
Análisis de Redes Sociales aplicado a las publicaciones científicas: el caso de la revista *El Profesional de la Información*

Social Network Analysis applied to scientific publications: the case of the journal Information Professional

José Luis ALONSO BERROCAL (1), Carlos G. FIGUEROLA (2)

Universidad de Salamanca, Instituto de Estudios de la Ciencia y la Tecnología, España (1) berrocal@usal.es. (2) figue@usal.es

Resumen

Se analiza, mediante el empleo de técnicas de análisis de redes sociales, la revista *El Profesional de la Información* (EPI). Se han analizado los años desde 2006 hasta 2023, mediante una descarga de los datos accesibles en el web, empleando un crawler diseñado ad-hoc. Los diferentes análisis de los datos, obtenidos a partir de los metadatos existentes, han permitido caracterizar las conexiones existentes entre los autores de las publicaciones, analizar la coocurrencia de palabras en los títulos de los artículos y detectar las comunidades existentes, tanto de autores como de palabras clave.

Palabras clave: Análisis de redes sociales. Detección de comunidades. Detección de tópicos. Revistas científicas. *Profesional de la Información*, *EI* (revista).

1. Introducción

Profesional de la Información es una de las revistas más importantes en el campo de la Ciencia de la Documentación (Library & Information Science) en español. Procedente de otra revista más antigua llamada *Information World* en Español (IVE), comienza como tal su andadura en 1998. Ha tenido una dilatada trayectoria, en la que su indización en las bases de datos más significativas es un hito importante.

Web of Science la indiza desde 2006 en los campos temáticos de Communication y en Information Science & Library Science:

<https://jcr.clarivate.com/jcr-jp/journal-profile?journal=PROF%20INFORM&year=2022>)

En el primero de ellos aparece en Q3 en 2019 y en Q1 en 2022, mientras que en el segundo estaba en Q3 en 2019 y en Q2 en 2022.

Scopus, por su parte, la indiza también en ambos campos desde 2006:

<https://www.scopus.com/sourceid/6200180164>

y la sitúa en 2022 en Q1 tanto con Communication como en Library & Information Science.

Abstract

The journal *El Profesional de la Información* (EPI) is analysed using social network analysis techniques. The years from 2006 to 2023 were analysed by downloading the data available on the web using an ad hoc crawler. The different analyses of the data, obtained from the existing metadata, made it possible to characterise the existing links between the authors of the publications, to analyse the co-occurrence of words in the titles of the articles and to detect the existing communities, both of authors and of keywords.

Keywords: Social network analysis. Detection of communities. Topic detection. Scientific journals. *Information professional* (journal).

Se trata, en consecuencia, de una revista importante dentro de lo que podemos considerar la investigación en Ciencia de la Documentación reciente (entendiendo por reciente las dos últimas décadas, aproximadamente). El hecho de que se trate de una revista radicada geográficamente en España y, como se verá más adelante, de que sus autores sean mayoritariamente españoles y pertenecientes a organizaciones españolas, sugieren una especial representatividad de la investigación española de calidad en el campo de las Ciencias de la Documentación.

Así pues, tiene sentido analizar la producción publicada en esta revista desde diversos puntos de vista. Entre ellos, desde las autorías y las redes de colaboración; y también desde la estructura temática y su evolución. En lo que sigue, este trabajo se estructura de la siguiente manera: primero, se ofrece un estado del arte en las técnicas relacionadas con las coautorías y el análisis y modelado de temas aplicando técnicas de redes. Después se explicará la metodología seguida para la recolección y preparación de datos. A continuación, se expondrán y discutirán los resultados obtenidos, para finalizar con unas conclusiones.

2. Estado de la cuestión

2.1. Análisis de coautoría y de coocurrencia de palabras

Los investigadores nos tenemos que enfrentar a un entorno que cambia de forma muy rápida, debido al desarrollo constante de técnicas, herramientas y métodos científicos. Los niveles de colaboración entre los investigadores han aumentado de forma significativa, con la idea principal de mejorar tanto los niveles de calidad, como la eficiencia y la visibilidad de la investigación científica (Beaver, 2001).

Este crecimiento de las colaboraciones se produce en diferentes especialidades (Cronin, 2004), así tenemos que no solamente crecen en ciencias físicas (Wuchty et al., 2007), si no en ciencias sociales (Ossenblok et al., 2014). Un mayor detalle se puede encontrar en Henriksen (2016), con un detallado estudio del aumento en diferentes disciplinas de las ciencias sociales.

Una buena forma de demostrar las redes sociales que se generan entre investigadores, en un determinado campo, es el análisis de la coautoría. Este es un campo de estudio muy popular en investigaciones bibliométricas, y que abarca estudios de cocitas entre los autores, la coautoría y también análisis de coocurrencia de palabras (Leydesdorff & Vaughan, 2006). En el análisis de la coautoría podemos observar la evolución existente en las redes sociales entre los investigadores. Un aspecto importante de esta línea de trabajo es identificar los actores principales y sus relaciones (Perianes-Rodríguez et al., 2010).

Las redes se componen de un conjunto de actores (los autores de los artículos) y las interacciones que se producen entre ellos. Las interacciones las forman los lazos que relacionan a los autores y que permiten construir la estructura social. Los componentes principales de la red social son el actor y su vínculo relacional. En Análisis de Redes Sociales los actores de la red social se corresponden con el término nodo (extraído de la teoría de grafos). Las conexiones relacionales son las relaciones entre los actores, que incluyen diferentes tipologías de las relaciones, de colaboración, de amistad, flujos personales de una organización, etc. (Molina, 2001; Wasserman & Faust, 2013). Los lazos de red son los patrones relacionales estables que representan la posición de un individuo dentro de una red (Stevenson & Greenberg, 2000).

Los trabajos publicados sobre el análisis de redes sociales de coautoría son muy variados y con numerosos puntos de vista analizados. (Newman, 2004) trabajó con redes de coautoría

en Biología, Física y Matemáticas. Los resultados reflejaron que la mayoría de los investigadores tienen solo unos pocos coautores, mientras que algunos tienen muchos (Adamic, 2000). Leydesdorff & Wagner (2008) trabajaron en la red de colaboraciones que se producía en artículos de Science Citation Index y llegaron a la conclusión de que la red de ciencia, en su conjunto, está aumentando.

Otros autores han trabajado en el análisis de las posiciones de los autores y cómo afecta a las redes de colaboración (Kong et al., 2019), y muchos trabajos que incorporan el cálculo de medidas propias del análisis de redes sociales, como intermediación, cercanía, diámetro, etc. para analizar desde otras perspectivas estas redes (Cheng et al., 2019).

2.2. Detección de temas emergentes

El Procesamiento del Lenguaje Natural (NLP) es una disciplina que nace en 1960 ligada a la Inteligencia Artificial y la Lingüística. Su principal objetivo es estudiar los problemas derivados de la generación y comprensión automática del lenguaje natural.

Gracias a las aportaciones realizadas en los últimos años en los métodos de trabajo, se pueden procesar enormes cantidades de información textual. En las áreas de Recuperación de Información Textual las técnicas de NLP son especialmente útiles.

En los sistemas de recuperación de información se emplea a menudo el procesamiento estadístico del lenguaje. A partir de este modelo cada documento está descrito por un conjunto de palabras clave que lo representan. Este concepto se conoce como bolsa de palabras (bag of words). Consiste en que las palabras de un documento se asignan como palabras claves con un peso en función de su importancia dentro del documento, basándose generalmente en la frecuencia de aparición. De esta manera, todas las palabras que tiene el documento se consideran como términos índices para ese documento. Además, se asigna un peso a cada término en función de su importancia, determinada normalmente por su frecuencia de aparición en el documento.

Una de las aplicaciones del NLP son las técnicas automáticas de clasificación de textos, que permiten asignar automáticamente categorías a una colección de documentos existentes. Uno de los puntos fundamentales de la aplicación de estas técnicas es la detección de topics, en inglés Topic Detection. Estos métodos son un conjunto de

algoritmos cuyo objetivo es descubrir los principales temas dentro de una colección estructurada de documentos. Por tanto, definir una estructura temática en una colección de documentos textuales. En los últimos tiempos estas técnicas se han aplicado a la web y a las redes sociales gracias a las nuevas herramientas tecnológicas (Blei, 2012). Debemos tener en cuenta que el Topic Detection and Tracking (TDT) es un track que comenzó en las Text Retrieval Conferences (TREC) (Allan, 2002).

