elecciones-generales/2019-10-05/pedro-sanchez-peso-campana-multiplica-medios-mitines-movilizacion_2270116/>

A3: 2 puntos

El artículo combina los contenidos curados publicados previamente por el propio medio con los de otras cabeceras, hasta un total de 8: de Antena 3 TV, La Razón, Aragón TV, PSOE, etc.). De esta forma, ofrece al lector una importante variedad de puntos de vista acerca de la noticia. En el parámetro se obtienen 2 puntos correspondientes a la evaluación positiva de los dos indicadores: contenido propio y externo.

A4. Parámetro: Fuente del contenido curado según tipo de organización

Definición:

Tipo de organización de la fuente productora del contenido curado

Explicación:

Existen diversas propuestas de clasificación para las fuentes de información utilizadas en periodismo, buena parte de ellas realizadas desde la especialidad de la documentación periodística y ahora también, de la curación periodística. A partir del análisis de, entre otros, los trabajos previos de Barnhurst, 2013; Cui y Liu 2017; Guallar y Codina 2018; Orero y Cebrian-Enrique, 2019, se presenta una propuesta que considera dos parámetros diferenciados para las fuentes, según la organización productora y según la estructura de la misma. En este apartado A4 de tipo de organización, se distingue entre fuentes oficiales (de la Administración pública), corporativas (de empresas u otras organizaciones privadas), medios de comunicación, ciudadanos particulares y comunidades online (wikis colaborativos, hashtags, foros online...).

Procedimiento:

Conforman este parámetro cinco indicadores, uno para cada una de las cinco categorías establecidas.

- A4.1. Contenido procedente de fuentes oficiales.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de fuentes oficiales, de las administraciones públicas?

Puntuación: 0-No; 1-Sí

- A4.2. Contenido procedente de fuentes corporativas.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de fuentes corporativas, de empresas u otras organizaciones privadas?

Puntuación: 0-No; 1-Sí

- A4.3. Contenido procedente de medios de comunicación.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de medios de comunicación?

Puntuación: 0-No; 1-Sí

- A4.4. Contenido procedente de ciudadanos particulares.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de fuentes personales, de ciudadanos particulares?

Puntuación: 0-No; 1-Sí

- A4.5. Contenido procedente de comunidades online.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de comunidades online (wikis colaborativos, hashtags, foros online...)?

Puntuación: 0-No; 1-Sí

Este parámetro, como en los dos casos anteriores, tiene indicadores no excluyentes y considera la variedad de tipos de fuentes en los contenidos curados como un criterio de calidad. Así, una puntuación de 1 indicará una mínima variedad de tipos de fuentes curadas por parte del producto analizado, mientras que 5 puntos, significaría una gran variedad de tipos de fuentes curadas.

Ejemplos:

-“¿Starbucks en Westeros? Las bromas sobre la aparición de un vaso en ‘Game of Thrones’” (Diana Rodríguez, El País, Verne, 7 mayo 2019) https://verne.elpais.com/verne/2019/05/06/mexico/1557171719_137538.html

A4: 3 puntos

El contenido curado en esta entrada toma en consideración publicaciones realizadas por usuarios en sus perfiles en redes sociales, junto con contenidos publicados por fuentes corporativas en redes sociales (Starbucks, Game of Thrones), así como noticias publicadas en el medio. En consecuencia, la fuente analizada obtiene en este parámetro 3 puntos, que corresponden a los indicadores: Ciudadanos particulares, Fuentes corporativas y Medios de comunicación.

-“El abogado que convirtió a miles de apátridas en ciudadanos con derechos en Kirguistán, Premio Nansen de ACNUR” (Eldiario.es, 2 octubre 2019)

<https://www.eldiario.es/desalambre/Kirguistan-Azizbek-Ashurov-Premio-ACNUR_0_947706060.html>

A4: 2 puntos

Esta curación se basa fundamentalmente en las fuentes de información oficiales de ACNUR, combinadas con noticias publicadas en medios. Se obtiene así 2 puntos, correspondientes a los indicadores de Fuentes Oficiales y de Medios de comunicación.

A5. Parámetro: Fuente del contenido curado según su morfología**Definición:**

Este parámetro evalúa la morfología o características de formato de la fuente del contenido curado.

Explicación:

A partir del análisis de la misma literatura que se ha comentado para el parámetro anterior A4, en este apartado A5 de tipo de organización de la fuente de contenido atendiendo a su morfología, se distingue entre: sitios web; blogs; redes sociales; y fuentes secundarias (como bases de datos, catálogos...).

Es importante señalar que estos dos parámetros A4 y A5 de fuentes del contenido curado según el tipo de organización y según su morfología, no son parámetros repetitivos, pues aunque ambos se refieren a fuentes de información, estas se han clasificado en base a criterios distintos. Así, por ejemplo, el contenido curado procedente de un blog (parámetro A5), podría ser de: una fuente oficial, una fuente corporativa, un medio de comunicación o un ciudadano particular (parámetro A4).

Procedimiento:

Para la valoración de este parámetro se considerarán cuatro indicadores, uno para cada una de las categorías establecidas.

- A5.1. Contenido procedente de sitios web.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de sitios web?

Puntuación: 0-No; 1-Sí

- A5.2. Contenido procedente de blogs.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de blogs?

Puntuación: 0-No; 1-Sí

- A5.3. Contenido procedente de plataformas sociales.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de plataformas sociales?

