

Comunicación digital. Tendencias y buenas prácticas

Coordinadores

Javier Guallar; Mari Vállez
Anna Ventura-Cisquella

CUVICOM

Curación, Visualización y Visibilidad en la Comunicación Digital

EPI

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.



CUVICOM

Curación, Visualización y Visibilidad en la Comunicación Digital

Mari Vázquez

Universitat de Barcelona
marivallez@ub.edu

Javier Guallar

Universitat de Barcelona
jguallar@ub.edu



Ediciones Profesionales de la Información SL

Mistral, 36 (Cortijo del Aire)
18220 Albolote (Granada), España
Tel.: +34-639 878 489/ 608 491 521
<https://www.scimagoepi.com>

© Cuvicom - Ediciones Profesionales de la Información SL

1ª edición, Agosto 2025

ISBN: 978-84-125757-5-0

<https://doi.org/10.3145/cuvicom.esp>

<https://www.scimagoepi.com/producto/comunicacion-digital-tendencias-y-buenas-practicas/>

Forma recomendada de citación:

Guallar, J., Vázquez, M., & Ventura-Cisquella, A. (Coords) (2025). *Comunicación digital. Tendencias y buenas prácticas*. Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicom.esp>

Javier Guallar; Mari Vázquez; Ana Ventura-Cisquella
Obra distribuida bajo una licencia CC BY-NC-SA 4.0
<https://creativecommons.org/licenses/by-nc-sa/4.0>



Claves de la comunicación digital de nuestra época (Introducción).....	4
<i>Javier Guallar, Mari Vállez, Anna Ventura-Cisquella</i>	
Visibilidad web y periodismo: consejos y buenas prácticas SEO.....	7
<i>Carlos Lopezosa, Rubén Alcaraz-Martínez, Mari Vállez y Mario Pérez-Montoro</i>	
Servicios de Google para periodistas y medios: recomendaciones para Google Discover y Google News	16
<i>Carlos Lopezosa, Douglas Cordeiro, Olaya López-Munuera y Javier Guallar</i>	
Curación de contenidos en newsletters independientes y de medios de comunicación. Buenas prácticas y recomendaciones	26
<i>Pere Franch y Javier Guallar</i>	
Curación de contenidos en redes sociales de revistas científicas. Tendencias y buenas prácticas.....	39
<i>Wileidys Artigas, Jesús Cascón-Katchadourian y Javier Guallar</i>	
Curación de fuentes y narrativas de igualdad en la agenda mediática	51
<i>Anna Ventura-Cisquella y Javier Guallar</i>	
SEOVisFrame: buenas prácticas y método para evaluar la optimización SEO de las visualizaciones de datos narrativas	72
<i>Rubén Alcaraz-Martínez, Mari Vállez, Raquel Escandell y Diana Bajaña</i>	
Contenido semántico y visualización de información: una propuesta de narrativa comunicacional basada en datos	89
<i>Mario Pérez-Montoro</i>	
¿Cómo movilizar las audiencias en medios de comunicación? Modelos de Activación de la Relevancia y el Engagement de Audiencias en Medios de Comunicación (MAREA)	110
<i>Llúcia Castells-Fos, Lluís Codina y Carles Pont</i>	
Factores para aumentar la visibilidad en repositorios digitales: metadatos, interoperabilidad, identificadores persistentes y optimización SEO-GEO.....	126
<i>Danilo Reyes-Lillo, Cristòfol Rovira y Alejandro Morales-Vargas</i>	
IA en salas de prensa iberoamericanas: casos de uso y mejores prácticas	142
<i>Alexis Apablaza-Campos y Jaime Andrés Wilches Tinjacá</i>	
IA y bancos de imágenes: una metodología de investigación.....	157
<i>Pere Freixa, Mar Redondo-Arolas, Lluís Codina y Carlos Lopezosa</i>	
Pensamiento crítico e inteligencia artificial en la academia: procedimiento de análisis matricial cualitativo para evaluar sistemas de IA.....	171
<i>Lluís Codina, Elisenda Aguilera-Cora, Carlos Lopezosa y Pere Freixa</i>	
Web of Science Research Assistant: análisis funcional y recomendaciones de uso	184
<i>Carlos Lopezosa, Elisenda Aguilera-Cora, Lluís Codina y Juan-José Boté-Vericad</i>	

Claves de la comunicación digital de nuestra época (Introducción)

Javier Guallar

Universitat de Barcelona, España

<https://orcid.org/0000-0002-8601-3990>

Mari Vállez

Universitat de Barcelona, España

<https://orcid.org/0000-0002-3284-2590>

Anna Ventura-Cisquella

Universitat de Barcelona, España

<https://orcid.org/0000-0002-3677-8055>

Guallar, J., Vállez, M., & Ventura-Cisquella, A. (2025). Claves de la comunicación digital de nuestra época (Introducción). En J. Guallar, M. Vállez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 4-6). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicecom.00.esp>

Este libro que hemos titulado *Comunicación digital. Tendencias y buenas prácticas*, pretende ofrecer al público interesado una serie de aportaciones concretas y prácticas, que son el fruto de algunas de las investigaciones académicas que se han realizado en los últimos cuatro años en el marco del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”.

Esta obra ha sido concebida para ser de interés y utilidad a un colectivo muy diverso: profesionales, académicos, estudiantes, estudiosos e interesados en general en el amplio y dinámico mundo de la comunicación digital.

A lo largo de sus capítulos, diferentes especialistas abordan desde la optimización en buscadores hasta la inteligencia artificial, la visualización de información o la curación de contenidos, guiados por una perspectiva y una intencionalidad práctica.

A continuación, señalamos un breve resumen a modo de muestra iniciática de cada capítulo.

- **Visibilidad web y periodismo: consejos y buenas prácticas SEO**, explora la integración estratégica del SEO en el periodismo digital, destacando su rol en la mejora de la visibilidad, credibilidad y sostenibilidad económica de los medios. Además, ofrece recomendaciones prácticas promoviendo un enfoque ético y de calidad en la producción de contenidos.
- **Servicios de Google para periodistas y medios: recomendaciones para Google Discover y Google News**, analiza Google News y Google Discover como plataformas clave en la actualidad en la visibilidad de los medios digitales. Se detallan sus funcionalidades e implicaciones, ofreciendo recomendaciones para optimizar el contenido.
- **Curación de contenidos en newsletters independientes y de medios de comunicación. Buenas prácticas y recomendaciones**, aborda el auge actual de las newsletters como producto comunicativo eficaz, enfatizando la importancia de la curación para conectar con audiencias y ofrecer información especializada. Además, se proponen recomendaciones de uso para medios y para profesionales.
- **Curación de contenidos en redes sociales de revistas científicas. Tendencias y buenas prácticas**, se centra en cómo es el uso de las redes sociales por parte de las revistas científicas, con foco en la curación de los contenidos y en especial en la red social X, ofreciendo recomendaciones prácticas para editores y gestores de revistas.
- **Curación de fuentes y narrativas de igualdad en la agenda mediática**, analiza el trabajo de las periodistas especializadas en igualdad en la prensa digital española, con énfasis en la curación de fuentes documentales y las temáticas que tratan, mostrando como visibilizan desigualdades, violencias de género, avances feministas y derechos LGTBIQ+.
- **SEOVisFrame: buenas prácticas y método para evaluar la optimización SEO de las visualizaciones de datos narrativas**, muestra los factores clave de posicionamiento para las visualizaciones de datos narrativas en los medios digitales, proponiendo la metodología SEOVisFrame para evaluar la optimización SEO de estos productos informativos.
- **Contenido semántico y visualización de información: una propuesta de narrativa comunicacional basada en datos**, presenta los principales contextos comunicacionales y contenidos semánticos que se pueden transmitir con gráficas, ofreciendo una metodología para una visualización de información adecuada.

- **¿Cómo movilizar las audiencias en medios de comunicación? Modelos de Activación de la Relevancia y el Engagement de Audiencias en Medios de Comunicación (MAREA)**, ofrece una metodología práctica para que los medios evalúen y fortalezcan su relevancia y *engagement*, aspectos cruciales para su sostenibilidad, desglosados en tres dimensiones: reputación, visibilidad y fidelización de la audiencia.
- **Factores para aumentar la visibilidad en repositorios digitales: metadatos, interoperabilidad, identificadores persistentes y optimización SEO-GEO**, explora las estrategias clave para mejorar la visibilidad de los repositorios digitales, que son fundamentales para aumentar el alcance e impacto de los contenidos académicos, como la optimización de metadatos o la interoperabilidad y la adopción de identificadores persistentes.
- **IA en salas de prensa iberoamericanas: casos de uso y mejores prácticas**, muestra la integración de la inteligencia artificial en las rutinas periodísticas iberoamericanas, destacando sus diferentes usos, como presentadores virtuales, redacciones automatizadas, conversión de audios, SEO asistido, resúmenes de artículos y *chatbots*.
- **IA y bancos de imágenes: una metodología de investigación**, ofrece un marco metodológico para analizar el sesgo de género y los estereotipos socioculturales en los bancos de imágenes profesionales. Se incluyen las imágenes generadas por IA, revelando algunos patrones estereotipados y clichés persistentes en el contenido visual.
- **Pensamiento crítico e inteligencia artificial en la academia: procedimiento de análisis matricial cualitativo para evaluar sistemas de IA**, propone el Procedimiento Matricial de Análisis de Sistemas de IA (MASIA), un método cualitativo para evaluar el rendimiento de la IA en entornos académicos. Este método está enfocado en la síntesis narrativa, uso de fuentes y formulación de prompts, fomentando el pensamiento crítico y la ética en la investigación.
- **Web of Science Research Assistant: análisis funcional y recomendaciones de uso**, ofrece, por último, un análisis funcional de Web of Science Research Assistant, una herramienta reciente de IA generativa de Clarivate integrada en su conocida base de datos que ayuda en diversas tareas como las revisiones de literatura y agiliza el descubrimiento científico.

En conjunto, *Comunicación digital. Tendencias y buenas prácticas* es una obra de utilidad práctica, que no solo diagnostica el estado actual de la comunicación digital, sino que también ofrece una amplia variedad de herramientas, metodologías, buenas prácticas y recomendaciones para profesionales y académicos. Se han tratado tanto los desafíos éticos como los técnicos de esta era digital, desde la evolución del SEO o la curación de contenidos en diferentes ámbitos hasta la necesidad de la integración responsable de la inteligencia artificial, señalando la relevancia del compromiso con la calidad informativa. Este libro quiere ser una lectura de apoyo, ayuda, referencia, o incluso inspiración, para quienes buscan estar al día y no desorientarse demasiado en estos tiempos de aguda transformación digital en el ámbito de la comunicación.

Javier Guallar, Mari Vállez y Anna Ventura-Cisquella
Barcelona, agosto de 2025

Visibilidad web y periodismo: consejos y buenas prácticas SEO

Carlos Lopezosa

Universitat de Barcelona, España

<https://orcid.org/0000-0001-8619-2194>

Rubén Alcaraz-Martínez

Universitat de Barcelona, España

<https://orcid.org/0000-0002-7185-0227>

Mari Vállez

Universitat de Barcelona, España

<https://orcid.org/0000-0002-3284-2590>

Mario Pérez-Montoro

Universitat de Barcelona, España

<https://orcid.org/0000-0003-2426-8119>

Lopezosa, C., Alcaraz-Martínez, R., Vállez, M., & Pérez-Montoro, M. (2025). Visibilidad web y periodismo: consejos y buenas prácticas SEO. En J. Guallar, M. Vállez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 7-15). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicom.01.esp>

Resumen

Este capítulo explora la integración estratégica del posicionamiento en buscadores (SEO) en el periodismo digital, destacando su papel en la mejora de la visibilidad, la credibilidad y la sostenibilidad económica de los medios de comunicación. Comienza definiendo el SEO y su evolución hacia el SEO semántico y la Optimización de la Experiencia de Búsqueda (SXO), subrayando el equilibrio entre la optimización algorítmica y la integridad periodística. El capítulo expone las mejores prácticas para la investigación de palabras clave, la optimización de titulares, el marcado semántico y las mejoras en la experiencia de usuario, ofreciendo recomendaciones prácticas para periodistas, editores y equipos técnicos. Finalmente, analiza el SEO como un componente crítico de los modelos de negocio sostenibles en los medios, defendiendo un enfoque ético y orientado a la calidad en la producción de contenidos digitales.