Son métodos estadísticos que analizan las palabras de los textos originales para describir los temas que están tras cada documento. Además, permite organizar los archivos o resumirlos de forma automática. Para explicar el funcionamiento del modelo partimos de una colección de documentos en los cuales queremos identificar los temas subyacentes para poder organizar la colección. Cada documento contiene una mezcla de diferentes temas. Un tema es un conjunto de palabras que poseen una distinta probabilidad de aparición en los documentos. Por tanto, se asignan palabras al azar a cada tema y se realizan distintas pruebas hasta generar los mejores resultados y más consistentes (Underwood, 2012).

Para que el modelo sea efectivo es necesario probar distintas variaciones en la elección de palabras vacías o sobre el número de temas. Por tanto, el modelo se puede adaptar a la perspectiva o necesidades del investigador.

Uno de los algoritmos que mejor resultado ofrece es el LDA ("Latent Dirichlet allocation") diseñado por (Blei, 2012). Es un modelo estadístico que trata de captar la intuición para definir los temas que subyacen en cada documento de la colección que se analiza. Es por tanto un conjunto de algoritmos capaces de detectar y extraer las relaciones semánticas latentes en cada documento.

Se trata de un modelo de aprendizaje supervisado para caracterizar el contenido de los mensajes (Benhardus & Kalita, 2013). En él, cada documento se representa mediante una bolsa de términos, que son la única variable observada, a través de la cual se intenta extraer los temas tendencia. Es decir, LDA tiene como entrada un conjunto de términos correspondientes a la representación de cada documento de la colección, dando como salida los temas latentes de cada conjunto de palabras, que es lo mismo que los temas latentes de cada documento. Formalmente, un documento se asocia con una distribución multinomial de temas que a su vez son distribuciones multinomiales de palabras.

Un topic o tema es un conjunto de palabras que tienden a coocurrir juntas en los mismos contextos. Para ello los programas analizan su frecuencia de aparición en cada documento. El número de temas se define en un primer momento y puede ser reconsiderado en función de los resultados. Todos los documentos de la colección comparten el mismo conjunto de temas, pero cada documento pertenece a un tema de forma mayoritaria. No siempre este proceso es fácil y algunos documentos pueden presentar similitudes hacia varios temas diferentes. Por ello cada documento muestra la proporción que posee para estar relacionado con ese topic. El propio sistema es el que clasifica cada documento dentro de cada topic, una tarea manual que para una gran colección de documentos sería impensable.

Por tanto, el modelo LDA representa cada documento como una mezcla de los temas en los que aparecen ciertas palabras con sus probabilidades. El modelo LDA funciona de la siguiente manera (Cheng et al., 2019). Partimos de la base de un conjunto de documentos en el que hemos elegido un número fijo de temas T para descubrir. Deseamos conocer la representación de los temas en cada documento y las palabras asociadas a cada tema. Básicamente, el análisis de temas con LDA presupone que, en una colección de documentos hay n temas, representados por un conjunto de términos; cada documento contiene una determinada proporción de cada tema. Esa proporción puede ser 0 para determinados temas en determinados documentos, aunque lo más habitual es que se trate de cantidades muy pequeñas.

El modelo LDA nos proporciona una herramienta poderosa para el descubrimiento y explotación de la estructura temática de los archivos. Sin embargo, la formulación del LDA como modelo probabilístico presenta más ventajas y posibilidades. Una ampliación a este modelo es la incorporación de metadatos en los documentos tales como autor, títulos, enlaces, etc. Por ejemplo, en muchas colecciones de documentos estos están unidos por citas o relaciones entre las páginas webs mediante hipervínculos.

En este punto el modelo relacional de temas supone que cada documento se modela como en el LDA y los vínculos entre los documentos dependen de la distancia hacia las proporciones de los temas de cada documento. Es una buena solución para la gestión, organización o descripción de una gran cantidad de archivos de texto. Podemos decir que este método de análisis de temas puede ser aplicado en diversos campos como: ciencia política, bibliometría, medicina, psicología, etc. Debemos considerar que este tema es un nuevo campo emergente en el aprendizaje de

máquina en el que aún queda mucho por investigar en el futuro.

2.3. Detección de comunidades

En el contexto de las redes, al hablar de comunidad nos referimos a un conjunto de nodos de la red que están más densamente conectados entre sí que con el resto de la red. Existen muchas técnicas para la detección de comunidades, como los algoritmos de agrupamiento jerárquico, métodos basados en cliques, agrupamiento por cortes, algoritmo Girvan-Newman, etc.

Un método ampliamente utilizado es el análisis de modularidad (el número de vínculos entre grupos es pequeño, dentro de grupos es alto), destacando el algoritmo Louvain.

Para finalizar con este apartado queremos mencionar otros trabajos que analizan publicaciones de revistas, en diferentes ámbitos, en su mayoría trabajando sobre los resultados de WoS: el trabajo de Guerrero-Castillo et al. (2023), donde se realiza un análisis de la misma revista, centrándose en el estudio de las palabras clave y generando las comunidades que se obtienen a partir de la coocurrencia de la mismas; el trabajo de Alonso Berrocal (2022) en el que se analizan la autoría y la detección de topics de los artículos publicados en la revista *Fonseca Journal of Communication*, en la década 2010-2020; el trabajo de Guallar et al. (2020) realiza un estudio bibliométrico de 1.226 publicaciones entre los años 2015-2019, pertenecientes a las revistas españolas indexadas en WoS en la categoría Information Science & Library Science; otro trabajo que analiza el Profesional de la Información, en el periodo 2006-2017 es el de López-Robles et al. (2019), en el que se trabaja sobre 1.308 artículos extraídos de WoS y aplicando el software SciMAT; el trabajo de Olmeda-Gómez et al. (2017) analiza los temas de 2.247 artículos de autores españoles indexados en revistas de Información y Documentación en WoS en el periodo 1985-2014; otros estudios que analizan 580 artículos publicados en las siete revistas españolas indexadas en Scopus o WoS en el periodo 2012-2014 son (Ferran-Ferrer et al., 2017; Guallar et al., 2017).

La aportación del trabajo, frente a los más recientes realizados sobre la misma publicación, ofrecen un mecanismo de recogida de datos no basado en la descarga de datos sobre WoS (se realiza mediante un crawler), se utiliza el cálculo de índices de red globales, que caracterizan la red completa (diámetros, densidad): índices de centralidad (intermediación, pagerank) que caracterizan las características de cada nodo individualmente y se aplica un algoritmo de detección de comunidades, que entendemos es más eficaz

con el mecanismo de recolección de datos realizado. Ofrece una visión desde una nueva perspectiva, dado que los estudios anteriores sobre la publicación utilizan procedimientos y software muy similar.

3. Metodología

La revista se publica en abierto en Internet, utilizando la plataforma Open Journal Systems (OJS) (Alperin et al., 2019) y se ha recogido la información que comprende los años 2006-2023, que en el momento de la recogida de datos suponía el 100% de los publicados. Como es habitual con este sistema, desde una página principal se accede a cada número publicado y, desde aquí, a los contenidos de cada número. Cada artículo, del tipo que sea, tiene su propia página.