Puntuación: 0-No; 1-Sí

- A5.4. Contenido procedente de fuentes secundarias.

Pregunta de análisis: ¿El producto tiene contenidos curados procedentes de fuentes secundarias, como bases de datos, catálogos...?

Puntuación: 0-No; 1-Sí

Como en el parámetro anterior, aquí se valora la variedad de tipos de fuentes según su morfología en los contenidos curados como un criterio de calidad, y así, una puntuación total en este parámetro de 1 indicará la mínima variedad de fuentes, y el 4, la máxima variedad posible.

Ejemplos

- "En los barrios pobres, las frutas y plantas salvan vidas" (Roberto Palomo, El País, 14 noviembre 2019)

A5: 3 puntos

<https://elpais.com/elpais/2019/11/14/planeta_futuro/1573741258_140453.html>

En esta noticia se enlazan contenidos procedentes de sitios web, de blogs y de bases de datos. Por ello, se obtienen en este parámetro 3 puntos, que se corresponden con los indicadores de: contenido procedente de sitios web, de blogs, y de fuentes secundarias.

- "Polémica en Argentina por una foto de una rider con su bebé" (El Periódico, 6 octubre 2019)

<<https://www.elperiodico.com/es/sociedad/20191006/la-polemica-en-argentina-por-una-foto-de-una-rider-con-su-bebe-7669208>>

A5: 2 puntos

En esta curación destacan los contenidos publicados en dos plataformas sociales (Facebook y Twitter) que se complementan con un enlace a un sitio web. En consecuencia, se obtienen 2 puntos en este parámetro, que se refieren a los indicadores de: contenido procedente de plataformas sociales, y de contenido procedente de sitios web.

B. Dimensión curación

B1. Parámetro: Autoría de la curación

Definición:

Este parámetro analiza si existe una autoría declarada del producto de curación estudiado.

Explicación:

Un elemento clave y fundamental para la evaluación de la calidad de cualquier producto o servicio digital es la identificación clara e inequívoca de su autoría. Así ha quedado establecido en buena parte de la bibliografía especializada en métodos de análisis o evaluación, ya desde las primeras propuestas existentes, como la de Cordina (2000).

Procedimiento:

En este parámetro se considera un indicador de valoración de la autoría del producto periodístico.

- B1.1. Autoría de la curación.

Pregunta de análisis: ¿Es visible la autoría del producto analizado?

Puntuación: 0-No; 1-Sí

En el caso de un producto periodístico de curación de contenidos, la valoración de su autoría tiene consideraciones similares a la de cualquier otro contenido digital. Así, se considera un elemento de calidad si está disponible de manera clara y precisa su identificación para el lector.

Ejemplo

- "Mientras dormías, 31 de julio de 2019" (Paolo Fava, El Español, 31 julio 2019) <https://www.elespanol.com/newsletter/>

B1: 1 punto

Esta *newsletter* diaria de curación de contenidos explicita claramente su autoría, a cargo de Paolo Fava. Ello la distingue y la destaca sobre otros boletines periodísticos que son más bien de tipo automatizado y no tienen una autoría identificada. Por tanto, se obtiene un punto en este parámetro.

B2. Parámetro: Técnicas de caracterización de contenidos o *sense making*

Definición:

En este parámetro se valora la utilización visible en el producto curado de técnicas de *sense making* o caracterización de contenidos.

Explicación:

En la bibliografía especializada sobre curación de contenidos se enfatiza la importancia de lo que se suele denominar *sense making* o caracterización de contenido (Deshpande, 2013; Guallar y Leiva-Aguilera, 2013; Martínez-Cañadas, 2017); esto es, se trata de añadir valor al contenido curado que se ofrece a la audiencia, poniéndolo en contexto mediante una o varias de las técnicas descritas en la bibliografía. Deshpande (2013) identifica seis técnicas en una taxonomía que ha tenido una considerable influencia, siendo adaptada y revisada posteriormente. Las seis técnicas de Deshpande, de menor a mayor dificultad para el curador según este autor, son: extraer, retitular, resumir o comentar, citar, *storyboarding* y paralelizar. La primera de ellas es poco más que una extracción automatizada, por lo que se ha prescindido de la misma en este método de análisis al considerar que no aporta valor más allá de la propia selección del contenido a curar. En esta investigación sí se consideran todas las otras, añadiendo además por nuestra parte una diferenciación entre resumir y comentar:

retitular: (habitualmente, solo para cuando se trata de una única fuente) curar el contenido con un título diferente al original

resumir: ofrecer un texto de resumen informativo u objetivo acerca del contenido curado

comentar: presentar un texto de resumen en un tono personal, subjetivo o de opinión

citar: incluir una cita textual del contenido curado

storyboarding: unir en un mismo producto diversas piezas de contenido en formatos diversos (por ejemplo, textos de artículos, fotos, tuits, videos incrustados...) intercalando texto propio.

paralelizar: presentar dos o más contenidos que no parecen tener vinculación entre ellos pero que el curador relaciona, justificando y explicando dicha vinculación.

Estas técnicas pueden presentarse independientemente o combinadas en un mismo producto de curación. Por ejemplo, las dos últimas, que requieren de gran elaboración, normalmente comportan su uso junto con otras técnicas de las anteriores.

Procedimiento:

Para la evaluación de este parámetro se consideran seis indicadores, uno para cada una de las categorías descritas anteriormente.

-B2.1. Retitular.