Palabras clave

SEO; Periodismo digital; Visibilidad web; SXO; Investigación de palabras clave; SEO semántico; Experiencia de usuario; Sostenibilidad de los medios; Redacciones; Optimización de contenidos.

I. Introducción

El posicionamiento en buscadores es un factor clave para los medios de comunicación online (Codina et al., 2016), ya que incide en cuántos posibles usuarios (lectores) puedan acceder a las noticias publicadas en internet (Dick, 2011). Para lograr que las noticias y reportajes lleguen al mayor número de usuarios posible, los periodistas deben tener en cuenta las buenas prácticas de posicionamiento en buscadores (SEO por sus siglas en inglés), que consisten en un conjunto de estrategias orientadas a mejorar la visibilidad web de los contenidos ya sean noticias, artículos, reportajes, etc., en buscadores como Google.

El SEO aplicado al sector periodístico implica cuestiones tales como la investigación de palabras clave relevantes y la optimización de titulares (Lopezosa y Codina, 2018), el uso óptimo de etiquetas y categorías (Giomelakis y Veglis, 2015a), la estructura clara de los textos (Iglesias-García, y Codina, 2016), y una adecuada utilización de elementos técnicos como, por ejemplo, las metadcripciones y las etiquetas ALT en imágenes (Giomelakis y Veglis, 2015b). Además, es esencial crear contenido original, bien documentado y de calidad (Lopezosa et al., 2019).

Implementar todas estas estrategias no solo permite incrementar la visibilidad y el tráfico orgánico, sino que además ayuda a fortalecer la credibilidad del medio y facilitar que los lectores encuentren y accedan a información fiable y relevante (Lopezosa y Codina, 2018).

Bajo este contexto nace este capítulo, que pretende ofrecer una guía práctica para comprender cómo el SEO puede integrarse de forma ética y eficaz en las redacciones periodísticas.

2. Fundamentos del SEO en medios digitales

Como ya hemos ido avanzando, el SEO es el conjunto de estrategias y técnicas que tienen como objetivo mejorar la visibilidad de un sitio web en los buscadores, especialmente en Google, que domina ampliamente el mercado global y condiciona, en gran medida, las prácticas de optimización (Giomelakis y Veglis, 2015a). En el contexto del periodismo, esto se traduce en maximizar las posibilidades de que una noticia, reportaje o contenido informativo aparezca en los primeros resultados cuando un usuario realiza una búsqueda relacionada (Iglesias-García y Codina, 2016).

En este sentido, la importancia del SEO para los medios tiene una doble vertiente: no solo facilita que una gran parte de los lectores accedan al contenido a través de las búsquedas realizadas (Asser, 2012), sino que además el SEO permite a los medios de comunicación llegar a múltiples públicos objetivos, aumentar su influencia como cibermedio y monetizar mejor los contenidos tanto mediante ingresos publicitarios, como a través de modelos de suscripción.

Se podría pensar entonces que los medios de comunicación están supeditados a pensar en buscadores y no en lectores, pero nada más lejos de la realidad. Es importante recalcar que integrar estrategias de posicionamiento en buscadores en las rutinas de producción periodística no significa subordinar la calidad informativa a las exigencias del algoritmo de Google, sino que se trata de adaptar el contenido para que sea más accesible, encontrable y visible (Smyrniotis y Rebillard, 2009).

Queda claro pues, que el auge de los buscadores como puerta de entrada principal a las noticias ha obligado a los medios a reconfigurar sus rutinas, pero, como veremos, al final lo que importa es la calidad del contenido y poner el foco central en la ciudadanía (Ferran-Ferrer et al., 2013).

A esto hay que sumarle que el SEO ha evolucionado significativamente desde sus inicios, es decir, la forma en la que se hacía es muy distinta de la que se hace ahora e incluso la que se hará en un futuro. De hecho, en la primera etapa, conocida como “SEO tradicional”, la optimización se basaba en aspectos como la densidad de palabras clave, la estructura HTML o el enlazado interno. Sin embargo, este enfoque más técnico y bastante rígido pronto mostró sus límites ante las diferentes actualizaciones de los algoritmos de Google, que comenzaron a valorar más cuestiones centradas en la calidad del contenido y experiencia de usuario (Smyrniotis, 2015).

De ahí surgió el SEO semántico y el SXO (*Search eXperience Optimization*). El SEO semántico pone el foco en el significado de los contenidos partiendo para ello del uso de datos estructurados, etiquetas semánticas (como las propuestas por Schema.org) y una redacción adaptada al lenguaje natural (Lopezosa et al., 2018). Paralelamente, el SXO se centra en la experiencia del usuario poniendo en relieve la necesidad de que el contenido sea útil, navegable, accesible y genere interacción entre el sitio web y el usuario (Alcaraz-Martínez, 2024).

Actualmente, ambos tipos (semántico y experiencial) se complementan y cada vez se integran más en las estrategias de visibilidad de los medios digitales (Giomelakis, 2023).

Las *Search quality rating guidelines* de Google, publicadas por primera vez en 2013 y, actualizadas constantemente (Google, 2025), proporcionan información de primera mano sobre los criterios que emplea este buscador para evaluar la calidad de un sitio web. Entre los conceptos más importantes recogidos en estas directrices encontramos los principios EEAT (*Experience*,

Expertise, Authority, and Trustworthiness) (Alcaraz-Martínez, 2024). Estos principios se centran en diversos aspectos relacionados con cada uno de los componentes del contenido, su autoría y el sitio web en conjunto, con el objetivo de determinar si la información presentada es fiable; es decir, si está elaborada por una persona con conocimientos pertinentes y si aporta una experiencia directa o profesional. Google aplica estos principios con especial atención en ámbitos sensibles como la salud, las finanzas, el derecho o el periodismo, donde la calidad, la veracidad y la responsabilidad informativa son esenciales para evitar la desinformación y proteger al usuario. Mediante la experiencia (*experience*), se evalúa si el autor tiene conocimientos de primera mano sobre el tema tratado. La pericia (*expertise*) se refiere a la formación, competencia o conocimiento especializado del autor sobre el tema. Con la autoridad (*authoritativeness*) se valora si el autor, la página o el sitio son reconocidos como referentes en su campo. Finalmente, la fiabilidad o confianza (*trustworthiness*) mide si el contenido es seguro, preciso, veraz y transparente.

3. Buenas prácticas SEO para periodistas y redacciones

A continuación, se detallan algunas de las prácticas SEO más habituales en medios de comunicación. Concretamente hablaremos de la selección de palabras clave y redacción optimizada, la doble titulación, el marcado semántico y la optimización del diseño, la navegación y la arquitectura web.

En gran medida, toda estrategia de posicionamiento en buscadores parte de la identificación de palabras clave relevantes (Smyrniais y Sire, 2014). En el sector periodístico, estas deben expresar con precisión el tema central de la noticia y, además, coincidir con los términos de búsqueda que los usuarios emplean para encontrar información. Para poder llevar a cabo la investigación de palabras clave, nos podemos apoyar en herramientas de terceros como Google Trends, SEMrush, Ahrefs, Majestic, Sistrix e incluso en las propias palabras sugeridas por el autocompletado de Google.

Una vez se han identificado las palabras clave principales y secundarias, por lo general estas pueden integrarse en el titular (encabezado H1), en el primer párrafo, en los subtítulos (encabezados H2, H3), el cuerpo del texto, en la URL y en las metadescripciones.

Esta integración debe hacerse de forma natural, es decir, sin forzar la redacción ni perjudicar la calidad del contenido. Por lo tanto, la fluidez narrativa y el rigor informativo siguen siendo prioritarios y están por encima de cualquier tipo de estrategia SEO.

Asimismo, una estrategia SEO muy utilizada en medios de comunicación es la práctica de la doble titulación (Asser, 2012). Consiste en aplicar un titular breve y directo en la portada del medio de comunicación, y otro más optimizado para SEO en el interior del contenido. Esta práctica, que inició la BBC y que acabó estandarizándose, permite combinar el impacto editorial con la visibilidad web (Smyrniais, 2015).

Además del titular, también se recomienda estructurar el contenido de manera estratégica. Normalmente se aconseja incluir subtítulos jerárquicos para guiar la lectura utilizando para ello encabezados H2 y H3, escribir párrafos cortos con un lenguaje claro y sencillo, incorporar listas e incluir enlaces internos y externos que aporten valor añadido al usuario que lee la noticia.

Adicionalmente, es fundamental aplicar correctamente marcado semántico a través de tecnologías como Schema.org. En medios de comunicación como mínimo se recomienda

utilizar el esquema *Newsarticle*. Esto puede implementarse directamente en el código HTML o mediante *plugins* especializados incorporados en los gestores de contenidos (CMS) de cada medio de comunicación. Como resultado, es posible optar a los denominados *rich snippets* o fragmentos enriquecidos, que permiten una representación más completa y atractiva en las páginas de resultados de los motores de búsqueda. Los resultados de búsqueda enriquecidos son resultados que incorporan elementos adicionales: valoraciones, imágenes, fechas, autores, categorías, grafos de conocimiento, etc., que ayudan, por un lado, a aumentar el número de clics y, por otro lado, a mejorar la percepción del medio como fuente de referencia (Giomelakis, 2023).

Como ya hemos avanzado, el SEO ha dejado de centrarse únicamente en posicionar contenidos a través de palabras clave o en la obtención de enlaces. En los últimos años, los buscadores y, especialmente Google, tienen en cuenta el comportamiento del usuario como indicador de calidad. Así, surgió el concepto de *Search eXperience Optimization* (Alcaraz-Martínez, 2024), que unifica SEO y usabilidad para garantizar que el contenido sea visible al mismo tiempo que también cumpla con las expectativas del lector.

En el sector periodístico, esto implica que una noticia no solo debe estar bien optimizada desde un punto de vista técnico, sino que debe facilitar una experiencia de lectura satisfactoria a través de un contenido claro, útil y de confianza, además de contar con una navegación óptima. Por lo tanto, si los usuarios abandonan rápidamente la página a la que han ingresado, no interactúan o no navegan hacia otros contenidos del medio, Google lo interpreta como una señal negativa y esto puede perjudicar al posicionamiento del sitio web.

Dentro del SXO, la arquitectura web también juega un papel fundamental ya que su objetivo es favorecer la retención del lector y mejorar su experiencia al navegar por el sitio web. En general, algunas recomendaciones clave para mejorar la experiencia de usuario en este sentido tienen que ver con:

- Diseño del sitio web adaptado a dispositivos móviles; es decir, que el medio digital se visualice correctamente en dispositivos móviles y tabletas.
- Velocidad y percepción del tiempo de carga optimizada, ya que una página lenta o percibida como tal disminuye la permanencia del usuario en el sitio web que visita.
- Estructura jerárquica clara, es decir, contar con menús, categorías y etiquetas bien organizadas.
- Migas de pan (*breadcrumbs*) para orientar de forma óptima al usuario sobre su ubicación e incrementar de forma natural el enlazado interno.
- Enlaces internos que inviten a los lectores a descubrir contenidos de valor relacionados, al mismo tiempo que se trabaja la distribución de la autoridad entre páginas.

Todas estas estrategias forman parte de una arquitectura semántica que no solo mejora la experiencia, sino que facilita la labor de los buscadores para interpretar y jerarquizar los contenidos de un sitio web.

A esto hay que sumarle que los buscadores utilizan múltiples métricas de interacción para definir si un contenido ha satisfecho al usuario y, por lo tanto, cumple con los estándares de calidad. Entre los indicadores más relevantes se encuentran:

- La tasa de rebote: es el porcentaje de usuarios que entran y salen sin visitar una segunda página dentro del dominio.