La página de cada artículo lleva metadatos siguiendo la lógica establecida por el propio sistema OJS: Dublin Core (Weibel & Koch, 2000), por un lado y, por otro, metadatos para la indexación por Google Scholar (Noruzi, 2005). Las Figuras 1 y 2 muestran algunos de los metadatos nativos de un artículo.

```
<meta name="gs_meta_revision" content="1.1"/>
<meta name="citation_journal_title" content="Profesional de la información" />
<meta name="citation_journal_abbrev" content="EPI"/>
<meta name="citation_issn" content="1699-2407"/>
<meta name="citation_author" content="Mariza Almeida"/>
<meta name="citation_author_institution" content="Universidad Federal del E" />
<meta name="citation_author" content="Igone Porto-Gómez" />
<meta name="citation_author_institution" content="Universidad del País Vasc" />
<meta name="citation_author" content="Loet Leydesdorff"/>
<meta name="citation_author_institution" content="University of Amsterdam"/>
<meta name="citation_title" content="Are Brazilian innovation systems innov" />
<meta name="citation_language" content="en"/>
<meta name="citation_date" content="2023/12/29"/>
<meta name="citation_volume" content="32"/>
<meta name="citation_issue" content="7"/>
<meta name="citation_doi" content="10.3145/epi.2023.dic.07"/>
```