Pregunta de análisis: ¿El producto analizado retitula el título original del contenido curado?

Puntuación: 0-No; 1-Sí

-B2.2. Resumir.

Pregunta de análisis: ¿El producto analizado contiene un texto de resumen informativo u objetivo del contenido curado?

Puntuación: 0-No; 1-Sí

-B2.3. Comentar.

Pregunta de análisis: ¿El producto analizado contiene un texto de resumen personal, subjetivo o de opinión acerca del contenido curado?

Puntuación: 0-No; 1-Sí

-B2.4. Citar.

Pregunta de análisis: ¿El producto analizado contiene una cita textual del contenido curado?

Puntuación: 0-No; 1-Sí

-B2.5. Storyboarding.

Pregunta de análisis: ¿El producto analizado une diversas piezas de contenido en formatos diversos?

Puntuación: 0-No; 1-Sí

- B2.6. Paralelizar.

Pregunta de análisis: ¿El producto analizado relaciona dos o más contenidos que no tenían antes vinculación entre ellos?

Puntuación: 0-No; 1-Sí

Este parámetro para la evaluación de la calidad de la curación puede hacer variar la puntuación del producto evaluado entre 0 puntos, si la curación es totalmente automatizada, y un máximo de 6 puntos, en el caso de combinar todas las técnicas presentadas. No obstante, una alta combinación de estas técnicas es muy poco frecuente.

Ejemplos:

- "El Despertador. 30 de mayo de 2019" (Paolo Faova, El Español, 30 mayo 2019)

<https://www.elespanol.com/newsletter/>

B2: 3 puntos

En esta *newsletter* se han aplicado tres técnicas de las descritas anteriormente: se ha realizado un resumen descriptivo de contenidos (Resumir), se ha mostrado la opinión del autor (Comentar) y se han añadido citas textuales (Citar). Se obtienen así 3 puntos en este parámetro, procedentes de los indicadores: Resumir, Comentar y Citar.

- "Cargas en Via Laietana y batalla campal en el aeropuerto de El Prat" (El Mundo, 14 octubre 2019) <<https://www.elmundo.es/cataluna/2019/10/14/5da407d321efa0a1248b468a.html>>

B2: 3 puntos

En el seguimiento de un evento de gran interés noticioso es habitual que los medios digitales publiquen noticias que se actualizan constantemente en directo, a medida que el evento evoluciona. En estos casos, la técnica del *storyboarding* es muy adecuada, ya que integra contenido de diversas fuentes y formatos de plataformas sociales, junto a la narración periodística de lo que va aconteciendo. En el caso del ejemplo referenciado, el formato externo más empleado han sido tweets, sobre todo de entidades y de políticos, así como videos y un gráfico interactivo del propio medio. Junto a ello, la descripción de la evolución de la jornada en un tono informativo encaja con la técnica de resumir y también se incluyen algunas declaraciones textuales. Se obtienen así 3 puntos, procedentes de los indicadores: *Storyboarding*, Resumir y Citar.

B3. Parámetro: Integración del contenido curado

Definición:

Este parámetro valora la forma en que se integra el contenido curado en el producto analizado

Explicación:

Nos fijamos exclusivamente en la forma del punto de acceso a cada contenido curado y no en la totalidad del producto periodístico de curación que se está analizando (artículo, página, sección, *newsletter*...). Se ha partido en este caso de la taxonomía de Guallar (2017a), y se ha modificado a partir de la observación de medios digitales realizada para esta investigación. En concreto, se distingue:

Hipervínculo dentro de un texto (el sistema más habitual en periodismo)

Hipervínculo en un apartado o sección de lista de enlaces o lista de etiquetas

Contenido incrustado (*embedded*): el contenido se integra dentro del producto periodístico y se puede consultar con las características propias de la publicación en la plataforma original (visionar, escuchar, etc.) sin salir del propio producto. Es el sistema habitual para visionar contenidos de algunas plataformas sociales como YouTube, Twitter...

En este parámetro es importante señalar que no se considera curación cuando en un determinado producto periodístico se hacen referencias a informaciones publicadas previamente a las cuales no se ofrece acceso. Es decir, si hay referencias a contenidos pero no existen hipervínculos o contenido incrustado, la puntuación sería 0. Asimismo, como en otros casos, se pueden presentar varias de estas posibilidades combinadas en un mismo producto de curación.

Procedimiento:

Para la valoración de este parámetro se utilizan tres indicadores, uno para cada una de las categorías establecidas.

- B3.1. Hipervínculo dentro del texto.

Pregunta de análisis: ¿El producto integra contenido curado mediante hipervínculo dentro del texto?

Puntuación: 0-No; 1-Sí

- B3.2. Hipervínculo en apartado de lista de enlaces o etiquetas.

Pregunta de análisis: ¿El producto integra contenido curado mediante hipervínculo en un apartado específico de lista de enlaces o de etiquetas?

Puntuación: 0-No; 1-Sí

-B3.3. Contenido incrustado (*embedded*).

Pregunta de análisis: ¿El producto integra contenido curado, incrustado o embebido manteniendo las funcionalidades de la plataforma original?

Puntuación: 0-No; 1-Sí

Los tres indicadores no son excluyentes y la puntuación del parámetro puede oscilar entre 1 y 3.