- El tiempo de permanencia (*dwell time*): es la duración media de la visita, lo que permite diferenciar entre *long-clicks* y *short-clicks*.
- Las páginas vistas por sesión: cuántos contenidos consume el usuario por visita.
- El CTR (*Click Through Rate*): es el porcentaje de clics que recibe un resultado respecto a su número de impresiones en la página de resultados de búsqueda.
- El *pogo-sticking*: ocurre cuando un usuario hace clic en un resultado y vuelve inmediatamente a la página de resultados para elegir otro enlace, lo que implica una señal de insatisfacción.
- Las métricas web principales (*core web vitals*): un conjunto de métricas específicas definidas por Google para evaluar aspectos clave de la experiencia de usuario, relacionadas con la velocidad, la interactividad y la estabilidad visual.
- Otras señales técnicas: compatibilidad móvil, uso de HTTPS...

Todos estos indicadores ayudan a entender y mejorar la experiencia de usuario y por tanto el SEO y el SXO. En el sector periodístico, esto se resume, en definitiva, en contenidos bien redactados, que aporten valor añadido, incluyan material visual atractivo, enlaces útiles y una propuesta clara de navegación.

4. Modelos de negocio en medios digitales y sostenibilidad a través del SEO

Más allá de su importancia como estrategia de visibilidad, el posicionamiento en buscadores es una herramienta clave para la sostenibilidad económica de los medios digitales. Es decir, un buen posicionamiento, además de atraer más lectores, mejora la capacidad del medio para monetizar ese tráfico web mediante publicidad, suscripciones, patrocinios y otros modelos híbridos.

El tráfico orgánico derivado del buen SEO (el SEO centrado en el usuario) es especialmente valioso porque por lo general es estable, segmentado y gratuito (en términos de adquisición).

Sin embargo, la sostenibilidad de los medios no puede lograrse únicamente desde lo técnico o comercial. Debe basarse en un compromiso con la calidad, la independencia y la función social del periodismo. En este sentido, el SEO puede ser un aliado estratégico, siempre que se haga un uso ético del mismo o lo que es lo mismo, se apliquen estrategias de posicionamiento que respeten la integridad del contenido, no suponga manipulación informativa y contribuya a que la información de calidad tenga el mayor alcance posible.

El reto, por lo tanto, está en integrar el posicionamiento en buscadores dentro de una cultura editorial responsable, donde las métricas no sustituyan a los valores periodísticos, sino que sean una herramienta más en favor de la calidad informativa.

5. Recomendaciones SEO para medios

La incorporación del SEO en las redacciones periodísticas no requiere una transformación radical, sino una integración inteligente y progresiva. En lo que sigue, presentamos una serie de recomendaciones prácticas que pueden servir como guía para consolidar una cultura del SEO en los medios.

Para redactores y periodistas:

- Usar títulos informativos claros con palabras clave relevantes y estratégicas. Optimizar en consonancia las etiquetas *title* y H1.
- Utilizar subtítulos jerarquizados, sobre todo encabezados H2 y H3 para organizar el contenido.
- Incluir enlaces internos estratégicos a otros contenidos del medio que ayuden a ampliar la información que se está consumiendo.
- Citar fuentes externas con enlaces relevantes y de calidad.
- Incluir imágenes con su correspondiente texto alternativo (atributo *alt*) y nombres de fichero descriptivos.
- Mejorar la percepción de precisión, veracidad y transparencia del contenido.

Para editores y responsables de contenidos:

- Implantar una política editorial que contemple criterios SEO.
- Supervisar las publicaciones en términos de extensión, frecuencia y calidad.
- Revisar periódicamente los títulos y metadescpciones de las entradas publicadas.
- Incorporar técnicas de doble titulación cuando sea pertinente.
- Mostrar las competencias y el conocimiento especializado de cada redactor mediante páginas de autor con biografías y contenidos.
- Formar a todo el equipo en SEO, usabilidad y experiencia de usuario.

Para técnicos y responsables web:

- Asegurar una arquitectura web clara y escalable.
- Optimizar la velocidad de carga y la compatibilidad a dispositivos móviles y tabletas.
- Integrar marcado semántico (Schema.org), como mínimo el esquema *NewsArticle*.
- Generar *sitemaps* actualizados (general y de noticias).
- Incorporar herramientas de análisis SEO (como Google Search Console, Ahrefs o SEMrush) para monitorear el posicionamiento de las noticias.

Sobre las herramientas, resulta importante destacar que existen numerosas herramientas — gratuitas y de pago— que facilitan la aplicación práctica del SEO en medios.

Tipo	Herramientas
Para auditorías SEO	Screaming Frog, SEMrush, Sistrix
Para análisis de comportamiento de usuario	Google Analytics, Hotjar, Matomo
Para análisis de <i>backlinks</i> y autoridad web	Majestic, Ahrefs

De este modo, aunque tener conocimiento sobre SEO dentro de las redacciones es muy importante, lo esencial es tratar de construir una mentalidad SEO dentro del medio. Para lograrlo, es clave fomentar la colaboración entre redactores, editores y técnicos SEO, así como transmitir al equipo que hay que priorizar siempre la calidad y la ética informativa frente al posicionamiento y la captación de clics.

En definitiva, una redacción orientada a la aplicación de estrategias de posicionamiento no solo ayuda a dar mayor visibilidad a su contenido, sino que además ayuda a mejorar la calidad de su producto informativo, amplía su impacto social y fortalece su sostenibilidad.

6. Conclusión

A lo largo de este capítulo se ha pretendido ofrecer un recorrido práctico por los conceptos, técnicas y recomendaciones fundamentales para aplicar el SEO en medios de comunicación.

Como hemos podido ver, integrar el SEO de forma ética y profesional dentro de las redacciones periodísticas no significa renunciar a los principios del buen periodismo, sino adaptarlos a las exigencias del presente. Por ejemplo, un titular optimizado no es incompatible con la veracidad, un enlace bien construido no rebaja la calidad de un reportaje, una arquitectura pensada para el usuario no empobrece la narrativa. Al contrario, todos estos elementos fortalecen el resultado final de los productos periodísticos.

No queremos acabar este capítulo sin reiterar que el SEO no debe entenderse como una amenaza, sino como una herramienta al servicio del periodismo. De hecho, consideramos que cuando se aplica con criterio, rigor y sensibilidad profesional, los contenidos de calidad acaban alcanzando a más personas, en el momento adecuado y en el formato más accesible para ellos. Por lo tanto, lejos de desvirtuar el oficio, el SEO puede contribuir a amplificar su impacto social, reforzando así el compromiso con la verdad, su utilidad pública y el derecho ciudadano a estar informado.

7. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

8. Referencias

- Alcaraz-Martínez, R. (2024). Integración de la optimización de la experiencia de búsqueda (SXO), la usabilidad, la arquitectura de la información y la accesibilidad web: explorando un terreno común para mejorar la visibilidad en buscadores. *Ibersid: Revista de Sistemas de Información y Documentación*, 18(1), 37–53. <https://doi.org/10.54886/ibersid.v18i1.4966>
- Asser, M. (2012). *Search engine optimisation in BBC News*. <https://bbc.in/2J8TzIR>
- Codina, L., Iglesias-García, M., Pedraza, R., & García-Carretero, L. (2016). *Visibilidad y posicionamiento web de informaciones periodísticas: El framework SEO-RCP*. DigiDoc-UPF. <http://hdl.handle.net/10230/26040>
- Dick, M. (2011). Search engine optimization in UK news production. *Journalism Practice*, 5(4), 462–77. <https://doi.org/10.1080/17512786.2010.551020>
- Ferran-Ferrer, N., Minguillón, J.; Pérez-Montoro, M. (2013). Key factors in the transfer of information-related competencies between academic, workplace and daily life contexts. *Journal of The American Society for Information Science and Technology*, 64(6), 1112–1121. <https://doi.org/10.1002/asi.22817>

- Giromelakis, D., & Veglis, A. (2015a). Employing search engine optimization techniques in online news. *Studies in Media and Communication*, 3(1), 22-33. <https://doi.org/10.11114/smc.v3i1.683>
- Giromelakis, D., & Veglis, A. (2015b). Investigating search engine optimization factors in media websites, the case of Greece. *Digital Journalism*, 4(3), 379-400. <https://doi.org/10.1080/21670811.2015.1046992>
- Giromelakis, D. (2023). Semantic search engine optimization in the news media industry: Challenges and impact on media outlets and journalism practice in Greece. *Social Media+ Society*, 9(3). <https://doi.org/10.1177/20563051231195545>
- Google (2025). *Search quality rating guidelines*. <https://static.googleusercontent.com/media/guidelines.raterhub.com/es//searchqualityevaluatorguidelines.pdf>
- Iglesias-García, M., & Codina, L. (2016). Los cibermedios y la importancia estratégica del posicionamiento en buscadores (SEO). *Opción*, 32(9), 929-944. <https://bit.ly/2J8Skcu>
- Lopezosa, C., & Codina, L. (2018). *Análisis de posicionamiento en medios de comunicación con herramientas SEO: cobertura informativa de los premios Oscar 2017*. Universitat Pompeu Fabra. Departament de Comunicació. (Serie Editorial DigiDoc); (EPI; 2)
- Lopezosa C., Codina, L., & Caldera-Serrano, J. (2018) SEO semántico: framework ISS para la optimización de sitios intensivos en contenidos. *Cuadernos de documentación multimedia*, 29, 97-123. <https://doi.org/10.5209/CDMU.60607>
- Lopezosa, C., Codina, L., & Rovira, C. (2019). *Visibilidad web de portales de televisión y radio en España: ¿qué medios llevan a cabo un mejor posicionamiento en buscadores*. Serie DigiDoc-EPI, 4. <http://hdl.handle.net/10230/36234>
- Smyrniaios, N. (2015). Google and the algorithmic infomediatioin of news. *Media Fields Journal*, 10, 1-10. <https://bit.ly/2XJIMN7>
- Smyrniaios, N., & Rebillard, F. (2009). L'actualité selon Google L'emprise du principal moteur de recherche sur l'information en ligne. *Communication & Langages*, 160, 95-109. <https://doi.org/10.4074/s0336150009002087>
- Smyrniaios, N., & Sire, G. (2014). The news according to Google how does algorithmic infomediatioin frame the work of French journalists? In *JSS-ECREA 2014 Conference*, Thessaloniki. <https://bit.ly/33bA5tN>

Servicios de Google para periodistas y medios: recomendaciones para Google Discover y Google News

Carlos Lopezosa

Universitat de Barcelona, España

<https://orcid.org/0000-0001-8619-2194>

Douglas F. Cordeiro

Universidade Federal de Goiás, Brasil

<https://orcid.org/0000-0002-5187-0036>

Olaya López-Munuera

Universitat Oberta de Catalunya, España

<https://orcid.org/0009-0000-6132-5384>

Javier Guallar

Universitat de Barcelona, España

<https://orcid.org/0000-0002-8601-3990>

Lopezosa, C., Cordeiro, D. F., López-Munuera, O., & Guallar, J. (2025). Servicios de Google para periodistas y medios: recomendaciones para Google Discover y Google News. En J. Guallar, M. Váñez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 16-25). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicom.02.esp>

Resumen

Este capítulo explora el papel de Google News y Google Discover como plataformas clave para mejorar la visibilidad de los medios digitales. A través de un enfoque práctico, analiza sus funcionalidades, implicaciones y diferencias, ofreciendo recomendaciones específicas para periodistas y editores que busquen optimizar contenidos en estos servicios. Google News funciona como un agregador de noticias centrado en la actualidad y la credibilidad, mientras que Google Discover ofrece recomendaciones de contenido personalizadas basadas en el comportamiento del usuario. Se proponen pautas de buenas prácticas para ambas plataformas, con el objetivo de mejorar la presencia digital y afrontar los retos algorítmicos sin comprometer la calidad periodística.