Figura 1. Metadatos Google Scholar

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<meta name="DC.Creator.PersonalName" content="Mariza Almeida"/>
<meta name="DC.Creator.PersonalName" content="Igone Porto-Gómez" />
<meta name="DC.Creator.PersonalName" content="Loet Leydesdorff"/>
<meta name="DC.Date.created" scheme="ISO8601" content="2023-12-29"/>
<meta name="DC.Date.dateSubmitted" scheme="ISO8601" content="2023-12-12"/>
<meta name="DC.Date.issued" scheme="ISO8601" content="2023-12-29"/>
<meta name="DC.Date.modified" scheme="ISO8601" content="2023-12-29"/>
<meta name="DC.Description" xml:lang="en" content="A knowledge-based economy" />
<meta name="DC.Description" xml:lang="es" content="A knowledge-based economy" />
<meta name="DC.Format" scheme="IMT" content="application/pdf"/>
<meta name="DC.Format" scheme="IMT" content="application/pdf"/>
<meta name="DC.Identifier" content="87536"/>
<meta name="DC.Identifier.DOI" content="10.3145/epi.2023.dic.07"/>
<meta name="DC.Identifier.URI" content="https://revista.profesionaldelainform" />
<meta name="DC.Language" scheme="ISO639-1" content="en"/>
<meta name="DC.Rights" content="Derechos de autor 2023 Profesional de la info" />
<meta name="DC.Source" content="https://creativecommons.org/licenses/by/4.0"/>
<meta name="DC.Source" content="Profesional de la información / Information I" />
<meta name="DC.Source.ISSN" content="1699-2407"/>
<meta name="DC.Source.Issue" content="7"/>
<meta name="DC.Source.Volume" content="32"/>
```

Figura 2. Metadatos Dublin Core

Se descargaron las páginas de los artículos y se extrajeron los metadatos de cada uno. No todos los campos son relevantes para el presente trabajo; pero uno especialmente importante es el

DC.Type.articleType. En la (Tabla I) se muestran los diferentes tipos contemplados a lo largo del período analizado y la cantidad de artículos para cada uno de ellos. Obviamente, no todos los tipos (Secciones de la revista, en realidad) se han mantenido en el tiempo. Algunos son muy específicos de un determinado período: Artículos de investigación Covid-19, por ejemplo.

Tipo	N.º
Agenda	28
Análisis	469
Artificial Intelligence	16
Artículo especial	2
Artículos de investigación	1091
Artículos de investigación Covid-19	49
Artículos de revisión	30
Editorial	114
Entrevistas	15
Especial sobre la revista Educación y biblioteca	3
Evaladores	2
Humor académico	1
Indicadores	41
Informe técnico	6
Letters	14
Nota técnica	2
Nota de investigación / Research note	4
Patrocinado / Sponsored	1
Reseñas	49
Software	8
Software documental	20

Tabla I. Diferentes tipos de campos contemplados

Igualmente, no todas las secciones son relevantes para el presente trabajo; Agenda, Humor Académico, son ejemplos claros. Se han considerado relevantes aquéllos que disponían de título específico, palabras clave, resumen y referencias.

Después de una revisión manual exhaustiva, los tipos o secciones que hemos considerado relevantes se muestran en la (Tabla II). Éste ha sido el material básico para nuestro estudio.

En general, los metadatos descargados adolecen de algunas inconsistencias en la codificación de los caracteres, aun cuando el propio código HTML en que van insertados esos metadatos especifique claramente la utilización de UTF-8. Se aplicaron utilidades estándar de conversión y, en los casos en que esas utilidades no fueron eficaces, se aplicaron scripts ad-hoc con expresiones

regulares para la detección y posterior sustitución de los caracteres con problemas de codificación.

Tipo	N.º
Análisis	469
Artificial Intelligence	16
Artículo especial	2
Artículos de investigación	1091
Artículos de investigación Covid-19	49
Artículos de revisión	30
Editorial	114
Especial sobre la revista Educación y biblioteca	3
Indicadores	41
Nota de investigación / Research note	4
Software	8
Software documental	20
TOTAL	1847

Tabla II. Tipos o secciones considerados relevantes

4. Resultados

La revista publica artículos en español y en inglés, y una pequeña parte en portugués la Figura 3 muestra la distribución temporal de ambas lenguas y puede apreciarse que los artículos se publican mayoritariamente en español, siendo la presencia de artículos en inglés testimonial hasta 2012. A partir de ese año la presencia de artículos en inglés crece lenta pero ininterrumpidamente, hasta llegar, actualmente, a la tercera parte de los que se publican.

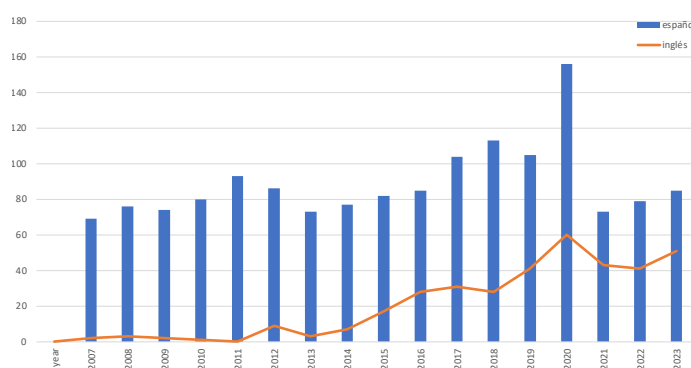


Figura 3. Distribución de idioma

Por lo que se refiere a los autores, es conocido el problema de la falta de homogeneidad formal en las autorías de la literatura académica (Medrano, 2020). A pesar de ello, y tras una revisión semi-manual selectiva, se han podido identificar 2.385 autores únicos. La media de autores firmantes

por artículo es de 2,36 para todo el período; sin embargo, salvo momentos puntuales, el número de autores por artículo tiende a crecer, como puede verse la (Figura 4).

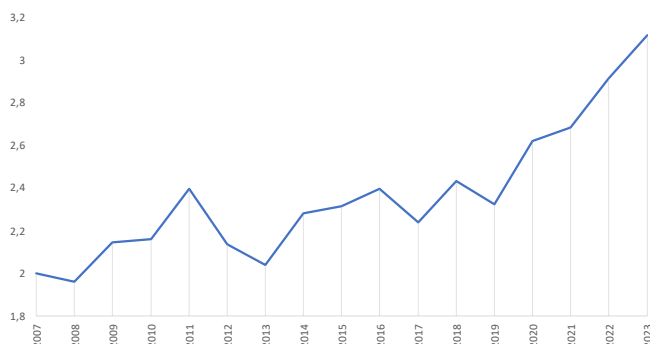


Figura 4. Promedio de autores por artículo

Autor	Artículos
Lluís Codina	44
Daniel Torres-Salinas	28
Javier Guallar	24
Félix De-Moya-Anegón	22
Emilio Delgado-López-Cózar	20
Blanca Rodríguez-Bravo	19
Ernest Abadal	18
Cristòfol Rovira	15
Mario Pérez-Montoro	15
Mike Thelwall	15
Xosé López-García	15
Andreu Casero-Ripollés	15
Carlos Arcila-Calderón	14
Antonia Ferrer-Sapena	13
Pere Masip	13
Enrique Orduña-Malea	13
Rafael Pedraza-Jiménez	13
Ramón Salaverría	13
Josep-Manuel Rodríguez-Gairín	13
Nicolás Robinson-García	13
Fernanda Peset	12
Loet Leydesdorff	12
Evaristo Jiménez-Contreras	11
Enrique Herrera-Viedma	11
Manuel Goyanes	11

Tabla III. Autores y frecuencia de publicación

De esos 2.385 autores, 1.600 firmaron un solo artículo en esta revista; 370 publicaron 2 artículos

en todo el período estudiado; 25 autores han firmado más de 10 artículos. La Tabla III muestra los 25 autores que firman artículos en esta revista con mayor frecuencia. Estos autores, teniendo en cuenta el periodo de tiempo y los números publicados por la revista, forman lo que podríamos llamar el núcleo central de la revista, por lo que a autorías se refiere.

Como se ha comentado antes, pocos artículos se deben a un solo autor; la colaboración de varios conforma una red de autorías que permite percibir estructuras de colaboración, pero también temáticas, dado que los autores se especializan en campos temáticos muy específicos.

La red de coautoría permite ir más allá del simple número de artículos que firma cada autor. Así, podemos construir una red en la que los nodos son los autores, que conectan entre sí cuando comparten la autoría de un artículo.

La red resultante tiene 553 nodos o vértices (es decir, hay 553 autores que co-firman con al menos otro autor). Las métricas más relevantes de esa red se proporcionan en (Tabla IV).

Métrica	Medida
Grado medio	4,007
Densidad	0,007
Diámetro	24
Componentes conexos	36

Tabla IV. Medidas de red