Ejemplos:

- "Twitter y Facebook eliminan 359 cuentas falsas atribuidas al PP para las últimas generales", (Borja Andrino, Daniele Grasso, Jordi Pérez Colomé, El País, 21 septiembre 2019) <https://elpais.com/tecnologia/2019/09/20/actualidad/1568972332_038053.html>

B3: 3 puntos

Este artículo contiene los tres tipos de integración de contenido curado que se han mencionado: 1) existen hipervínculos dentro del texto que remiten a otros artículos, 2) hay asimismo hipervínculos en las etiquetas (*tags*) que acompañan al artículo, y por último, 3) se muestran publicaciones en Twitter incrustadas dentro del cuerpo del artículo. Se obtiene así 3 puntos correspondientes a los 3 indicadores del parámetro.

- "Fernando Alonso sufre en su estreno en Marruecos tras un triple pinchazo y pierde 45 minutos con Al-Attiyah" (20 Minutos, 5 octubre 2019) <<https://www.20minutos.es/deportes/noticia/fernando-alonso-sufre-marruecos-pierde-al-attiyah-3790354/0/>>

B3: 2 puntos

El artículo inserta contenido procedente de Instagram junto a hipervínculos en el texto. Se obtiene así 2 puntos, que corresponden a los indicadores: hipervínculo dentro del texto, y contenido incrustado.

B4. Parámetro: Función periodística o informativa del contenido curado

Definición:

Este parámetro valora la función que cumple el contenido curado dentro del producto a analizar, desde el punto de vista periodístico o informativo.

Explicación:

Nos referimos en este apartado a la intencionalidad o propósito con el que cada contenido curado aparece dentro del producto periodístico analizado. A partir de Barnhurst (2013) y de Cui y Liu (2017), se pueden distinguir diferentes tratamientos o enfoques periodísticos en una curación periodística. De Cui y Liu se toman sus tres categorías principa-

les, *Sourcing curation*, *Contextualizing curation* e *Interpreting curation* y se añaden cuatro más, obtenidas a partir de la observación de medios digitales realizada para esta investigación. Como en el parámetro anterior, la función se analiza exclusivamente para cada hipervínculo o contenido incrustado, y no en el resto de contenido redactado por el curador y que completa el producto en cuestión. Se consideran las siguientes categorías:

Sin modificar: contenido original sin modificar por parte del curador. Se han identificado dos tipos: hipervínculos en los que el texto corresponde al título original del documento, sin modificación por parte de este; y los documentos incrustados (*embedded*) procedentes de plataformas sociales.

Describir: Contenido curado como descripción o resumen. Coincide con *sourcing curation*: "Los hipertextos de los artículos reflejan parcial o totalmente el contenido de enlazado. Sin clicar el contenido original el lector debe ser capaz de tener una idea general del mismo" (Cui y Liu, 2017, 855). Este tratamiento periodístico está relacionado directamente con la técnica de *sense making* de resumir.

Contextualizar: el contenido curado se utiliza fundamentalmente para contextualizar o documentar una información (*contextualizing curation*): sin que ese contenido esté relacionado directamente con la historia de la noticia, puede ayudar sin embargo a entender los contextos sociales, culturales, históricos... de la misma. Se apoya especialmente en la curación de fuentes de información atemporal o retrospectiva.

Interpretar: el contenido curado es interpretado por el curador (*interpreting curation*), es decir, los hipervínculos a las fuentes no son frases descriptivas o de contexto sino interpretaciones o evaluaciones del contenido ("el contenido curado sirve para apoyar o reafirmar una opinión subjetiva o interpretativa del autor del texto", Cui y Liu, 2017, 859). Este tratamiento periodístico está relacionado directamente con la técnica de *sense making* de comentar.

Citar Fuente: cuando el texto del hipervínculo no se refiere a aquello de lo que trata el contenido al que se accede, sino que el hipervínculo señala cual es la fuente donde se ha publicado.

Citar Autor: como en el caso anterior, pero señalando al autor.

Llamar a la acción: la popular técnica del *Call to Action* (CtA) del marketing digital se utiliza también en los medios de comunicación digitales. Consiste en una llamada a entrar en un hipervínculo o en un contenido incrustado, a través de un texto del tipo "clica aquí".

Procedimiento:

Para la evaluación de este parámetro se definen siete indicadores, uno para cada una de las categorías establecidas.

- B4.1. Sin modificar

Pregunta de análisis: ¿El producto analizado presenta acceso a contenido curado sin modificación en su título o directamente como contenido incrustado?

Puntuación: 0-No; 1-Sí

- B4.2. Describir.

Pregunta de análisis: ¿El producto analizado presenta acceso a contenido curado con un hipervínculo o incrustación de tipo descriptivo o de resumen?

Puntuación: 0-No; 1-Sí

- B4.3. Contextualizar.

Pregunta de análisis: ¿El producto analizado presenta acceso a contenido curado con un hipervínculo o incrustación del tipo de contextualización?

Puntuación: 0-No; 1-Sí

- B4.4. Interpretar.

Pregunta de análisis: ¿El producto analizado presenta acceso a contenido curado con un hipervínculo o incrustación de tipo interpretativo o de opinión?

Puntuación: 0-No; 1-Sí

- B4.5. Citar fuente.

Pregunta de análisis: ¿El producto analizado presenta acceso a contenido curado mediante indicación de la fuente?

Puntuación: 0-No; 1-Sí

- B4.6. Citar autor.

Pregunta de análisis: ¿El producto analizado presenta acceso a contenido curado mediante indicación del autor?