Palabras clave

Google News; Google Discover; Periodismo digital; SEO; Visibilidad web; Algoritmos; Medios de comunicación; Estrategias editoriales.

I. Introducción

La necesaria adaptación del sector periodístico a los avances en Internet ha provocado, a lo largo de los años, transformaciones importantes en las rutinas productivas de los medios de comunicación (Salaverría y García-Avilés, 2008), especialmente en lo que respecta a los modos de difusión de las noticias mediante su visibilidad en buscadores (Iglesias-García y Codina, 2016) y en plataformas de agregación de contenido (Lopezosa et al., 2019).

Bajo este escenario, empresas como Google, con servicios específicos como Google Discover y Google News, se han convertido en herramientas fundamentales para que los medios de comunicación aumenten su visibilidad, amplíen su alcance y, en consecuencia, mejoren sus posibilidades de monetización (Lopezosa et al., 2022; Cordeiro et al., 2024).

Google News es una plataforma de agregación de contenido que recopila y organiza contenidos informativos procedentes de una amplia variedad de fuentes de medios, facilitando a los lectores el acceso a noticias actuales y fiables (Guallar, 2015). Por lo tanto, aparecer en Google News representa para los medios de comunicación no solo un aumento en el número de lectores, sino también una mejora en su credibilidad (Lopezosa et al., 2024).

Por su parte, Google Discover es una plataforma de recomendación de contenidos que se ha consolidado como una de las principales fuentes de tráfico para numerosos medios digitales, especialmente en dispositivos móviles (Lopezosa et al., 2022) y, desde mediados de 2025, en la versión de escritorio de algunos países (González, 2025). Su capacidad para hacer sugerencias personalizadas y modificar el *feed* en tiempo real y sin que el usuario realice una búsqueda activa, influye en las estrategias editoriales, ya que permite a los medios alcanzar nuevas audiencias (Lopezosa et al., 2023).

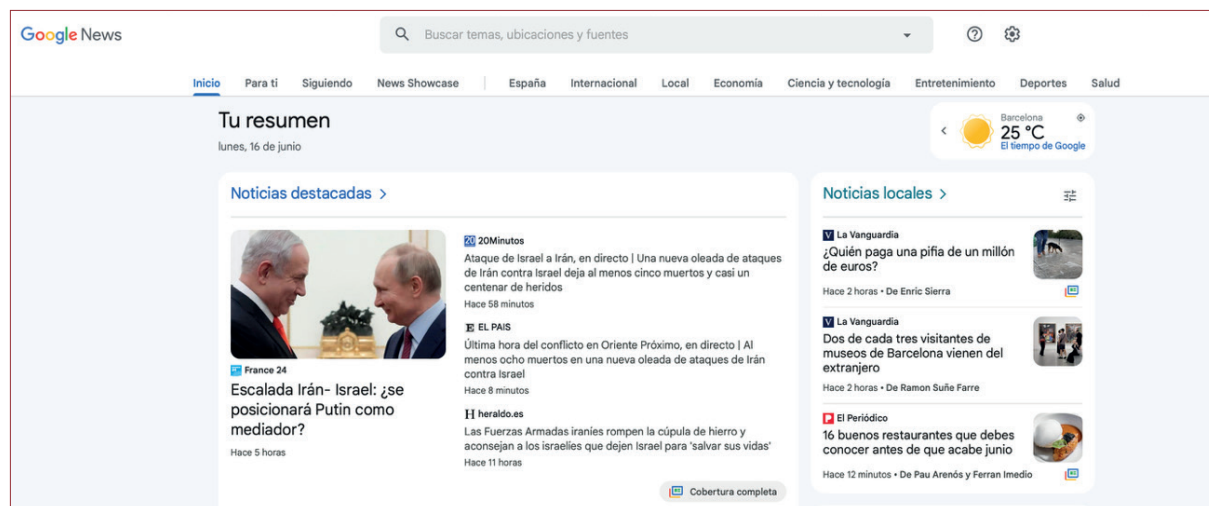
Bajo este contexto, nace este capítulo que trata de analizar ambos servicios de Google desde una perspectiva práctica, centrándose en cómo debe optimizarse el contenido para que los medios de comunicación puedan mejorar la presencia en estas dos plataformas.

2. Qué es Google News: cómo funciona y qué implica para los medios

Google News es un agregador de contenido que recopila diariamente titulares de cientos de fuentes informativas y las jerarquiza y estructura en función del perfil de cada usuario (lector). Se trata de un servicio gratuito y que por tanto carece de publicidad (Lopezosa et al., 2024).

Figura 1

Portada de Google News (16/06/2025).



De este modo, Google News actúa como un difusor de contenidos procedentes de un amplio conjunto de fuentes informativas que se caracterizan por ser medios digitales seleccionados y acreditados como tales. Por lo tanto, el objetivo principal de Google News es ofrecer a los usuarios una visión panorámica y actualizada de los acontecimientos más relevantes tanto a nivel local como a nivel nacional e internacional.

Es importante destacar que Google News no produce contenido propio y por lo tanto no es una redacción periodística. Básicamente utiliza algoritmos automatizados que se encargan de rastrear y agrupar artículos bajo un mismo tema y diferentes fuentes con el objetivo de (1) ofrecer múltiples visiones sobre una misma noticia; (2) detectar tendencias que afecten a la *agenda setting*; y (3) clasificarlos por relevancia, actualidad y confianza de la fuente (Support Google, 2025a).

El resultado es una especie de portada de diario digital personalizado. De hecho, los usuarios personalizan su experiencia de diferentes formas, ya que pueden seleccionar de manera específica infinidad de temas, fuentes, ubicaciones e intereses (Support Google, 2025b). A esto hay que sumar que Google también adapta los contenidos en función del historial de búsqueda y preferencias del usuario, aunque siempre da prioridad a fuentes de información con buena reputación y que además cumplen con sus propias directrices de calidad (Support Google, 2025a; 2025b).

Aparecer en Google News puede representar tanto una oportunidad como un gran desafío para los medios de comunicación. Por un lado, facilita una enorme visibilidad global, lo que se traduce en una mayor cuota de lectores (Cordeiro, 2024), pero por otro, obliga a los medios a adaptarse a lógicas algorítmicas que pueden beneficiar a los grandes grupos de comunicación frente a los pequeños, lo que puede poner en riesgo la diversidad mediática si no se regula o gestiona adecuadamente (Cordeiro, 2024).

Aunque ya hemos avanzado que aparecer en Google News puede ser un sello de calidad para un medio y que además puede ayudar aumentar notablemente el alcance y el tráfico hacia el sitio del medio, lo cierto es que la competencia por aparecer es alta y está mediada por criterios algorítmicos (Support Google, 2025c). Ante esta circunstancia, los medios están obligados a optimizar su contenido siguiendo unas prácticas recomendadas que veremos a lo largo de este capítulo.

3. Qué es Google Discover: cómo funciona y qué implica para los medios

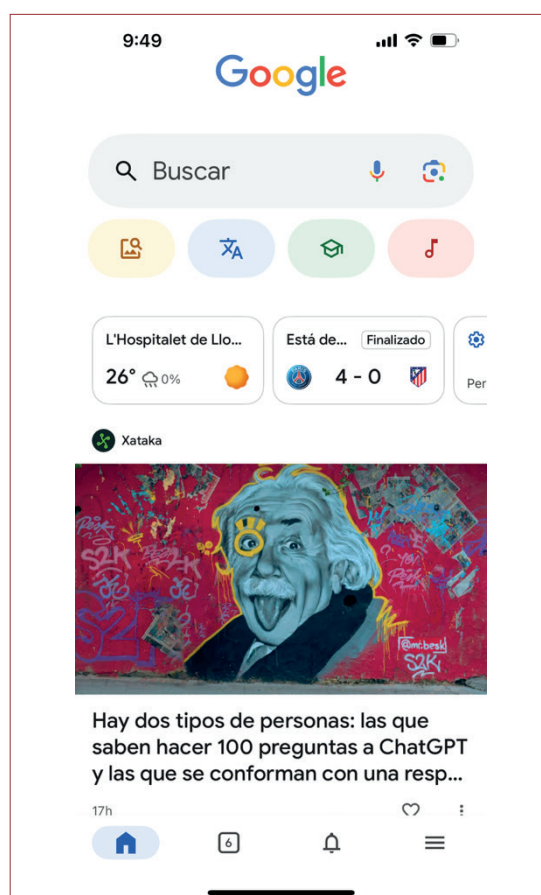
Google Discover es un servicio de Google que recomienda contenidos como noticias (Absi-Flores y León-Ferreyros, 2024), artículos, vídeos o anuncios directamente en el teléfono móvil del usuario, sin que este tenga que realizar una búsqueda activa (Lopezosa et al., 2022; Google Developers, 2020).

A diferencia del buscador tradicional, donde el usuario escribe una consulta, Discover presenta un *feed* personalizado que aparece automáticamente en dispositivos Android, dentro de la aplicación de Google y en el navegador (Support Google, 2025d). Desde mediados de 2025, también ha comenzado a estar disponible la versión de escritorio en algunos países (González, 2025).

El contenido que aparece en Google Discover es seleccionado por algoritmos, que operan en concreto con una técnica de *machine learning* llamada *reinforcement learning*, que optimiza las recomendaciones en tiempo real según como se comporta cada usuario (Strzelecki y Rizun, 2023): por ejemplo, qué noticias ha leído antes, cuánto tiempo pasa en ellas, su actividad en web y aplicaciones o su ubicación geográfica (Support Google, 2025d.; Lopezosa et al., 2022). Gracias a este funcionamiento, Discover se considera una herramienta de curación algorítmica (Lopezosa et al., 2022).

Aunque Google establece directrices para priorizar contenido útil, fiable y centrado en el usuario (Google for Developers, 2025), análisis sectoriales (Del Castillo, 2024; Polo, 2023) y académicos (Lopezosa et al., 2023) han advertido de la presencia significativa de titulares sensacionalistas y estrategias de *clickbait* en el *feed* como estrategias adaptativas agresivas a la métrica del CTR. Al respecto Google ha ejecutado recientemente actualizaciones y penalizaciones dirigidas a reducir el contenido engañoso y de baja calidad (Gonzalez, 2024), especialmente tras los *core updates* de marzo (Nelson, 2024) y agosto de 2024 (Mueller, 2024).

Figura 2.
Portada de Google Discover (16/06/2025)



3.1. Cómo funciona

Google Discover es una herramienta que muestra automáticamente una selección de contenidos personalizados a cada usuario, sin que este tenga que realizar ninguna búsqueda por palabras clave. En lugar de escribir lo que quiere encontrar (como sucede en el buscador de Google), el usuario se encuentra con un “feed” de artículos, noticias y otros formatos que Google posiciona considerando la afinidad temática, según su historial reciente de navegación, búsquedas o ubicación (Support Google, 2025d) y la evaluación algorítmica (Strzelecki y Rizun, 2023; Infobae, 2019).

La interacción del usuario con Discover es fundamentalmente pasiva aunque, a fecha de realización de este capítulo, puede indicar si un contenido le ha gustado, así como indicar si quiere ver menos contenido similar, si no le interesa la entidad, si no quiere que se muestre más contenido de la fuente así como denunciar y enviar comentarios a Google sobre un resultado en cuestión. Esto refuerza una dinámica unidireccional: es Google quien decide, en base a sus sistemas, qué mostrar, y el usuario consume sin que él u otros usuarios (como ocurre en las redes sociales) participen activamente del proceso.

Esta característica obliga a los periodistas a adaptar sus contenidos a las lógicas del *engagement* y la relevancia contextual, más que a una estrategia específica de posicionamiento en buscadores (Lopezosa et al., 2022).

Desde el punto de vista de los medios y los profesionales que gestionan sus webs (editores, periodistas o *webmasters*), Discover representa una fuente potencialmente muy valiosa de tráfico. En muchos casos, supone una parte importante de las visitas que recibe un medio digital (Lopezosa et al., 2023).