Contrasta el nivel de grado con el diámetro de la red, considerablemente elevado, y más en una red de sólo 553 nodos. Esto sugiere una red estructurada en clusters muy autónomos, pero con sus nodos fuertemente interconectados entre sí, internamente a cada cluster; el número de componentes conexos apoya esta idea.

La propia silueta de esa red (Figura 5, en la página siguiente) muestra comunidades bien marcadas, con débiles conexiones entre unas y otras, al igual que pequeñas colaboraciones claramente periféricas.

La Tabla V muestra los 15 autores con mayor Page Rank y los 15 con mayor Intermediación (Freeman, 1977; Gleich, 2015); algunos de ellos son los mismos, aunque en diferente posición, salvo algunas excepciones. Una mejor posición en la tabla de Intermediación sugiere una mayor actividad multitemática, una mayor capacidad de colaborar con otros campos temáticos.

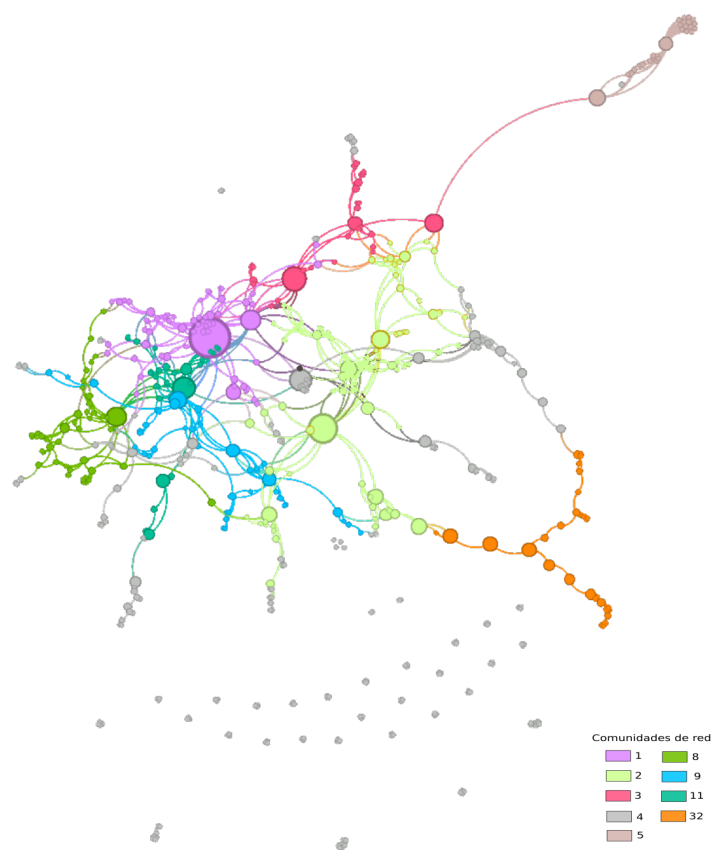


Figura 5. Visualización de comunidades de autores

Page Rank	Intermediación
Lluís Codina	Lluís Codina
Javier Guallar	Rafael Repiso
Daniel Torres-Salinas	Ernest Abadal
Xosé López-García	Ramón Salaverría
Félix De-Moya-Anegón	Félix De-Moya-Anegón
Ramón Salaverría	Javier Guallar
Antonio Castillo-Esparcia	Xosé López-García
Enrique Herrera-Viedma	Marta Somoza-Fernández
Javier Díaz-Noci	Enrique Orduña-Malea
Carlos Arcila-Calderón	Javier Díaz-Noci
Enrique Orduña-Malea	Daniel Torres-Salinas
Nicolás Robinson-García	Concepción Rodríguez-Parada
Ernest Abadal	Carmen Llorente-Barroso
Blanca Rodríguez-Bravo	Tamara Vázquez-Barrio
Pere Masip	Ruth Rodríguez-Martínez

Tabla V. Autores con mayor PageRank e Intermediación

En efecto, muchos autores se especializan en campos temáticos específicos, y esto tiene su reflejo en las comunidades de red. Si acudimos a la detección de comunidades de red (Plantíe & Crampes, 2012) con los algoritmos habituales —Louvain o Leiden en nuestro caso (Traag et al., 2019)—, es fácil detectar comunidades de autores enfocados claramente en el mismo aspecto temático. La Tabla VI muestra las comunidades de autores más significativas.

N	Autores
1	Lluís Codina, Cristòfol Rovira, Guillermo López-García, Rafael Pedraza-Jiménez, Assumpció Huertas, Mario Pérez-Montoro, Santiago Tejedor-Calvo, Frederic Guerrero-Solé, Santiago Giraldo-Luque, Germán Llorca-Abad, Carles Pont-Sorribes, Ariadna Fernández-Planells, Isabel Villegas-Simón, Carlos Lopezosa, María Díez-Garrido
2	Rafael Repiso, Enrique Orduña-Malea, Daniel Torres-Salinas, Nicolás Robinson-García, Enrique Herrera-Viedma, Elea Giménez-Toledo, Adoración Merino-Arribas, Evaristo Jiménez-Contreras, Magdalena Trillo-Domínguez, Emilio Delgado-López-Cózar, María-Dolores Olvera-Lobo, Nuria Lloret-Romero, Isidro F. Aguillo, Luis Rodríguez-Yunta, Javier Salvador-Bruna
3	Ernest Abadal, Javier Guallar, Pere Masip, Núria Ferran-Ferrer, Julià Minguiñón, Pere Franch, Manuel-Jesús Cobo-Martín, Juan-José Boté-Vericad, Remedios Meler, José-Ricardo López-Robles, Jaume Suau, Carlos Ruiz-Caballero, Sue Aran-Ramspott, Nadia-Karina Gamboa-Rosales, Lluís Anglada
4	Félix De-Moya-Anegón, Zaida Chinchilla-Rodríguez, Loet Leydesdorff, Lutz Bornmann, Benjamín Vargas-Quesada, Wenceslao Arroyo-Machado, Han-Woo Park, Rodrigo Sánchez-Jiménez, María-Victoria Nuño-Moral, Antonio Perianes-Rodríguez, Esteban Romero-Frías, Milan Martić, Veljko Jeremic, Vicente P. Guerrero-Bote, Carlos Olmeda-Gómez
5	Concepción Rodríguez-Parada, Blanca Rodríguez-Bravo, Ana-Reyes Pacios, Marina Vianello-Osti, Paz Fernández-y-Fernández-Cuesta, Andrés Fernández-Ramos, Marta De-la-Mano-González, David Nicholas, Eti Herman, Abdullah Abrizah, Chérifa Boukacem-Zeghmour, Hamid R. Jamali, Jie Xu, Marzena Swigon, Carol Tenopir

N	Autores
6	Ruth Rodríguez-Martínez, Jesús Díaz-Campo, Salvador Gómez-García, María-Ángeles Chaparro-Domínguez, Amparo López-Merí, Xavier Ramon-Vegas, Salomé Berrocal-Gonzalo, Ana González-Neira, Andreu Casero-Ripollés, José-Luis Rojas-Torrijos, Marcel Mauri-Ríos, Teresa Piñeiro-Otero, Beatriz Feijoo-Fernández, Francisco Segado-Boj, Lluís Mas-Manchón
8	Xosé López-García, Carmen Costa-Sánchez, Francisco Campos-Freire, María-Isabel Míguez-González, José Rúas-Araújo, Valentín-Alejandro Martínez-Fernández, Miguel Túniz-López, Pablo Vázquez-Sande, María-José Establés, Olga Blasco-Blasco, Berta García-Orosa, María-José Ufarte-Ruiz, Vicente Coll-Serrano, Ana-María López-Cepeda, Marta Rodríguez-Castro
9	Javier Díaz-Noci, Carlos Arcila-Calderón, Ana Serrano-Tellería, Mar Iglesias-García, Elías Said-Hung, Ignacio Aguaded, Rosa Berganza, Lucía García-Carretero, Marta Martín-Llaguno, Luis-Mauricio Calvo-Rubio, Jordi Morales-i-Gras, Dolores Palau-Sampio, Félix Ortega-Mohedano, Sonia Parratt-Fernández, Ainara Larrondo-Ureta
10	Marta Somoza-Fernández, Alexandre López-Borrull, Josep-Manuel Rodríguez-Gairín, Roberto García, Tomás Saorín, Fernanda Peset, Mari-Carmen Marcos, Rafael Aleixandre-Benavent, Aurora González-Teruel, Toni Granollers-Saltiveri, Rosângela-Schwarz Rodrigues, Eva Ortoll, Candela Ollé-Castellà, Tomás Baiget, Gregorio González-Alcaide
11	Ramón Salaverría, Josep-Lluís Micó-Sanz, Santiago Justel-Vázquez, José-Alberto García-Avilés, Pilar Sánchez-García, Montse Bonet, Bienvenido León, María-Carmen Erviti, Alicia De-Lara-González, Miguel Carvajal-Prieto, Gustavo Cardoso, Iván Lacasa-Mas, Toni Sellas-Güell, Alba Diez-Gracia, Félix Árias-Robles
12	Charo Sádaba-Chalezquer, Alfonso Vara-Miguel, Manuel Goyanes, Homero Gil de Zúñiga, Javier Serrano-Puche, Gloria Rosique-Cedillo, Alberto Ardèvol-Abreu, María-Pilar Martínez-Costa, Alejandro Barranquero-Carretero, Brigitte Huber, Florencia Claes, Alba Córdoba-Cabús, Luis Deltell, Carmen Rodríguez-Wangüemert, Eduardo-Francisco Rodríguez-Gómez
13	Antonio Castillo-Esparcia, Ana Castillo-Díaz, Iván Puentes-Rivera, Carmen Carretón-Ballester, Carmen Quiles-Soler, Miguel De-Aguilera-Moyano, Alba-María Martínez-Sala, Lucía Caro-Castaño, Ana-Belén Fernández-Souto, Ana Almansa-Martínez, Paula Pineda-Martínez, Concepción Campillo-Alhama, Inmaculada Berlanga-Fernández, Isabel Ruiz-Mora, Justo Villafañe-Gallego
14	Inmaculada J. Martínez-Martínez, Juan-Miguel Aguado-Terrón, Claudio Feijoo-González, José-Luis Gómez-Barroso, Sergio Ramos-Villaverde, Francisco-José Sarabia-Sánchez, Yannick Boeykens
15	José-Antonio Cordón-García, Raquel Gómez-Díaz, Julio Alonso-Arévalo, Helena Martín-Rodero, Javier Merchán-Sánchez-Jara, Almudena Mangas-Vega, Araceli García-Rodríguez
18	Beatriz Catalina-García, Esther Martínez-Pastor, Ricardo Vizcaino-Laorga, Manuel Montes-Vozmediano, Joaquín López-del-Ramo, José-Antonio Merlo-Vega, Gema Alcolea-Díaz, Natalia Arroyo-Vázquez, José-Antonio Gómez-Hernández, Tony Hernández-Pérez, Pilar Beltrán-Orenes, David Rodríguez-Mateos, Miriam Rodríguez-Pallares, Aurora Cuevas-Cerveró, Manuel Fernández-Sande
19	María-Teresa Fernández-Bajón, José-Antonio Moreiro-González, Juan-Antonio Pastor-Sánchez, Diego Martín-Campo, Tránsito Ferreras-Fernández, Sonia Sánchez-Cuadrado, Jorge Morato-Lara, Carlos-Miguel Tejada-Artigas, María-Jesús Colmenero-Ruiz, Adilson-Luiz Pinto, Julián Urbano, Mónica Marrero, Audilio Gonzales-Aguilar, Miguel-Ángel Marzal-García-Quismondo, Valentín Moreno-Pelayo