Puntuación: 0-No; 1-Sí

- B4.7. Llamar a la acción.

Pregunta de análisis: ¿El producto analizado presenta acceso a contenido curado mediante una llamada explícita a la acción?

Puntuación: 0-No; 1-Sí

Se trata de funciones no excluyentes y que se pueden complementar, aunque difícilmente se puede dar la combinación de todas ellas en un mismo

producto; el mínimo de puntuación del parámetro es 1 y el máximo, 7.

Ejemplos:

- "El Despertador, 9 de setembre: Deures per fer" (Ferran Cases, Nació Digital, 9 septiembre 2019)

<<https://www.naciodigital.cat/noticia/186984/deures/fer>>

B4: 5 puntos

Los artículos y *newsletters* de este autor muestran una variada combinación de tratamientos periodísticos en sus contenidos curados. En este ejemplo se puede apreciar cómo: 1) describe, 2) contextualiza ("Així va explicar Macià..."), 3) interpreta ("tal com demostra..."), 4) cita autor ("article conjunt d'Oriol Junqueras i Marta Rovira") y 5) llama a la acción ("Recupereu-lo aquí"). Se obtiene así en este parámetro un total de 5 puntos correspondientes a los indicadores de: Describir, Contextualizar, Interpretar, Citar autor y Llamar a la acción.

- "Así hemos contado las primeras reacciones a la sentencia del 'procés' en Cataluña" (El Confidencial, 15 de octubre de 2019) https://www.elconfidencial.com/espana/cataluna/2019-10-15/manifestaciones-el-prat-sentencia-proces-directo-cortes-vias-tren_2279696/

B4: 4 puntos

En esta pieza periodística, del tipo de seguimiento en directo de un acontecimiento, se presenta el enlace al contenido curado de las siguientes maneras: 1) sin modificar (el más habitual en los contenidos curados procedentes de plataformas sociales, como tweets), 2) mediante hipervínculos descriptivos, 3) mediante citas al autor del contenido ("Lo cuenta Juanma Romero"), y 4) con llamadas a la acción ("Lea aquí la noticia completa"). Se obtienen así en este parámetro un total de 4 puntos correspondientes a los indicadores: Sin modificar, Describir, Citar autor, y Llamar a la acción.

4. CONCLUSIONES

El sistema que aquí se presenta es la primera propuesta existente en la literatura académica para el análisis y la evaluación de la curación de contenidos en periodismo. Hasta el momento se habían descrito casos, analizado muestras de diarios o propuesto taxonomías, pero no existía una propuesta de un sistema de análisis y evaluación, que se pueda aplicar a todo tipo de diarios o de productos de noticias, y que permita apreciar qué nivel de calidad existe en la curación de contenidos de un determinado diario o un determinado pro-

ducto periodístico, ya sea un tipo de artículo, una sección, una *newsletter*... Este sistema de análisis va por tanto más allá de la simple descripción de una noticia o de un medio, pues permite, además de identificar sus componentes, establecer valoraciones de calidad, y realizar productos derivados, por ejemplo, ránquines, etc. En este sentido, nuestro sistema de análisis permite llevar a cabo al menos tres tareas distintas:

- realizar análisis comparativos entre diferentes productos noticiosos que utilicen la curación de contenidos;

- utilizar el sistema como herramienta para evaluar la calidad de la curación de contenidos por un medio concreto con el objetivo de identificar aquellos aspectos que son mejorables;

- utilizar el sistema como guía para el desarrollo de nuevos productos de calidad basados en la curación de contenidos.

En este sistema se ha potenciado la valoración de la variedad de usos de la curación, con el convencimiento de que una mayor variedad implica una mayor calidad de la misma. Así, por ejemplo, si observamos el parámetro A4, *Fuentes del contenido curado atendiendo al tipo de organización*, se considera que un producto periodístico que cure contenidos provenientes de varios tipos de fuentes informativas, tales como Organismos oficiales, Organizaciones privadas, Asociaciones, Ciudadanos y Medios de comunicación, tendrá una calidad superior a si solo ofrece contenidos de un solo tipo de fuentes de información, como por ejemplo Medios de comunicación (algo, esto último, muy habitual). En este caso, por tanto, cuanto mayor variedad de tipos de fuentes, mayor calidad en la curación. Así, en líneas generales consideramos aquí que la existencia de una mayor variedad de uso de cada uno de los elementos que conforman una curación, desde los tipos de fuentes a los tipos de técnicas de *sense making*, repercutirá en una mayor calidad de la curación y por extensión, en una mayor calidad del periodismo que se ofrece a la ciudadanía.

Como con cualquier sistema de análisis y evaluación, es importante señalar que éste debe es-

tar sujeto a pruebas y refinamientos sucesivos. Por nuestra parte, como futuras líneas de trabajo, en relación directa con este trabajo, se propone la aplicación de este sistema a muestras acotadas de medios digitales o de productos periodísticos; por ejemplo a: 1) medios digitales de un determinado país, principales medios de una área geográfica determinada o a nivel mundial; 2) productos periodísticos determinados, como *newsletters*, reportajes multimedia, redes sociales de medios...; o 3) contenidos informativos temáticos, como información política, deportiva, especializada en tecnología...