Sin embargo, lograr visibilidad en este canal no es sencillo: el sistema no permite registrarse ni solicitar ser incluido, y los criterios que determinan qué contenidos se muestran no son públicos. Aun así, hay ciertas prácticas recomendadas que ayudan a mejorar las posibilidades de aparecer: crear contenido original, utilizar imágenes destacadas de buena calidad, firmar los artículos con nombres reales y cuidar lo que Google llama EEAT: experiencia, especialización, autoridad y fiabilidad (Google Developers, 2025).

3.2. Qué implica para los medios

Google Discover tiene implicaciones directas sobre los medios de comunicación por varios motivos. En primer lugar, porque es una fuente de tráfico web altamente personalizada lo cual puede atraer una gran cantidad de lectores de calidad, y en segundo lugar, porque puede llegar a representar una proporción creciente del tráfico orgánico para medios de comunicación que ya cuentan con un buen posicionamiento en Google.

En este sentido, a diferencia del posicionamiento en buscadores, con Discover, los medios no compiten por palabras clave, sino por captar el interés del usuario con contenido relevante y visualmente atractivo. Además, se priorizan estrategias centradas en que los contenidos cuenten con imágenes destacadas, titulares atractivos y claros, información contrastada y la experiencia del autor/periodista sobre el tema, más que en las estrategias clásicas de SEO. Es decir, es más importante la capacidad de mantener al usuario interesado que la actualidad del tema.

Si bien no se puede garantizar la aparición en el *feed*, aplicar ciertas prácticas recomendadas como las que veremos más adelante pueden mejorar la visibilidad, especialmente aquellas relacionadas con imágenes, titulares y estructura técnica del contenido (Lopezosa et al., 2023; Google Developers, 2024).

Estas estrategias para competir por la atención en el *feed* (como el uso de titulares o imágenes llamativas), también pueden afectar a la calidad del contenido, aunque Google ha implementado actualizaciones recientes para reducir el contenido engañoso y de baja calidad. A esto se suman otros desafíos como la inestabilidad del tráfico, ya que los cambios en el algoritmo pueden hacer que un medio pierda audiencia de repente; o la dificultad para la planificación editorial, ya que es difícil saber con precisión qué factores prioriza el sistema.

4. Recomendaciones para periodistas y medios

Antes de ofrecer una serie de consejos a periodistas y editores sobre cómo conseguir posicionar las noticias en Google News y Google Discover, vale la pena pararnos a recapitular las principales características de ambos servicios para ilustrar sus diferencias y similitudes.

Tabla 1
Características de Google News y Google Discover.

Características	Google News	Google Discover
Enfoque principal	Noticias actuales y cobertura de eventos relevantes a nivel local, nacional y mundial.	Contenido personalizado basado en intereses y comportamiento del usuario, no solo noticias.
Personalización	Personalización mediante selección de temas, fuentes y ubicaciones; secciones como "Para ti".	Personalización automática basada en historial de búsqueda, navegación, ubicación e interacción.
Tipo de contenido	Noticias de medios reconocidos, artículos, videos y cobertura completa de eventos.	Noticias, artículos, blogs, videos, productos, anuncios o entretenimiento con formato noticioso.
Actualización del <i>feed</i>	Actualizaciones periódicas con las noticias más relevantes y recientes.	<i>Feed</i> dinámico y en tiempo real, cambia según el comportamiento e intereses del usuario.
Interacción del usuario	Seguir/dejar de seguir temas o fuentes, marcar artículos como "me gusta" o "no me gusta".	Ajustar preferencias con "me gusta", "mostrar más/menos", seguir o dejar de seguir temas en algunos países, ocultar fuentes.
Cobertura y profundidad	Ofrece "Cobertura completa" para ver múltiples perspectivas sobre un mismo tema.	No hay "cobertura completa", el contenido es más variado y sin agrupación temática.
Acceso y disponibilidad	App móvil (Android/iOS), web (news.google.com), integración con otros productos de Google.	Google app, Chrome en móvil, google.com en móvil, pantalla de inicio de algunos dispositivos y desde 2025 en escritorio en algunos países.
Algoritmo de recomendación	Basado en intereses seleccionados y preferencias explícitas del usuario.	Basado en inteligencia artificial y <i>machine learning</i> , analiza múltiples señales de comportamiento.
Propósito	Mantener informado al usuario sobre la actualidad global y local.	Sugerir contenido relevante y entretenido búsqueda previa, basado en predicción algorítmica.

A continuación, extraemos una serie de recomendaciones prácticas para mejorar la presencia en Google News y Google Discover.

4.1. Decálogo para optimizar la presencia de medios digitales en Google News

1. Solicitar la inclusión del sitio web de noticias a Google News a través de Google Publisher Center. Antes de solicitar esta inclusión será importante asegurarse de que se cumplen los requisitos técnicos y editoriales que recogen en su guía.
2. Garantizar la transparencia del medio de comunicación, es decir, publicar información sobre los periodistas, las políticas editoriales y contar con un espacio de información de contacto.
3. Publicar contenidos originales y evitar el contenido duplicado.
4. Actualizar el sitio web con frecuencia, manteniendo un ritmo regular de publicación de noticias y de cobertura de temas noticiosos.
5. Utilizar etiquetas y categorías coherentes. De ese modo se facilita la clasificación algorítmica del contenido por parte de Google.
6. Cumplir con las políticas de contenido de Google News relacionadas con el uso de titulares sensacionalistas, la desinformación y el lenguaje ofensivo.
7. Optimizar el rendimiento técnico del sitio web, lo que incluye garantizar tiempos de carga rápidos y el diseño responsivo para dispositivos móviles.
8. Incluir datos estructurados con el estándar schema.org que faciliten la interpretación del contenido por parte de Google.
9. Optimizar el archivo robots.txt para que evite bloqueos de rastreo que puedan impedir la indexación del sitio por parte de los bots de Google.
10. Apostar por una estrategia editorial clara que combine cobertura de actualidad con profundidad analítica. De este modo se facilita que el contenido pueda ser incluido en diferentes secciones de Google News, ya sea "Internacional", "Local", "Economía", "Ciencia y tecnología", "Entretenimiento", etc.

4.2. Decálogo para optimizar noticias en Google Discover

1. Publicar contenidos con formato noticioso, con temáticas basadas en tendencias en tiempo real o atemporales (*evergreen*), que aporten un valor añadido y sean susceptibles de viralizarse.
2. Redactar titulares atractivos y coherentes con el contenido, que optimicen la visibilidad en un entorno de alta competencia sin trasgredir las políticas de Google.
3. Incluir imágenes grandes (al menos 1200px de ancho), originales y de calidad. Deben ser relevantes para el contenido y estar correctamente firmadas.
4. Optimizar la ficha de autor y firmar los contenidos con autoría real que aseguren la EEAT (Experiencia, Especialización, Autoridad y Fiabilidad).
5. Aplicar buenas prácticas de SEO, datos estructurados, diseño responsive adaptado al entorno móvil y buena velocidad de carga.
6. Publicar con alta frecuencia y enriquecer los contenidos incluyendo varios formatos con vídeos, tablas, infografías, listas.
7. Medir y analizar el rendimiento en Search Console. Usar métricas clave (impresiones, clics, CTR) para ajustar estrategias y entender cómo y cuándo se genera el tráfico desde Discover. Recurrir a herramientas como Google Trends o Discover Snoop para observar las tendencias activas fuera del medio.

8. Potenciar el tráfico a través de redes sociales, portada, *hot topics*, notificaciones *push* (WhatsApp, Telegram, etc), newsletter (correo electrónico).
9. Especializarse temáticamente y construir autoridad, especialmente en áreas como salud, finanzas o ciencia (YMYL).
10. Cumplir las políticas editoriales de Google Discover: evitar prácticas penalizadas, contenido generado por IA sin supervisión humana, plagio, titulares engañosos, falta de transparencia editorial o perfiles de autor ficticios.

5. Conclusiones

En la actualidad, Discover y News son dos pilares esenciales para la mejora de la visibilidad web de los medios de comunicación online. Ambos servicios de Google han transformado, por un lado, la manera en que los usuarios acceden a la información y, por otro, cómo los medios digitales pueden distribuir sus contenidos a gran escala.

A lo largo de este capítulo hemos visto que Google News ofrece una vía estructurada y confiable para la difusión de noticias de actualidad permitiendo a los medios posicionar su marca y su rigor periodístico ante una audiencia global interesada en información veraz y relevante. Por su parte, Google Discover ha revolucionado el consumo informativo: por un lado, incorporando una lógica algorítmica que anticipa el interés del lector al decidir qué le recomienda y, por otra, disparando la audiencia potencial de la prensa digital e influyendo en las redacciones, donde el funcionamiento opaco y la dificultad para controlar la visibilidad es uno de los principales retos al usar esta herramienta.

Este trabajo ha pretendido poner en valor ambas plataformas como herramientas para conectar a los medios con audiencias amplias y segmentadas, y en consecuencia para impulsar el crecimiento de lectores y abrir nuevas oportunidades de monetización.

Si bien el aprovechamiento estratégico de Discover y News es fundamental para que los medios consoliden su presencia digital, no queremos concluir sin señalar el desafío de aplicar estas estrategias con una mirada crítica hacia la tensión que generan las exigencias de la visibilidad algorítmica con los estándares de calidad informativa.

En este escenario, el reto consiste en aplicar estas estrategias para optimizar la presencia digital de los medios poniendo al lector en el centro, y sin renunciar a los compromisos sociales del periodismo.

Así pues, corresponde a periodistas y editores adaptarse de forma ética a los constantes cambios en los algoritmos, velando siempre por la integridad y los principios del periodismo. Será de esta forma en la que los medios digitales podrán sostenerse en estos dos servicios de Google sin renunciar a su esencia.

6. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUCOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

7. Referencias

- Absi Flores, L. M., & León Ferreyros, A. (2024). Posicionamiento de contenido digital en Google Discover: el caso del diario peruano El Comercio. *Estudios sobre el Mensaje Periodístico*, 30(3), 625-634. <https://doi.org/10.5209/emp.96024>
- Cordeiro, D. F., Lopezosa, C., Guallar, J., & Vállez, M. (2024). Análisis de la cobertura de Google News: un estudio comparativo de Brasil, Colombia, México, Portugal y España. *Contratexto*, (42), 177-208. <https://doi.org/10.26439/contratexto2024.n42.7212>
- Cordeiro, D. F. (2024). Perspectivas en contraste: análisis comparativo cuantitativo España y Brasil de la cobertura del conflicto israelí-palestino en Google News. *Documentación de las Ciencias de la información*, 47, 15-25. <https://revistas.ucm.es/index.php/DCIN/article/view/92187>
- Del Castillo, C. (2024). El 'periodismo de algoritmo' ataca de nuevo: así es como titulares engañosos y supermercados inundan los medios. *elDiario.es*, 29 October. https://www.eldiario.es/tecnologia/periodismo-algoritmo-ataca-nuevo-titulares-enganosos-supermercados-inundan-medios_1_11734185.html
- González, D. (2024). Google Discover tumba tu tráfico si titulas así. Te envía un mensaje como este. Es su nueva cruzada contra el clickbait. *Red de Periodistas*. <https://www.reddeperiodistas.com/google-discover-desplomara-tu-trafico-si-titulas-asi-y-te-envia-un-mensaje-como-este-en-su-nueva-cruzada-contra-el-clickbait>
- González, D. (2025) Google Discover en escritorio llega en estos días, pero viene con truco. *Red de Periodistas*. <https://www.reddeperiodistas.com/google-discover-en-escritorio-llega-en-estos-dias-pero-viene-con-truco>
- Google Developers. (2020). Aparecer en Discover. *Google Developers*. <https://developers.google.com/search/docs/appearance/google-discover?hl=es>
- Google Developers. (2025a). Crear contenido útil, fiable y centrado en las personas. <https://developers.google.com/search/docs/fundamentals/creating-helpful-content>
- Google Developers. (2025b). Experience, Expertise, Authoritativeness, and Trust (E-E-A-T). In *General Guidelines* (pp. 26-28). <https://guidelines.raterhub.com/searchqualityevaluatorguidelines.pdf>
- Guallar, J. (2015). Prensa digital en 2013-2014. *Anuario ThinkEPI*, 9, 153-160. <https://doi.org/10.3145/thinkepi.2015.37>
- Infobae (2019). Cómo usar Google Discover para ver las noticias que te interesan en tu celular. *Infobae*. <https://www.infobae.com/america/tecno/2019/09/02/como-usar-google-discover-para-ver-las-noticias-que-te-interesan-en-tu-celular>
- Iglesias-García, M., & Codina, L. (2016). Los cybermedios y la importancia estratégica del posicionamiento en buscadores (SEO). *Opción*, 32(9), 929-944.
- Lopezosa, C., Codina, L., & Rovira, C. (2019). *Visibilidad web de portales de televisión y radio en España: ¿qué medios llevan a cabo un mejor posicionamiento en buscadores*. Barcelona: Universitat Pompeu Fabra. Departament de Comunicació, Serie DigiDoc-EPI, 4. <http://hdl.handle.net/10230/36234>