24	Jordi Sánchez-Navarro, Leila Mohammadi, Pablo Lara-Navarra, Daniel Aranda, Silvia Martínez-Martínez, Ferran Lalueza, Elisenda Estanyol, Mireia Montaña-Blasco, Cristina Aced-Toledano, Judith Clares-Gavilán, Elena Neira, David Maniega-Legarda, Eva-Patricia Fernández-Manzano, Susana Miquel-Segarra

N	Autores
25	María-José Cantalapiedra-González, Íñigo Marauri-Castillo, Juan-José Gutiérrez-Cuesta, Leire Iturregui-Mardaras, María Ruiz-Aranguren
26	Natalia Papi-Gálvez, Alejandra Hernández-Ruiz, Juan-José Perona-Páez, María-Luz Barbeito-Veloso, Ana-María Enrique-Jiménez, Estrella Barrio-Fraile, Marta Perlado-Lamo-de-Espinosa, Sonia López-Berna
29	Antonio Hidalgo-Nuchera, Julián Chaparro-Peláez, Ángel Hernández-García, Santiago Iglesias-Pradas, Alberto Urueña-López, Antonio Fumero-Reverón
31	Emili Prado, Óscar Coromina, Celina Navarro, Matilde Delgado, Adrián Padilla, Belén Monclús, Núria García-Muñoz
32	Carmen Llorente-Barroso, Tamara Vázquez-Barrio, Rebeca Suárez-Álvarez, Mónica Viñarás-Abad, Francisco Cabezuelo-Lorenzo, María Sánchez-Valle, Javier Sierra-Sánchez, Francisco García-García, Vanessa Rodríguez-Breijo, Francisco-Javier Herrero-Gutiérrez, Daniel Barredo-Ibáñez, Núria Simelio-Solà, Pedro Molina-Rodríguez-Navas, Jorge Gallardo-Camacho, Luis Mañas-Viniegra
33	Sara Martínez-Cardama, Mercedes Caridad-Sebastián, Ana-María Morales-García, Ana R. Pacios, Fátima García-López
35	Concha Pérez-Curiel, Mar García-Gordillo, Ricardo Domínguez-García, Rubén Rivas-de-Roca
38	Jorge Clemente-Mediavilla, Nuria Villagra, Abel Monfort, Rebeca Antolín-Prieto

Tabla VI. Comunidades de autores ordenadas por intermediación

4.1. Instituciones de afiliación

De cara a establecer redes de colaboración, son importantes las instituciones a las que están afiliados los autores que colaboran en los diferentes artículos. Los metadatos de EPI reservan un campo para este dato; sin embargo, su utilización plantea algunos problemas.

El primero es que no todos los artículos ofrecen información al respecto; sólo 776 tienen información válida sobre esta cuestión. En el resto, en la mayor parte de los casos figura como única institución la expresión 'El Profesional de la Información' y, en otros, simplemente está vacío.

El segundo problema es el de la falta de uniformidad en la formulación de las Instituciones. Este problema es bien conocido (Kim & Diesner, 2015), y, de alguna manera, parecido al planteado por el de los nombres de los autores. Sin embargo, en este caso, carecemos todavía de mecanismos similares al ORCID (Haak et al., 2018).

Adicionalmente, tenemos la cuestión de los diferentes niveles de agregación: en unos casos, por ejemplo, figura como institución una universidad determinada, mientras en otros figura una facultad concreta de esa universidad, o incluso un departamento de una facultad de una universidad, y así sucesivamente.

Grado medio	5,72
Densidad	0,022
Diámetro	9
Componentes conexos	5

Tabla VII. Medidas de red para instituciones

<i>Organizaciones con mayor grado</i>	
Universidad Complutense de Madrid	
Universidad de Granada	
Universitat Pompeu Fabra	
Universidad de Málaga	
Universidad Carlos III de Madrid	
Universitat Autònoma de Barcelona	
Universidad de Salamanca	
Universitat Oberta de Catalunya	
Universidade de Santiago de Compostela	
Universitat de Barcelona	
Universidad Rey Juan Carlos	
Universidad de Navarra	
Universidad de Valladolid	
Universidad del País Vasco	
Universidad Internacional de La Rioja	

Tabla VIII. Instituciones ordenadas según su grado

<i>Organizaciones con mayor Page Rank</i>	
Universidad Complutense de Madrid	
Universidad de Granada	
Universitat Pompeu Fabra	
Universidad Carlos III de Madrid	
Universitat de Barcelona	
Universidad de Málaga	
Universidad Rey Juan Carlos	
Universidad de Salamanca	
Universitat Autònoma de Barcelona	
Universitat Oberta de Catalunya	
Universidade de Santiago de Compostela	
Universidad de Navarra	
Universidad Internacional de La Rioja	
Universidad del País Vasco	
Universidad de Valladolid	

Tabla IX. Instituciones ordenadas según su PageRank

Con todo, y tras una revisión manual de las instituciones más frecuentes, es posible obtener informaciones útiles. Así, de 548 instituciones diferentes, unificando manualmente los niveles de agregación, nos quedamos con 235 instituciones.

Construimos una red de coincidencias con las instituciones que colaboran (coocurren) en los mismos artículos, generando una pequeña red de 235 nodos y 596 enlaces. En la Tabla VII tenemos las medidas de red más relevantes y en las Tablas VIII y IX las instituciones ordenadas por valor de grado y de PageRank. En la Figura 6 podemos ver las redes de colaboración bien definidas.

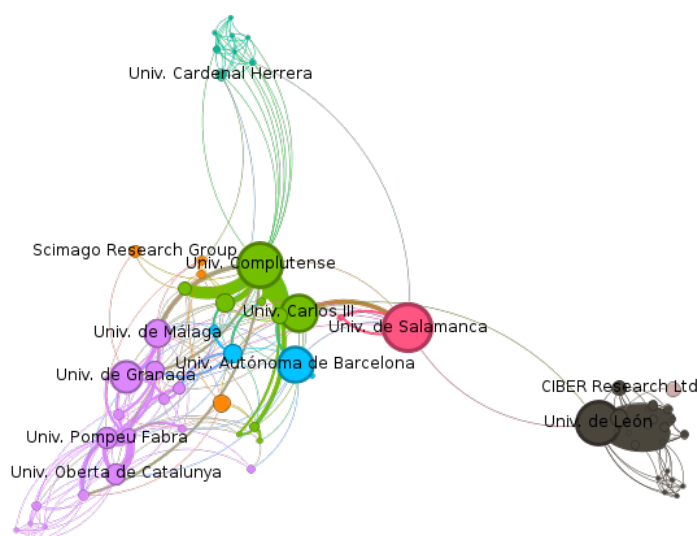


Figura 6. Red de colaboraciones entre instituciones

4.2. Palabras clave

Los metadatos de los artículos contienen palabras clave descriptivas de su contenido; se trata de palabras clave del autor o autores, por lo que no cabe esperar normalización ni uso consistente (Peset et al., 2020). En efecto, encontramos un total de 15.433 palabras clave (8,35 palabras clave por artículo, en promedio), y 6.019 palabras clave únicas.

Las palabras clave coocurren en determinados artículos, formando parejas coincidentes. De un total de 72.204 coincidencias, 62.457 suceden una sola vez. Así, se puede constatar una elevada dispersión en el uso de palabras clave, lo cual es consistente con lo observado en otros trabajos y lleva a plantear la utilidad de tales palabras clave.

Con palabras clave y sus coincidencias o coocurrencia en el mismo artículo, es posible también construir una red (no dirigida) en la que cada palabra clave es un nodo y la coocurrencia de dos de ellas constituye un arco o enlace entre dichos

nodos. La red resultante es marcadamente dispersa; pero con una simple poda de los enlaces que representan una única coocurrencia y la consiguiente de los nodos que resultan con grado igual a 0, resulta una red de sólo 1.334 nodos o palabras clave. Este enfoque ha sido aplicado en otros trabajos, como el de Lozano et al. (2019).

Grado medio	7.9
Grado medio con pesos	22.32
Diámetro	9
Componentes conexos	23

Tabla X. Medidas de red para palabras clave

espana, internet, comunicacion, periodismo, medios de comunicacion, periodismo digital, publicidad, medios digitales, prensa digital, diarios]
covid-19, social media, social networks, coronavirus, journalism, pandemias, disinformation, communication, pandemics, fake news
redes sociales, twitter, medios sociales, comunicacion politica, television, web 2.0, facebook, instagram, participacion, desinformacion
revistas cientificas, acceso abierto, scopus, bibliometria, produccion cientifica, web of science, bases de datos, repositorios, universidades, impacto
bibliotecas universitarias, bibliotecas, digitalizacion, bibliotecas publicas, indicadores, perfiles profesionales, documentacion audiovisual, evolucion, bibliotecarios, usuarios
big data, transparencia, sitios web, datos abiertos, marketing, usabilidad, webs, datos masivos, administracion publica, datos