En suma, con esta propuesta y su aplicación sucesiva, se pretende ayudar a comprender mejor las potencialidades de la curación de contenidos en el periodismo, lo que puede revertir, tanto a nivel particular, en la mejora de su uso por parte de los medios digitales, como a nivel global, en la mejora de la calidad del periodismo, y tanto en escenarios profesionales como académicos y de investigación.

5. AGRADECIMIENTOS

Este estudio ha sido realizado en el marco del grupo de investigación consolidado "Cultura i Continguts Digitals" (SGR 2017-422), financiado por la Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) de la Generalitat de Catalunya y forma parte del proyecto "Narración interactiva y visibilidad digital en el documental interactivo y periodismo estructurado" RTI2018-095714-B-C21 (MICINN/FEDER), Ministerio de Ciencia, Innovación y Universidades (España).

ACKNOWLEDGEMENTS

This study has been carried out within the framework of the consolidated research group "Cultura i Continguts Digitals" (SGR 2017-422), financed by the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) of the Generalitat de Catalunya and is part of the project "Interactive narration and digital visibility in interactive documentary and structured journalism" RTI2018-095714-B-C21 (MICINN / FEDER), Ministry of Science, Innovation and Universities (Spain).

6. REFERENCIAS

- Anton, L.; Guallar, J. (2014). Análisis de los archivos audiovisuales en internet de las televisiones autonómicas españolas. *Revista Española de Documentación Científica*, 37 (1), e033. <https://doi.org/10.3989/redc.2014.1.1044>
- Barnhurst, KG. (2013). Newspapers experiment online: Story content after a decade on the web. *Journalism*, v. 14, 3-21. <https://doi.org/10.1177/1464884912448898>

- Bhargava, R. (2009). Manifesto For The Content Curator: The Next Big Social Media Job Of The Future?. *Rohit Bhargava.com*, 30 September. <http://www.rohitbhargava.com/2009/09/manifesto-for-the-content-curator-the-next-big-social-media-job-of-the-future-.html>

- Bhaskhar, M. (2016). *Curation. The power of selection in a world of excess*. [ebook] ISBN: 978-0349408699