Lopezosa, C., Giomelakis, D., Pedrosa, L. & Codina, L. (2023). Google Discover: uses, applications and challenges in the digital journalism of Spain, Brazil and Greece". *Online Information Review*. <https://doi.org/10.1108/OIR10-2022-0574>

Lopezosa, C., Guallar, J., & Santos-Hermosa, G. (2022). Google Discover: entre la recuperación de información y la curación algorítmica. *SCIRE. Representación y organización del conocimiento*, 28(2), 13-22. <https://doi.org/10.54886/scire.v28i2.4796>

Lopezosa, C., Vállez, M., & Guallar, J. (2024). The vision of Google News from the academy: scoping review. *Doxa Comunicación*, 38. <https://doi.org/10.31921/doxacom.n38a1891>

Mueller, J. (2024). Qué debes saber sobre nuestra actualización principal de agosto del 2024 | Blog central de la Búsqueda de Google. *Google for Developers*. <https://developers.google.com/search/blog/2024/08/august-2024-core-update?hl=es>

Nelson, C. (2024, marzo 5). Qué deben saber los creadores web sobre la actualización principal de marzo del 2024 y las nuevas políticas de spam. *Google Developers*. <https://developers.google.com/search/blog/2024/03/core-update-spam-policies?hl=es>

Polo, S. (2023). Google Discover resucita la edad dorada del 'clickbait': "Prioriza lo más leído y eso es una ventana al sensacionalismo". *El Mundo*, May 16. <https://www.elmundo.es/television/medios/2023/05/16/6461f0fae4d4d8907a8b459c.html>

Salaverría, R., & García-Avilés, J. A. (2008). La convergencia tecnológica en los medios de comunicación: retos para el periodismo. *Trípodos*, (23), 31-47.

Strzelecki, S., & Rizun, M. (2023). Exploring the Impact of Google Discover on Users and Publishers: A Data-driven Study. *31° International Conference of information System Development*. <https://doi.org/10.62036/ISD.2023.28>

Support Google (2025a). *Empezar a utilizar Google News*. <https://support.google.com/googlenews/answer/9005669?hl=es&co=GENIE.Platform%3DAndroid>

Support Google (2025b). *Cómo empezar a usar Google*. <https://support.google.com/googlenews/answer/9005669?hl=es-419&co=GENIE.Platform%3DAndroid>

Support Google (2025c) *News content across Google*. <https://support.google.com/news/publisher-center/answer/9607025?hl=en>

Support Google (2025d). *Personalizar el contenido de Discover*. <https://support.google.com/websearch/answer/2819496?hl=es&co=GENIE.Platform%3DAndroid>

Curación de contenidos en *newsletters* independientes y de medios de comunicación. Buenas prácticas y recomendaciones

Pere Franch

Universitat Ramon Llull, España

<https://orcid.org/0000-0002-4180-1168>

Javier Guallar

Universitat de Barcelona, España

<https://orcid.org/0000-0002-8601-3990>

Franch, P., & Guallar, J. (2025). Curación de contenidos en *newsletters* independientes y de medios de comunicación. Buenas prácticas y recomendaciones. En J. Guallar, M. Váñez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 26-38). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicecom.03.esp>

Resumen

El reciente auge de las newsletters da prueba de su eficacia como producto digital de información. Una de sus principales aportaciones consiste en que, gracias a la curación de contenidos, las newsletters permiten a los periodistas conectar con usuarios interesados en un área temática determinada y proporcionarles información específica o especializada sobre su ámbito de interés. En este capítulo se presentan las principales conclusiones de seis estudios centrados en la curación de contenidos de boletines publicados por medios tradicionales y digitales, así como por periodistas independientes. Se ofrece, asimismo, un compendio de buenas prácticas y recomendaciones para mejorar la calidad de las newsletters periodísticas. Los boletines de calidad ayudan a los usuarios a orientarse ante la abrumadora cantidad de información que ofrece Internet, dispersa, inmediata, fragmentada y descontextualizada. Al mismo tiempo, las newsletters permiten a los medios estrechar lazos con su audiencia, fidelizar a los usuarios y aumentar su base de suscriptores. Para mejorar su calidad, recomendamos que las newsletters tengan autoría identificada y abundante contenido curado procedente de fuentes variadas en cuanto a origen (interno y externo), rango temporal y morfología. También se aconseja que haya variación en el uso de técnicas de curación y en las funciones de los enlaces.

Palabras clave

Curación de contenidos; Newsletters; Boletines independientes; Periodismo digital; Medios de comunicación; Fuentes de información; COVID-19; Coronavirus.

I. Introducción

La *newsletter* es un producto de comunicación antiguo y a la vez novedoso que ha experimentado un importante auge en los últimos años. Es probablemente el medio más extendido y reconocido de curación de contenidos digitales y ha sido objeto de varias investigaciones recientes.

Este capítulo resume las principales conclusiones de los estudios que sobre este tema se han realizado en el marco del proyecto de investigación CUVICOM (<https://www.ub.edu/cuviceom/>), en concreto sobre la curación de contenidos de *newsletters* de medios de comunicación y de editores independientes. También propone recomendaciones y buenas prácticas para la elaboración de *newsletters* y su curación.

Nuestra investigación en este campo ha dado lugar, desde 2021, a seis artículos que se centran en los siguientes aspectos o temas:

- *Newsletters* informativas de medios españoles tradicionales y nativos digitales (Guallar et al., 2021a).
- *Newsletters* independientes españolas (Cascón-Katchadourian et al., 2022).
- *Newsletters* especializadas en la pandemia de SARS-CoV-2 (Guallar et al., 2022).
- *Newsletters* independientes latinoamericanas (Lopezosa et al., 2023).
- La visión de los editores independientes (Lopezosa et al., 2024).
- La oferta de boletines de *The New York Times* (Guallar et al., 2025).

Estos estudios combinan dos metodologías: entrevistas semiestructuradas a editores, y análisis de muestras de newsletters utilizando el método CAS (*Curation Analysis System*) (Tabla 1). Este método, descrito inicialmente en Guallar et al. (2021b), observa dos dimensiones –contenido y curación–, fijándose en los siguientes parámetros: del contenido se estudia la cantidad, el rango temporal, el origen (propio o externo) y las fuentes (según organización y morfología); de la curación se observa la autoría, la técnica de *sense-making* y la función de los enlaces.

2. Principales conclusiones de la investigación

2.1. Newsletters de periódicos españoles

El primer estudio, centrado en los boletines publicados por medios *legacy* y nativos digitales españoles (Guallar et al., 2021a), concluyó que, en el momento en que se recogieron los datos (en 2020), todos los medios seguían priorizando los enlaces a sus propios contenidos frente a los contenidos externos. Ello deriva en una oferta poco variada de fuentes de información y una presencia sorprendentemente baja de contenidos de las redes sociales, a pesar de que éstas se consideran un canal propicio para la curación de noticias.

En cuanto a las características de la curación, la investigación inicial estableció dos grandes tipos de curación de newsletters: intelectual y automatizada. La intelectual la realiza un autor profesional que se identifica ante el usuario. Utiliza una o varias técnicas de *sense-making* e incorpora enlaces a diversas fuentes de información. En cambio, en la automatizada, la aportación periodística es menor (o inexistente) y no se identifica la autoría; se basa en una serie de enlaces a los que no se aplica ninguna técnica de *sense-making* y no hay variedad en el uso informativo de los enlaces. Se puede relacionar esta dicotomía con la aportación posterior de Andringa (2022), que distingue dos modelos de boletines: el funcional y el relacional. El modelo funcional tiene como principal o único objetivo proporcionar a la audiencia información útil y relevante; suele tener una audiencia muy específica y no fomenta una relación directa entre el autor y los suscriptores. La mayoría de los boletines automatizados siguen este modelo. El modelo relacional, en cambio, busca construir relaciones con la audiencia a través de contenidos personales y prácticas de *engagement*. Estas newsletters tienen por objetivo conseguir un mayor grado de interacción con su audiencia.

El estudio también reveló que la mayoría de las newsletters publican una cantidad moderada de contenidos, a menudo en torno a diez piezas por boletín. El contenido curado es principalmente información de actualidad publicada en las últimas horas, casi exclusivamente de medios de comunicación, sobre todo digitales. Otras fuentes, como las redes sociales o los sitios web oficiales, tienen una presencia muy limitada.

2.2. Newsletters independientes españolas

El estudio de los boletines independientes españoles (Cascón-Katchadourian et al., 2022) arrojó las siguientes conclusiones: como principales características destacan que son gratuitos, creados por autores identificados, y suelen publicarse semanalmente o incluso con mayor frecuencia; cubren una amplia gama de temas, desde la actualidad general hasta la información especializada en áreas y nichos específicos. Mantienen un alto nivel de fidelidad

de la audiencia, con tasas de apertura superiores al 50% en muchos casos. Y algunos tienen productos relacionados, como podcasts.

En cuanto a la calidad de la curación de sus contenidos, todas las newsletters estudiadas en esta investigación recurren a la curación intelectual o profesional, más que a la automatizada. Esta es una de las principales diferencias con respecto al estudio anterior.

También cabe destacar que existe diversidad de rangos temporales en los contenidos curados, con un predominio de contenidos recientes (últimos días), lo que está claramente relacionado con la gran abundancia de boletines de periodicidad semanal, así como una presencia regular de contenidos retrospectivos y actuales (últimas horas).

Los boletines independientes españoles presentan también un uso variado de técnicas de *sense-making*, combinando mayoritariamente dos o tres de ellas por boletín. La técnica "resumir" está presente en todos los casos, y "citar" y "comentar" se utilizan con frecuencia. Entre la variedad de usos de los enlaces, destacan las funciones "describir", "sin modificar" y "citar fuente de contenido", por este orden.

En cuanto a las fuentes, los boletines independientes enlazan mayoritariamente con contenidos externos, lo que revela dos tendencias emergentes en su uso: en primer lugar, algunas newsletters citan predominantemente sitios web de medios de comunicación, dejando poco espacio para otras fuentes; en segundo lugar, otras citan con más frecuencia medios sociales (redes sociales y blogs).

Este estudio incluyó entrevistas semiestructuradas con los editores de los diez boletines analizados. A partir de sus respuestas, se llegó a la conclusión de que su principal preocupación es aumentar el número de suscriptores. La mayoría se muestra reticente a ofrecer un producto de pago y a la vez afirma que el periodo de auge de los boletines podría llevar a la saturación y a un posterior reajuste. Los editores también creen que el futuro de la edición de contenidos pasa por una mayor especialización y la ampliación a distintos formatos.