radio, audio communication, podcasting, podcasts, audio, digital audio, europa, business models, podcast, platforms
spain, scholarly communication, open access, china, transparency, interviews, poland, france, research, early career researchers
relaciones publicas, comunicacion corporativa, comunicacion organizacional, gestion del conocimiento, empresas, profesionales, reputacion, comunicacion interna, responsabilidad social corporativa, organizaciones
web semantica, ontologias, wikipedia, sistemas de informacion, gestion documental, tesauros, clasificaciones, skos, vocabularios controlados, datos enlazados

Tabla XI. Palabras clave con mayor PageRank por comunidades

Algunos de los datos más relevantes de esa red aparecen en la Tabla X. El peso de los arcos se

ha considerado como la simple frecuencia con que cada par de palabras clave coocurre. La formación de comunidades de palabras clave en esta red sugiere bloques de especialización temática, aunque el uso de palabras clave claramente transversales, por un lado; y otras de contenido muy amplio, por otro, desdibujan un tanto esas comunidades. La mezcla de palabras en inglés y español introduce también ruido.

Así, la Tabla XI muestra las palabras clave con mayor Page Rank de las comunidades más importantes obtenidas mediante el algoritmo de Leiden. En la Figura 7 (en la página siguiente) podemos observar la red formada por las palabras clave.

En cualquier caso, técnicas más sofisticadas de análisis de temas o topic modeling pueden ser aplicadas, dado que los propios metadatos incluyen los resúmenes de los artículos. El análisis automático de temas basado en los textos de los resúmenes ha sido aplicado en diversos trabajos de tipo bibliométrico, como en Figuerola et al. (2017) o en García-Marco et al. (2020) y se basa en la aplicación de técnicas de LDA (Latent Dirichlet Allocation).

Existen programas que realizan los cálculos e identifican los temas en los documentos de forma relativamente sencilla (McCallum, 2002).

En este trabajo se manejó un número inicial de 20 topics que, tras analizar los resultados, se re-fundieron o agruparon en 9. Como la técnica nos proporciona el porcentaje o proporción de cada tema en cada resumen de artículo y disponemos de la fecha de publicación de cada uno de esos artículos, es posible establecer la importancia de esos temas y su evolución temporal.

Los grandes temas presentes en la revista EPI son (Tabla XII):

Educación, formación especializada, universidad
Periodismo y medios de comunicación
Bibliotecas
Producción científica, bibliometría
Perspectiva de género
El libro y la edición, incluyendo libro electrónico
Redes sociales
Información y salud, incluyendo lo relacionado con el COVID
Conocimiento Abierto, transparencia

Tabla XII. Temas predominantes en EPI

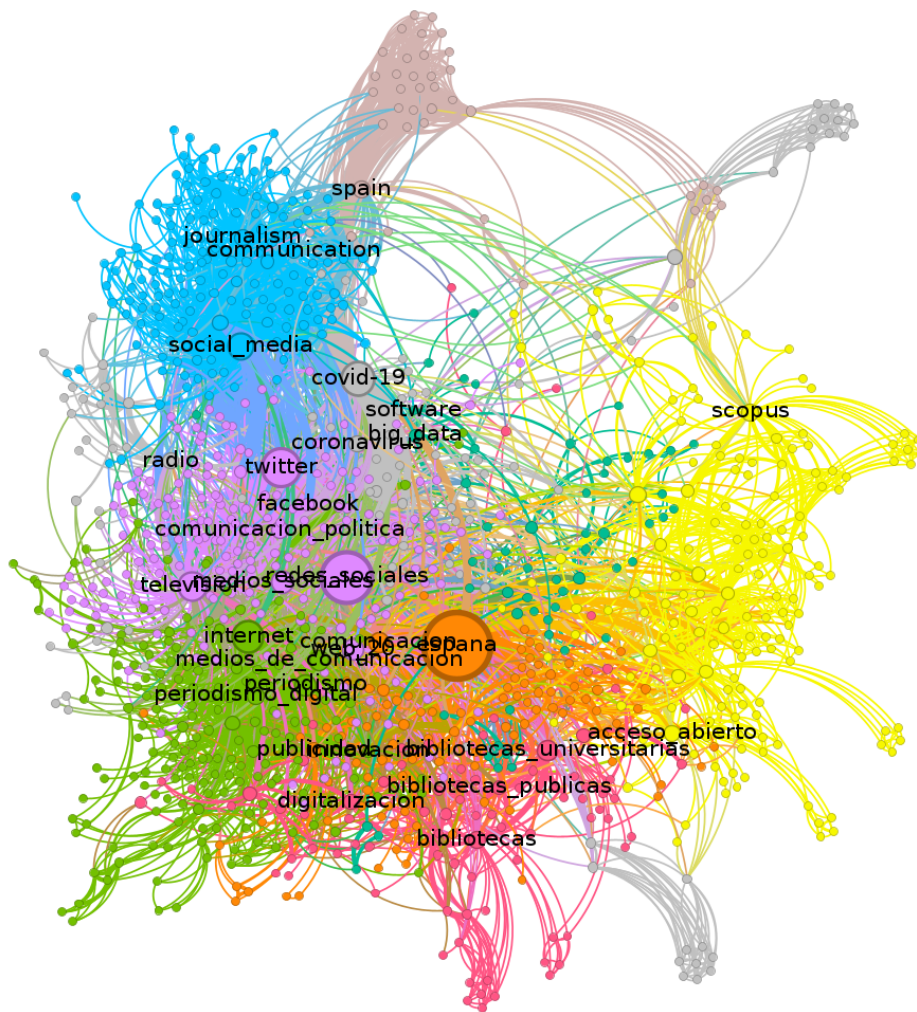


Figura 7. Visualización de la red de palabras clave

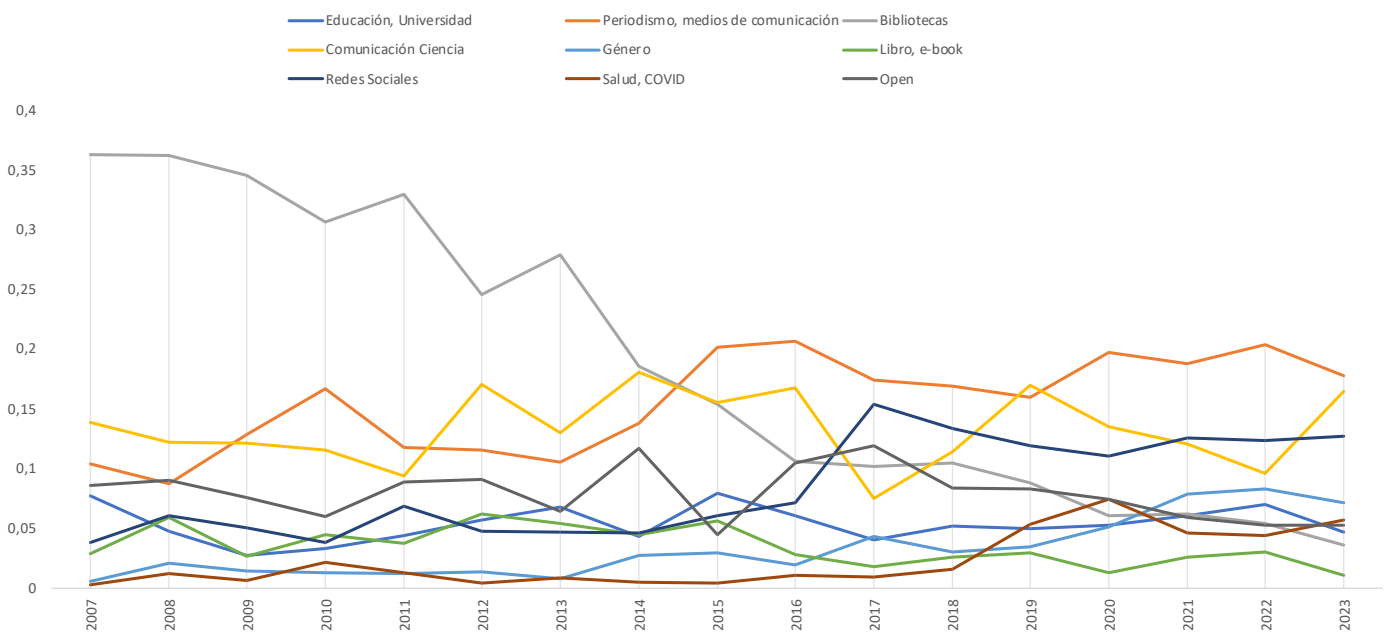


Figura 8. Evolución temporal de los temas

La Figura 9 muestra la importancia o intensidad de cada uno de los nueve temas, mientras que la Figura 8 (en la página anterior) muestra su evolución.

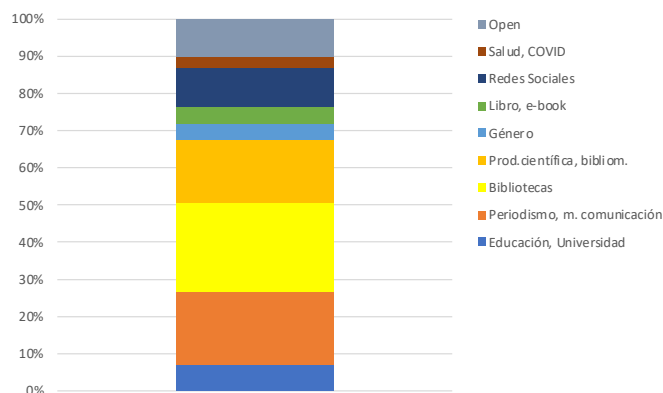


Figura 9. Importancia de los temas principales

Los bloques temáticos, en conjunto, están constituidos por lo relacionado con el periodismo y la comunicación; las bibliotecas y asuntos relacionados; y la bibliometría. Los temas relacionados con las bibliotecas muestran una línea claramente descendente, de manera bastante más suave en los años más recientes. Mientras, lo relacionado con periodismo y medios de comunicación parece evolucionar de manera opuesta. La bibliometría, con picos puntuales, parece tener un nicho bien definido que se mantiene durante todo el período. En los demás temas, menos representados, lo relacionado con las redes sociales muestra un crecimiento notable a partir de 2017; los estudios de género, o relacionados, muy minoritarios hasta 2013, experimentan a partir de entonces un crecimiento considerable.

Lo relacionado con la salud tiene también un incremento notable en los últimos años, relacionado directamente con la pandemia; recordemos también que EPI publica un número especial dedicado a ella. Lo relacionado con el libro y la edición mantiene una línea estable a lo largo del tiempo, al igual que los temas relacionados con la educación y la formación. El conocimiento abierto, en general, parece perder, en los años más recientes, un poco de interés, aunque de forma leve.

5. Conclusiones

La aplicación de las técnicas de análisis de redes sociales permite caracterizar los contenidos de las publicaciones científicas, facilitando la extracción de información, que, de otra manera, dado el volumen de datos existentes, pasaría desapercibida.

A través de las medidas de red de los autores, tenemos unos niveles de conexión extremadamente bajos (densidad), indicando que los niveles de interacción entre autores son relativamente bajos, lo normal son publicaciones entre los mismos grupos de autores. A su vez el alto valor del diámetro es consistente con esta idea.

El promedio de autores ha aumentado en los últimos años, fruto de las colaboraciones más extensas y favorecidas por los equipos multidisciplinares. Las comunidades de autores, nos permiten ver las colaboraciones más fuertes entre los mismos.

Las instituciones a las que pertenecen los autores, si bien no están disponibles en todos los casos, y con las particularidades destacadas, ofrecen un modelo de red en el que podemos ver las instituciones más relevantes, destacando la Universidad Complutense, la Universidad de Granada y la Universidad Pompeu Fabra. La Figura 6 (en la página 21) refleja las interacciones entre dichas instituciones.

Del estudio de las palabras clave y la aplicación de técnicas de detección de topics, hemos obtenido los temas primordiales de la revista. Destacan los temas de bibliotecas, comunicación, redes, producción científica y a pesar de estar limitado en el tiempo el tema sobre el COVID ha dado un impulso a los temas relacionados con información y salud.

Este esquema de trabajo, puede aplicarse a cualquier publicación científica y con la obtención de datos normalizados facilitar la comparación entre publicaciones o analizar una misma publicación a lo largo del tiempo.

Referencias

- Adamic, L. A. (2000). Zipf, power-laws, and pareto-a ranking tutorial. Xerox Palo Alto Research Center, Palo Alto, CA, <http://ginger.hpl.hp.com/shl/papers/ranking/ranking.html>.
- Allan, J. (2002). Introduction to topic detection and tracking. // Topic detection and tracking. Springer. 1-16.
- Alonso Berrocal, J. L. (2022). Análisis de coautoría y detección de topics en la revista Fonseca Journal of Communication (2010-2020). // Gutiérrez San Miguel, M. B. (Ed.). La transferencia del conocimiento en las revistas científicas: Estudio de caso Fonseca Journal of Communication. Tirant Humanidades. 17-34.