- Bradshaw, P. (2013). Journalism *is* curation: tips on curation tools and techniques. *Online Journalism Blog*, 30 September <http://onlinejournalismblog.com/2013/09/30/curation-tools-tips-advice-journalism>
- Bruns, A. (2018). *Gatewatching and news curation: Journalism, social media, and the public sphere*. New York: Peter Lang. ISBN 978 1433143984
- Cappelletti Júnior, M.; Domínguez Quintas, S. (2014). La curaduría de contenidos y la narrativa colaborativa en el Ciberperiodismo: estudio del caso de Storify en el diario digital elpais.com. *Estudios del mensaje periodístico*, 20 (1) <https://revistas.ucm.es/index.php/ESMP/article/view/45216>. https://doi.org/10.5209/rev_ESMP.2014.v20.n1.45216
- Carbonell Villena, N.; Pedraza-Jiménez, R. (2016). "Imágenes en prensa escrita". En: Pedraza-Jiménez, R.; Codina, L.; Guallar, J. *Calidad en sitios web: método de análisis general, e-commerce, imágenes, hemerotecas y turismo*, 93-128. Editorial UOC. Colección EPI Scholar.
- Cascón-Katchadourian, J.; Ruiz-Rodríguez, A. A.; Alberich-Pascual, J. (2018). Revisión, análisis y evaluación de sistemas para la gestión de activos multimedia en organizaciones. *Revista Española de Documentación Científica*, 41 (1): e196. <https://doi.org/10.3989/redc.2018.1.1481>
- Codina, L. (2000). Evaluación de recursos digitales en línea. *Revista Española de Documentación Científica*, v. 23, n. 1, 9-44. <https://doi.org/10.3989/redc.2000.v23.i1.315>
- Codina, L. (2018). *Curación de contenidos para periodistas: definición, esquema básico y recursos*. Barcelona: Universitat Pompeu Fabra, Facultad de Comunicación, Área de Formatos Digitales e Interactivos, 32 p. <https://repositori.upf.edu/handle/10230/34369>
- Creswell, J.W. (2009). *Research Design. Qualitative, Quantitative and Mixed Methods Approach*. London: Sage.
- Cui, X.; Liu, Y. (2017). How does online news curate linked sources? A content analysis of three online news media. *Journalism*, 18 (7), 852-870. <https://doi.org/10.1177/1464884916663621>
- Denzin, N. K.; Lincoln, Y. S. (2011). *The Sage handbook of qualitative research*. Thousand Oaks, California: Sage.
- Deshpande, P. (2013). Six content curation templates for content annotation. *The Curata blog*, 13 agosto <http://www.curata.com/blog/6-content-curation-templates-for-content-annotation/>
- Díaz Arias, R. (2015). Curaduría periodística, una forma de reconstruir el espacio público. *Estudios del mensaje periodístico*, 21. <http://revistas.ucm.es/index.php/ESMP/article/view/51129>
- Díaz-Campo, J. (2014). Las cadenas de televisión españolas en Internet: un estudio sobre la calidad de sus sitios web. *Estudios sobre el Mensaje Periodístico*, 20 (1), 273-287. <http://revistas.ucm.es/index.php/ESMP/article/view/45219>. https://doi.org/10.5209/rev_ESMP.2014.v20.n1.45219
- DIR Confidencial (2019). El ranking de audiencias de medios digitales, muy cerca de tocar techo tras estrecharse. DIR Confidencial, 23 enero. <https://bit.ly/2BIrsie>
- Ferran-Ferrer, N.; Guallar, J.; Abadal, E.; Server, A. (2017). Research methods and techniques in Spanish library and information science journals (2012-2014). *Information Research*, 22 (1). <http://eprints.rclis.org/32191/>
- Fondevila-Gascón, J F.; Segura-Jiménez, H. (2012). El peso de la hipertextualidad en el periodismo digital: estudio comparativo. *Cuadernos de información*, 30, 31-40. <https://www.redalyc.org/pdf/971/97124309004.pdf>
- Glaser, B.; Strauss, A.L. (1967). *The Discovery of Grounded Theory*. Chicago: Aldine.
- Guallar, J. (2017a). Artículos de curación de contenidos. Categorías y ejemplos. *Anuario ThinkEPI*, 11, 210-216 <http://eprints.rclis.org/31800/>. <https://doi.org/10.3145/thinkepi.2017.38>
- Guallar, J. (2017b). Content curation in digital media: Between retrospective and real-time information. In: Campos Freire, F. et al. *Media and Metamedia Management*, Springer International Publishing, 37-46. ISBN: 9783319460666. https://doi.org/10.1007/978-3-319-46068-0_6
- Guallar, J.; Abadal, E. (2009). Evaluación de las hemerotecas de prensa digital: indicadores y ejemplos de buenas prácticas. *El profesional de la información*, 18 (3), 255-269. <https://doi.org/10.3145/epi.2009.may.02>
- Guallar, J.; Abadal, E.; Codina, L. (2012). Sistema de análisis de hemerotecas de prensa digital. *Tripod*, 31, 37-64. <http://raco.cat/index.php/Tripodos/article/view/262072>
- Guallar, J.; Codina, L. (2018). Journalistic content curation and news librarianship: Differential characteristics and necessary convergence. *El profesional de la información*, 27 (4), 778-791. <https://doi.org/10.3145/epi.2018.jul.07>
- Guallar, J.; Ferran-Ferrer, N.; Abadal, E.; Server, A. (2017). Revistas científicas españolas de información y documentación: análisis temático y metodológico. *El profesional de la información*, 26 (5), 947-960. <https://doi.org/10.3145/epi.2017.sep.16>
- Guallar, J.; Leiva-Aguilera, J. (2013). *El content curator. Guía básica para el nuevo profesional de internet*. Barcelona: Editorial UOC, 2013. ISBN 978-84-9064-018-0
- Guerrini, F. (2013). *Newsroom curators & independent storytellers: content curation as a new form of journalism*. Reuters Institute for the Study of Journalism, University of Oxford. <https://reutersinstitute.politics.ox.ac.uk/our-research/newsroom-curators-and-independent-storytellers-content-curation-new-form-journalism>
- Karlsson, M.; Clerwall, C.; Örnebring, H. (2015). "Hyperlinking practices in Swedish online news 2007-2013: The rise, fall, and stagnation of hyperlinking as a journalistic tool". *Information, communication & society*, 18 (7), 847-863. <https://doi.org/10.1080/1369118X.2014.984743>
- Linares, J.; Codina, L.; Abadal, E.; Guallar, J. (2016). Periodismo en Bases de Datos y Buscabilidad de la Información: protocolo de análisis. *Hipertext.net: Anuario Académico Sobre Documentación Digital y Comunicación Interactiva*, 14, 1-44. <https://doi.org/10.2436/20.8050.01.33>
- Martínez-Cañadas, E. (2017). *Curación de contenidos para bibliotecas*. Barcelona: Editorial UOC, col. *El profesional de la información*, n. 42. ISBN: 9788490293706

- OJD Interactiva (2019). *Medios digitales*. <https://www.ojdinteractiva.es/medios-digitales>
- Orero, P.; Cebrián-Enrique, B. J. (2019). Evolución del uso del hipertexto en la prensa española (2002-2016). *El profesional de la información*, 28 (2), e280225. <https://doi.org/10.3145/epi.2019.mar.25>
- Parra-Valero, P. (2017). Curación de contenidos desde bibliotecas: competencias, herramientas y aplicaciones. *Ciência da Informação*, 45 (2), 103-117. <http://revista.ibict.br/ciinf/article/view/3805>
- Pedraza-Jiménez, R.; Codina, L.; Guallar, J. (2016). *Calidad en sitios web: método de análisis general, e-commerce, imágenes, hemerotecas y turismo*. Barcelona: Editorial UOC. Colección EPI scholar.
- Perpinyà Morera, R.; Cid-Leal, P. (2018). Los portales de archivos españoles: transparencia, interoperabilidad y orientación a los usuarios. *Revista Española de Documentación Científica*, 41 (3): e212. <https://doi.org/10.3989/redc.2018.3.1507>
- Rodríguez-Martínez, R.; Codina, L.; Pedraza-Jiménez, R. (2010). Cibermedios y web 2.0: modelo de análisis y resultado de aplicación. *El profesional de la información*, 19 (1), 35-44. <https://doi.org/10.3145/epi.2010.ene.05>
- Rodríguez-Martínez, R.; Codina, L.; Pedraza-Jiménez, R. (2012). Indicadores para la evaluación de la calidad en cibermedios: análisis de la interacción y de la adopción de la Web 2.0. *Revista Española de Documentación Científica*, 35 (1), 61-93 <https://doi.org/10.3989/redc.2012.1.858>
- Roig, N.; Pedraza-Jimenez, R. (2016). Comercio electrónico. En: Pedraza-Jimenez, R.; Codina, L.; Guallar, J. (coords.), *Calidad en sitios web. Método de análisis general, e-commerce, imágenes, hemerotecas y turismo*. Colección EPI scholar. Barcelona: Editorial UOC
- Rosenbaum, S. (2011). *Curation Nation: how to win in a world where consumers are creators*. McGraw-Hill, 284 p. ISBN 978-0-07-176039-3
- Sanabre, C.; Pedraza-Jiménez, R.; Codina, L. (2018). WebSite Canvas Model: propuesta de un modelo visual para la ideación estratégica de sitios web. *Revista Española de Documentación Científica*, 41 (4): e221. <https://doi.org/10.3989/redc.2018.4.1542>