2.3. Newsletters sobre coronavirus de periódicos de cinco países

El trabajo sobre newsletters especializadas en noticias de Covid-19 de medios de comunicación de Francia, Alemania, España, Reino Unido y Estados Unidos, concluyó que estos boletines comparten la mayoría de las características observadas en estudios previos (Rojas-Torrijos y González-Alba, 2018; Guallar et al., 2021a), si bien se observa una ligera mejora en todos los indicadores observados en comparación con las newsletters generalistas: se aprecia una mayor presencia de autoría identificada, diferentes rangos temporales, mayor variedad de fuentes y un abanico más amplio de técnicas de *sense-making*.

En cuanto al uso de fuentes, los medios *legacy* ofrecen poco contenido de fuentes externas. Por el contrario, los *pure digital* son los que ofrecen más contenido externo, a veces incluso más enlaces a webs externas que a la propia. También cabe destacar que todos los medios analizados ofrecen enlaces a sitios web de medios de comunicación y, en mucha menor medida, a fuentes oficiales y corporativas (a menudo del ámbito sanitario) y a fuentes ciudadanas.

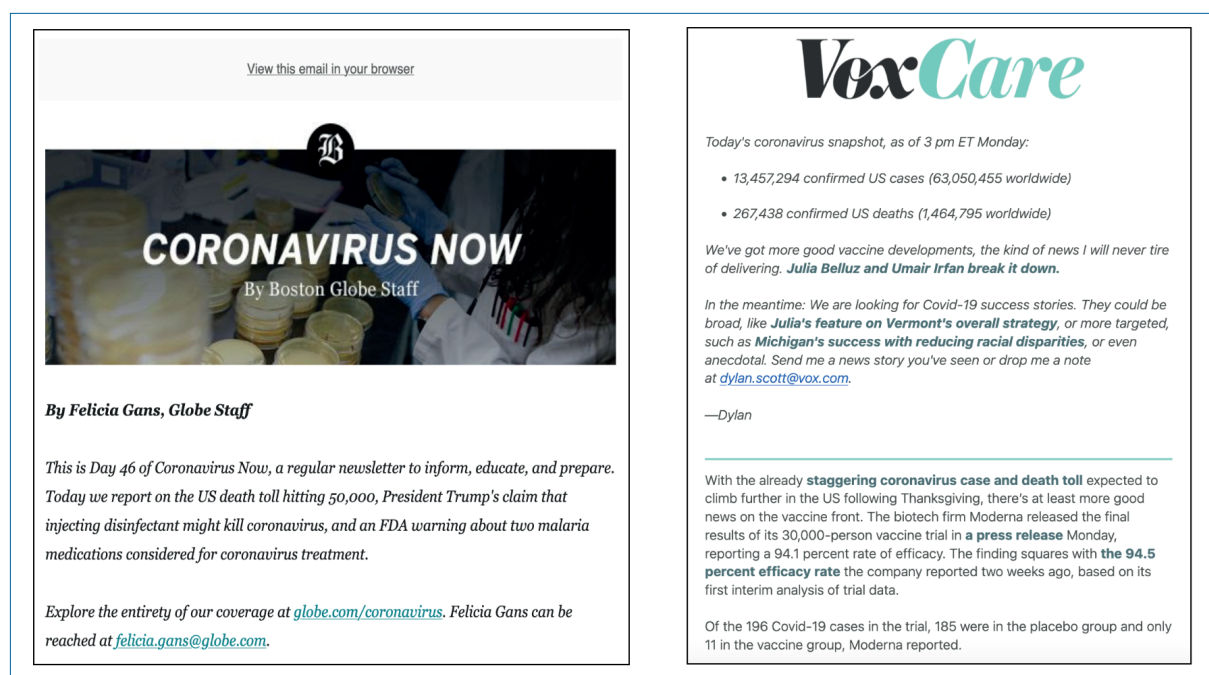
El origen del contenido curado y el formato es más diverso en medios digitales como *elDiario.es*, *Politico*, *Vox* y *BuzzFeed*, que incluyen no sólo páginas web, sino también blogs y redes sociales. Los enlaces a otros medios suelen ser a grandes empresas internacionales

de referencia, como periódicos, cadenas de televisión y agencias de noticias. Los enlaces a contenidos especializados son, en la mayoría de los casos, a organizaciones como la ONU o la OMS, o a sitios web de expertos en salud y medicina. De estos resultados concluimos que los medios de comunicación tradicionales que han hecho la transición al entorno digital utilizan principalmente esta tecnología para aumentar el tráfico web. En cambio, los *pure digital* utilizan los hipervínculos y la web para proporcionar al usuario la oportunidad de acceder a fuentes diversas y variadas, incluidas las redes sociales.

Las newsletters con contenido automatizado y una gran cantidad de enlaces a su propia página web pretenden principalmente aumentar el tráfico sin personalizar el contenido del boletín. Esta tendencia se ha observado más en los medios tradicionales que en los digitales. Aunque su finalidad es informativa, el principal objetivo es fomentar las suscripciones, las visitas al sitio web, los micropagos y las descargas de contenidos para ganar y retener usuarios (Silva-Rodríguez, 2021).

Figura 1

Ejemplos de newsletters informativas sobre la pandemia de SARS-CoV-2 de un medio tradicional y uno pure digital: The Boston Globe y Vox, respectivamente.



2.4. Newsletters independientes de América Latina

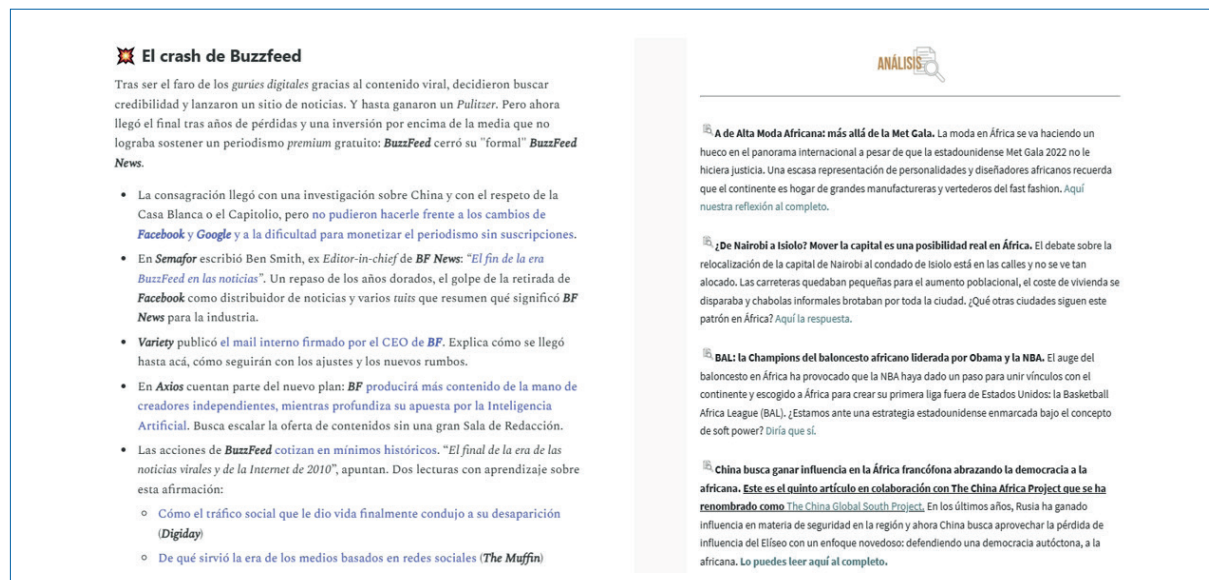
Esta investigación consistió en un análisis de las principales características y la calidad de la curación de contenidos de 12 newsletters independientes de referencia en América Latina, e identificó patrones interesantes. Se detectó que publican una cantidad significativa de contenido curado por newsletter (entre 11 y 47 elementos), y –en comparación con estudios anteriores– un mayor uso de contenido atemporal, una mayor presencia de contenido de fuentes externas, y un predominio de técnicas de *sense-making* como “comentar” y “resumir”.

Todas las newsletters analizadas ofrecen un alto nivel de calidad, como demuestra su puntuación en comparación con otros boletines analizados en estudios anteriores. En todos los casos, la curación es elaborada, “intelectual” o de autor, nunca automatizada.

El estudio revela también tres grupos principales de fuentes en la curación de boletines independientes latinoamericanos: redes sociales (principalmente Twitter, YouTube y LinkedIn); medios de comunicación de todo el mundo o de la región latinoamericana (por ejemplo, *The Guardian*, *El País*, *O Globo*, *Folha*, *El Universal* y *The New York Times*); y los propios boletines.

Figura 2

Ejemplos de curación de dos newsletters independientes, de España y América Latina: *Lunes* y *África Mundi*.



2.5. Entrevistas con editores

Este artículo, basado en entrevistas con editores independientes de newsletters, revela varios puntos clave:

Los editores se inspiran en contenidos online, herramientas para organizar contenidos y métodos de curación, todo lo cual influye en la estructura y el diseño de las newsletters. Ello concuerda con otros estudios que subrayan la importancia de la curación de contenidos en el proceso de contar historias.

Los editores perciben un auge positivo de las newsletters, hecho que atribuyen a aspectos innovadores estratégicos como la aparición de comunidades nicho, el valor de los contenidos originales, la diversidad de los modelos de monetización, la integración multimedia y la mejora y expansión continuas en este ámbito de las newsletters. Los editores también destacan el atractivo de los boletines al proporcionar información personalizada y facilitar la producción por parte del usuario, así como su capacidad para escapar al control algorítmico.

Las newsletters también destacan por su capacidad para construir imagen de marca, concretamente por su importancia estratégica a la hora de establecer conexiones directas con los usuarios, atraer suscriptores y complementar a los medios tradicionales. Aunque algunos entrevistados anticipan que los boletines seguirán siendo productos de nicho, su capacidad única para ofrecer una comunicación cercana, incluso "íntima", fomenta la lealtad a la marca personal del editor. Los editores también reconocen que las newsletters son herramientas eficaces para captar suscriptores y que se han convertido en competidoras potenciales de los medios tradicionales.

Por último, los entrevistados valoran el papel crucial que significa la selección de contenidos que ofrecen las newsletters frente a la gran saturación de información que hay en la red. Reconocen el impacto permanente del progreso tecnológico en este ámbito y son partidarios de la colaboración entre curadores para ofrecer recomendaciones de contenidos y fomentar una sólida y acreditada identidad digital.

2.6. Las newsletters de *The New York Times*

El estudio de las newsletters de *The New York Times* concluye que la mayoría de ellas (73%) son de acceso abierto. Este hallazgo refuerza la idea de que los boletines tienen como objetivo ampliar el contenido del NYT, aumentar su impacto y hacer crecer la audiencia del medio y, en consecuencia, las suscripciones de pago. Una alta frecuencia y regularidad de publicación son claves para la fidelidad de los usuarios, como se señala en la literatura pertinente (Isaac, 2019; Kim et al., 2023; Mondéjar, 2024; Santos-Silva y Granado, 2019; Seely y Spillman, 2021).

La mayoría de los boletines publican opiniones o noticias, segmentadas por tiempo, geografía, idioma o tema. Pocos siguen un modelo relacional o híbrido, ya que la mayoría son funcionales y se limitan a proporcionar información sin interacción directa con el usuario. Aunque los autores varían, las newsletters sólo para suscriptores suelen tener un estilo más personal, con periodistas que ofrecen contenido experto de forma directa e interactiva.