- Alperin, J. P.; Willinsky, J.; Owen, B.; MacGregor, J.; Smecher, A.; Stranack, K. (2019). The Public Knowledge Project. // Connecting the Knowledge Commons: From Projects to Sustainable Infrastructure: The 22nd International Conference on Electronic Publishing—Revised Selected Papers, 151.
- Beaver, D. Deb. (2001). Reflections on scientific collaboration (and its study): Past, present, and future. // Scientometrics. 52:3, Article 3. <https://doi.org/10.1023/A:1014254214337>

- Benhardus, J.; Kalita, J. (2013). Streaming trend detection in twitter. // *International Journal of Web Based Communities*. 9:1, Article 1.
- Blei, D. M. (2012). Probabilistic topic models. // *Communications of the ACM*, 55:4, 77-84.
- Cheng, F.-F.; Huang, Y.-W.; Tsaih, D.-C.; Wu, C.-S. (2019). Trend analysis of co-authorship network. // *Library Hi Tech*. 37:1.
- Cronin, B. (2004). Bowling alone together: Academic writing as distributed cognition. // *Journal of the American Society for Information Science and Technology*. 55:6, Article 6. <https://doi.org/10.1002/asi.10406>
- 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://InformationR.net/ir/22-1/paper741.html>
- Figuerola, C. G.; García Marco, F. J.; Pinto, M. (2017). Mapping the evolution of library and information science (1978–2014) using topic modeling on LISA. // *Scientometrics*. 112, 1507-1535.
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. // *Sociometry*. 40:1, 35-41.
- García-Marco, F.-J.; Figuerola, C. G.; Pinto, M. (2020). Análisis de la evolución temática de la investigación sobre Información y Documentación en español en la base de datos LISA mediante modelado temático (1978-2019). // *Profesional de la información*. 29:4, <https://doi.org/10.3145/epi.2020.jul.27>.
- Gleich, D. F. (2015). PageRank beyond the web. // *Sian review*. 57:3, 321-363.
- Guallar, J.; Ferran-Ferrer, N.; Abadal, E.; Server, A. (2017). Revistas científicas españolas de información y documentación: Análisis temático y metodológico. // *Profesional de la información*. 26:5, 947-960. <https://doi.org/10.3145/epi.2017.sep.16>
- Guallar, J.; López-Robles, J.-R.; Abadal, E.; Gamboa-Rosales, N.-K.; Cobo, M.-J. (2020). Revistas españolas de Documentación en Web of Science: Análisis bibliométrico y evolución temática de 2015 a 2019. // *Profesional de la Información*. 29:6. <http://dx.doi.org/10.3145/epi.2020.nov.06>
- Guerrero-Castillo, P.; Nuño-Moral, M.-V.; Guerrero-Bote, V. P.; De-Moya-Anegón, F. (2023). New map of the research published in *Profesional de la Información* (2006-2023). // *El Profesional de la información*. 32:7, e320708. <https://doi.org/10.3145/epi.2023.dic.08>
- Haak, L. L.; Meadows, A.; Brown, J. (2018). Using ORCID, DOI, and other open identifiers in research evaluation. // *Frontiers in Research Metrics and Analytics*. 3, 28.
- Henriksen, D. (2016). The rise in co-authorship in the social sciences (1980–2013). // *Scientometrics*. 107:2, Article 2.
- Kim, J.; Diesner, J. (2015). The effect of data pre-processing on understanding the evolution of collaboration networks. // *Journal of Informetrics*. 9:1, Article 1.
- Kong, X.; Mao, M.; Jiang, H.; Yu, S.; Wan, L. (2019). How does collaboration affect researchers' positions in co-authorship networks? // *Journal of Informetrics*, 13:3, Article 3.
- Leydesdorff, L.; Vaughan, L. (2006). Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment. // *Journal of the American Society for Information Science and Technology*. 57:12, Article 12. <https://doi.org/10.1002/asi.20335>
- Leydesdorff, L.; Wagner, C. S. (2008). International collaboration in science and the formation of a core group. // *Journal of Informetrics*. 2:4, Article 4. <https://doi.org/10.1016/j.joi.2008.07.003>
- López-Robles, J.-R.; Guallar, J.; Otegi-Olaso, J.-R.; Gamboa-Rosales, N.-K. (2019). El profesional de la información (EPI): Bibliometric and thematic analysis (2006-2017). // *El profesional de la información*, 28:4, e280417. <https://doi.org/10.3145/epi.2019.jul.17>
- Lozano, S.; Calzada-Infante, L.; Adenso-Díaz, B.; García, S. (2019). Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature. // *Scientometrics*. 120, 609-629.
- McCallum, A. K. (2002). Mallet: A machine learning for language toolkit. <http://mallet.cs.umass.edu>.
- Medrano, J. F. (2020). Enfoque híbrido para la correcta identificación de autores en bases de datos bibliográficas de libre acceso: El caso de Google Scholar. // XXI Simposio Argentino de Inteligencia Artificial (ASAI 2020)-JAIIO 49 (Modalidad virtual).
- Molina, J. L. (2001). El análisis de redes sociales: Una introducción. Ediciones Bellaterra.
- Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. // *Proceedings of the national academy of sciences*. 101(suppl 1), Article suppl 1.
- Olmeda-Gómez, C.; Ovalle-Perandones, M.-A.; Perianes-Rodríguez, A. (2017). Co-word analysis and thematic landscapes in Spanish information science literature, 1985–2014. // *Scientometrics*. 113:1, 195-217. <https://doi.org/10.1007/s11192-017-2486-8>
- Ossenblok, T. L. B.; Verleysen, F. T.; Engels, T. C. E. (2014). Coauthorship of journal articles and book chapters in the social sciences and humanities (2000-2010): Coauthorship of Journal Articles and Book Chapters in the Social Sciences and Humanities (2000-2010). // *Journal of the Association for Information Science and Technology*. 65:5, Article 5. <https://doi.org/10.1002/asi.23015>
- Perianes-Rodríguez, A.; Olmeda-Gómez, C.; Moya-Anegón, F. (2010). Detecting, identifying and visualizing research groups in co-authorship networks. // *Scientometrics*. 82:2, Article 2. <https://doi.org/10.1007/s11192-009-0040-z>
- Peset, F.; Garzón-Farinós, F.; González, L.-M.; García-Massó, X.; Ferrer-Sapena, A.; Toca-Herrera, J. L.; Sánchez-Pérez, E. A. (2020). Survival analysis of author keywords: An application to the library and information sciences area. // *Journal of the Association for Information Science and Technology*. 71:4, 462-473.
- Plantié, M.; Crampes, M. (2012). Survey on social community detection. // *Social media retrieval*. 65-85. Springer.
- Stevenson, W. B.; Greenberg, D. (2000). Agency and Social Networks: Strategies of Action in a Social Structure of Position, Opposition, and Opportunity. // *Administrative Science Quarterly*. 45:4, Article 4. <https://doi.org/10.2307/2667015>
- Traag, V. A.; Waltman, L.; Van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. // *Scientific Reports*. 9:1, 5233. <https://doi.org/10.1038/s41598-019-41695-z>
- Underwood, T. (2012). Topic modeling made just simple enough. // *The Stone and the Shell*, 7.
- Wasserman, S.; Faust, K. (2013). Análisis de redes sociales: métodos y aplicaciones (Vol. 10). Madrid: CIS-Centro de Investigaciones Sociológicas.
- Weibel, S. L.; Koch, T. (2000). The Dublin core metadata initiative. // *D-lib magazine*. 6:12, 1082-9873.
- Wuchty, S.; Jones, B. F.; Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. // *Science*. 316(5827), Article 5827. <https://doi.org/10.1126/science.1136099>

Enviado: 2024-03-31. Segunda versión: 2024-05-13.
Aceptado: 2024-05-17.