NOTICIAS /NEWS

ENDORSE (European Data Conference on Reference Data and Semantics)

Del 16 al 19 de marzo de 2021 se celebró la primera conferencia virtual ENDORSE (European Data Conference on Reference Data and Semantics)¹, organizada por la Oficina de Publicaciones de la Unión Europea en colaboración con el programa ISA² de la Comisión Europea.

Con 650 participantes registrados, e incluso más espectadores en línea, la conferencia ofreció la posibilidad conocer a fondo las novedades en el desarrollo de los recursos semánticos elaborados por la Oficina de Publicaciones y otros organismos europeos, los proyectos en desarrollo, y muchas innovaciones tecnológicas relacionadas con los datos semánticos.

A lo largo de los cuatro días del evento se pudo constatar que la creación, conservación e intercambio de datos de referencia (listas, taxonomías, tesauros y ontologías) son tareas básicas para la gestión del conocimiento. La conferencia reunió a diferentes agentes dedicados al desarrollo de recursos semánticos en varios dominios, con gran protagonismo del Derecho. Así mismo, hubo sesiones dedicadas a soluciones tecnológicas relacionadas con el procesamiento del lenguaje natural, el aprendizaje automático y las técnicas de inteligencia artificial.

El primer día de la conferencia se abordó la gobernanza de datos en entornos inter-organizacionales y, en relación con esto, se destacaron los beneficios de la reutilización de datos para no duplicar trabajos. Además, se trató de la importancia de la interoperabilidad y su papel en que el conocimiento se descubra más fácilmente, y de la trascendencia de la publicación multilingüe. Por otra parte, se analizaron los nuevos perfiles profesionales que será necesario formar en las universidades, como

el científico de datos o el "ingeniero de conocimientos". También las intervenciones de los ponentes hicieron hincapié en la necesidad de implementar buenas prácticas para que los datos sean de calidad y fiables, y estén actualizados. Finalmente, destacaron la importancia de los estándares para modelar, enriquecer y compartir esos datos.

En el segundo y tercer día de la conferencia se impartieron varios talleres de interés, y hubo muchas intervenciones relacionadas con las ontologías desarrolladas para el ámbito legal y económico. Cabe destacar que las comunicaciones dejaron de manifiesto cómo la inteligencia artificial, empleada para actualizar y reasignar automáticamente los grafos de conocimiento, abre ahora interesantes posibilidades de interoperabilidad semántica, lo que permite una mayor eficiencia, precisión y explotación de los datos.

El cuarto día de la conferencia, Daniele Rezzi, especialista de la Comisión Europea y responsable de la política de datos abiertos en la Oficina de Publicaciones, presentó la estrategia europea de datos. Entre sus objetivos está asegurar la fiabilidad en la transacción de datos, la regulación del poder basado en datos, el desarrollo del potencial socio-económico de los datos como bienes públicos, y asegurar la legitimidad e imparcialidad de los datos albergados en los espacios comunes de datos europeos. Además de detallar las acciones para la regulación en cuanto a la gobernanza de los datos, Rezzi se refirió a los dominios en los que los datos tendrán gran importancia en los próximos años, como la agricultura, la energía y el 'pacto verde' o 'acuerdo ecológico' (*green deal*).

A lo largo de la conferencia muchas intervenciones señalaron que, a pesar de todos de todos los

esfuerzos, los depósitos de datos aislados tienden a permanecer, lo que perjudica la interoperabilidad y la recuperación de la información. Por otra parte, los expertos y asistentes pudieron compartir experiencias y reflexionaron sobre buenas prácticas y lecciones aprendidas, y esto puso de manifiesto la necesidad de colaborar más dentro de la comunidad de "expertos semánticos".

Hubo propuestas a favor de una gestión de datos semánticos colaborativa, capaz de involucrar a usuarios, expertos, creadores, legisladores, etc. En el marco europeo, una gestión así de la terminología y los datos crearía una fuente de referencia que podría servir a los Estados miembros, y también atraer la participación del sector privado, fomentando la interoperabilidad transfronteriza e intersectorial.

Las tecnologías semánticas, la inteligencia artificial y una terminología actualizada, pueden jugar un papel esencial en áreas de interés como el aprendizaje automático, la digitalización, la traducción automática, etc. En este sentido, la conferencia ENDORSE ha permitido conocer proyectos innovadores y soluciones, y ha facilitado el necesario intercambio de experiencias.

31 de marzo de 2021

NOTAS

1. <https://op.europa.eu/en/web/endorse>

Teresa Abejón Peña

Centro de Ciencias Humanas y Sociales
Consejo Superior de Investigaciones Científicas
(CSIC)