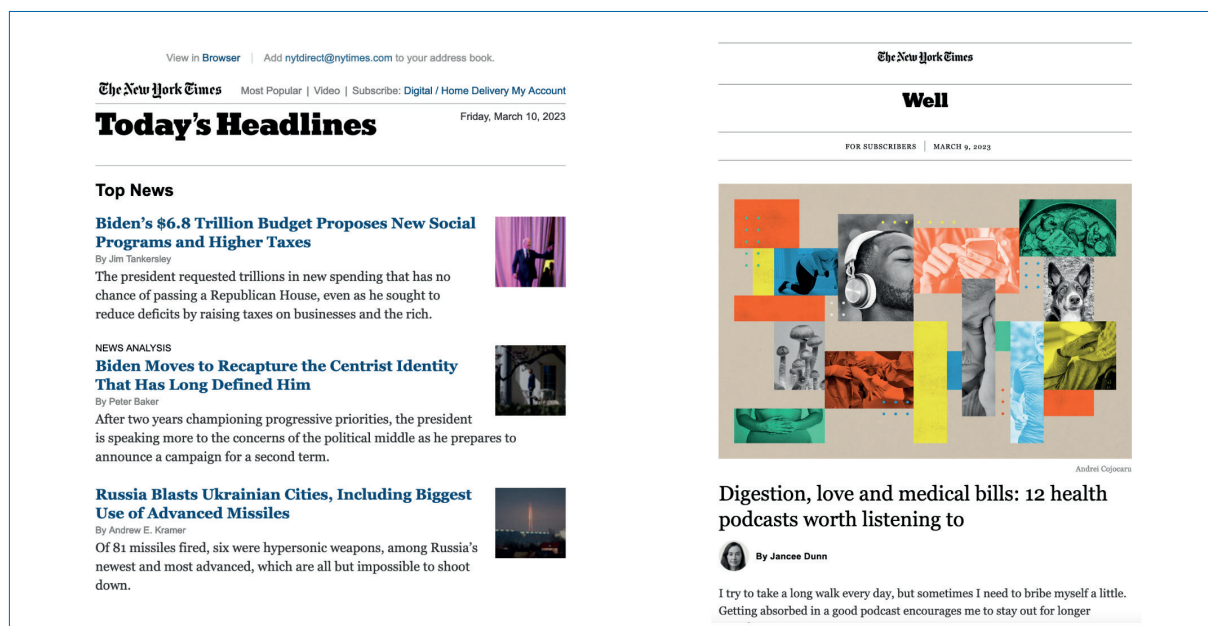
La cantidad de contenidos curados es menor que en estudios anteriores, especialmente si lo comparamos con el de las newsletters independientes. Hay más contenido atemporal y retrospectivo, hecho directamente relacionado con el carácter informativo y contextual que tienen las newsletters del NYT. Hay muy pocos contenidos en tiempo real, reservados al seguimiento de acontecimientos importantes e inesperados (*breaking news*). La mayoría de los boletines sólo ofrecen contenidos de *The New York Times*, lo que se traduce en un aumento del tráfico y una mejora en los rankings de audiencia. Esto plantea la cuestión de si existe una tendencia endogámica o un afán corporativista en los grandes medios de comunicación. De hecho, el contenido procedente de medios predomina en los enlaces publicados; sin embargo, las newsletters de noticias con autor específico ofrecen en general más fuentes externas y casi todos son sólo para suscriptores. En la mayoría de los casos estos contenidos externos son de calidad y aportan mayor valor añadido a los usuarios.

En cuanto a la morfología de los contenidos curados, predominan los sitios web. También hay poca variación en las técnicas de curación, siendo “resumir” la más frecuente. En cuanto a la función de los enlaces, “sin modificar” y “describir” son las dos más utilizadas, seguidas de “call to action”.

En general, la calidad de las newsletters del NYT es inferior a la observada en otros estudios. Las newsletters independientes latinoamericanas obtienen las puntuaciones más altas (Lopezosa et al., 2023), y resultados similares arroja el estudio de newsletters de medios españoles (Guallar et al., 2021a). Por otro lado, los boletines del NYT con autoría identificada obtienen puntuaciones más altas, lo que corrobora lo observado en estudios previos sobre newsletters independientes españolas y latinoamericanas (Cascón-Katchadourian et al., 2022; Lopezosa et al., 2023), en los que se constataba que las newsletters de autor ofrecen al usuario contenidos de mayor variación y calidad.

Figura 3

Ejemplos de newsletters de The New York Times. A la izquierda, un ejemplo de newsletter automatizada, sin autoría y modelo funcional. A la derecha, un ejemplo de newsletter relacional con autora identificada y contenido específico, en este caso, salud y bienestar.



3. Buenas prácticas

Los estudios presentados confirman que las newsletters se están consolidando como un medio que facilita y fomenta la curación de contenidos, lo que puede redundar en una mejora de la calidad del periodismo. Constituyen un producto personalizado acorde con los intereses específicos de cada usuario, lo que conlleva que las newsletters con contenidos curados de calidad sean una alternativa eficaz a la saturada oferta de material que ofrece la red. La mayoría de las newsletters estudiadas consisten en productos con contenidos específicos, organizados temáticamente, con información jerarquizada y limitada en el tiempo, presentada en un formato concreto y de acuerdo con las preferencias del suscriptor, quien las recibe en la privacidad de su correo electrónico (Carr, 2014; Pell, 2021; Rourke, 2021; Suárez, 2020).

Las newsletters de calidad son un recurso utilizado por los medios tradicionales para estrechar lazos con su audiencia, fidelizar a los usuarios, fomentar su *engagement* con el medio y, en último extremo, aumentar el número de suscriptores (Isaac, 2019; Mondéjar, 2024; Santos-Silva y Granado, 2019). Su objetivo final es ofrecer valor añadido a la información pura y dura —lo que se considera *hard news*—, con lo que el autor del boletín se convierte en un agente personal de los usuarios, establece una relación cercana (Andringa, 2022) o íntima (Mondéjar, 2024) con ellos, y se convierte en gestor, coproductor y distribuidor de información, al tiempo que adopta el papel de creador de redes (Upmanyu, 2024; Weder et al. 2023).

A continuación se exponen algunas de las buenas prácticas para crear newsletters informativas de calidad:

El método CAS (tabla 1) evalúa la calidad del boletín en función de la presencia o ausencia de características específicas, por lo que es una buena práctica ofrecer la mayor variedad posible en todos los parámetros analizados en todas las categorías. Esto incluye generosidad en la cantidad de contenidos curados, variedad en el rango temporal y equilibrio entre contenido

original y externo. También se valoran los enlaces a contenidos curados de calidad procedentes de diversas fuentes ajenas al medio. La variedad de fuentes se valora cuando éstas no son sólo medios de comunicación, sino también fuentes oficiales, corporativas o ciudadanas. Y se valora positivamente el tipo de fuente cuando no sólo se ofrecen sitios web sino también blogs o contenidos de redes sociales.

En términos de curación, tener un autor identificado se considera un signo de calidad para las newsletters. Los estudios aquí presentados establecen una doble dicotomía estrechamente relacionada entre sí. En primer lugar, se distingue entre newsletters con autor –autoría intelectual– y newsletters automatizadas. La calidad del contenido y de la curación suele considerarse mayor en las primeras, ya que presentan el valor añadido que supone el trabajo intelectual del autor. La curación automatizada de contenidos, en cambio, sólo proporciona enlaces (normalmente de medios de comunicación o de la propia organización) a los usuarios. En esta línea, Andringa (2022) introduce una taxonomía similar y distingue dos modelos de boletines: el relacional fomenta una relación directa entre el autor y el consumidor, mientras que el funcional puede ser el resultado de un buen trabajo del autor y puede proporcionar enlaces de calidad, pero sin establecer esta proximidad o complicidad con el usuario.

Así pues, una práctica recomendada para crear newsletters de calidad es ofrecer newsletters de autor que sigan el modelo relacional, con contenidos específicamente alineados con los intereses concretos del nicho del público objetivo.

En cuanto a las otras dos categorías de curación analizadas, se valora la variedad de técnicas de *sense-making*, aunque la técnica más utilizada es “resumir”. Por tanto, es una buena práctica presentar el contenido del boletín en un texto claro y conciso para que el lector sepa qué ofrece el producto. Por último, la variedad en las funciones también es un signo de calidad en el texto de los enlaces. Cabe señalar que la función más común en la mayoría de los enlaces analizados es “describir”. En consecuencia, una práctica recomendada es presentar el texto de los enlaces de forma que permita al lector identificar fácilmente el contenido que encontrará una vez acceda al sitio web enlazado.

El uso de fuentes internas o externas merece un análisis más detallado. Diversos estudios han demostrado que los medios tradicionales (*legacy*) suelen ofrecer más contenido propio que externo y ofrecen mucho menos enlaces a contenido externo que los medios digitales (*pure digital*). Las newsletters sobre el coronavirus de *The Boston Globe* son paradigmáticas en este sentido. Ello sugiere que su objetivo puede ser generar tráfico web más que ofrecer un servicio de calidad a los lectores. En cualquier caso, esto muestra una actitud autorreferencial típica del efecto de cámara de eco observado ante la proliferación de las redes sociales (Teren y Borge, 2021). Por el contrario, los medios nativos digitales y los boletines independientes ofrecen más contenidos externos. Algunos incluso no tienen problema en ofrecer enlaces a medios de la competencia, lo que se valora positivamente, ya que permite a los usuarios ampliar sus conocimientos y fuentes de información sobre los temas que les interesan.

Tabla 1

Dimensiones, parámetros e indicadores observados por el método CAS.

Dimensión	Parámetro	Indicador
A. Contenido	A1. Número de contenidos	Cantidad
		Información retrospectiva o atemporal
	A2. Rango temporal	Información reciente
		Información actual
		Información en tiempo real
	A3. Origen	Contenido propio
		Contenido externo
	A4. Fuentes por tipo de organización	Fuentes oficiales
		Fuentes corporativas
		Medios
		Ciudadanos
	A5. Fuentes según morfología	Páginas web
		Blogs
		Redes sociales
		Fuentes secundarias
B. Curación	B1. Autoría	Autoría
	B2. Técnica de <i>sense-making</i>	Resumir
		Comentar
		Citar
		<i>Storyboarding</i>
	B3. Función del enlace	Sin modificar
		Describir
		Contextualizar
		Interpretar
		Citar fuente
		Citar al autor
		Llamar a la acción

Fuente: Elaboración propia a partir de Guallar et al. (2021b).

4. Recomendaciones

Por último, en los siguientes párrafos se presentan algunas implicaciones prácticas en forma de recomendaciones respecto a los medios de comunicación, las *newsletters* y los periodistas curadores de contenidos:

4.1. En relación con los medios digitales

- Centrarse en la selección de contenidos de calidad. Los medios de comunicación podrían mejorar su calidad ofreciendo *newsletters* con contenidos curados de alta calidad y enlaces externos de alto valor añadido.

- (b) Boletines de autor para acercarse al público. Las *newsletters* con autoría identificada y un estilo personal puntúan más alto en calidad de curación, por lo que los medios podrían promoverlos para conectar mejor con su audiencia.
- (c) Enfoque temático y segmentación. Se observa que publicar boletines con segmentación temática variada permite llegar a audiencias específicas, por lo que se recomienda a los medios seguir esta estrategia.
- (d) Estrategia de tráfico web. Una elevada presencia de enlaces internos, principalmente en los medios tradicionales, sugiere que se sigue una estrategia para potenciar el tráfico web, por lo que los medios podrían utilizar las *newsletters* para dirigir a los usuarios a sus sitios web y aumentar así el número de suscriptores.

4.2. En relación con las *newsletters*

- (a) La importancia del contenido curado con valor añadido. Las *newsletters* de calidad ofrecen contenidos curados que aportan contexto y análisis, en la mayoría de los casos mediante la técnica de “resumir”.
- (b) Incluir enlaces externos. Aunque los enlaces internos son importantes para el tráfico web, los enlaces externos de fuentes fiables y acreditadas añaden valor y aumentan la credibilidad.
- (c) Diversidad de técnicas de curación. La mayoría de las *newsletters* estudiadas (excepto algunas independientes) se basan en una única técnica, por lo que utilizar una variedad de técnicas, como “comentar”, “citar” o “storyboarding”, podría enriquecer la experiencia del usuario.
- (d) Modelo híbrido para una mayor conexión. Combinar el modelo funcional de proporcionar información muy acreditada con un enfoque relacional que fomente la interacción con la audiencia puede mejorar la calidad de las *newsletters* y reforzar las relaciones con los usuarios.

4.3. Sobre el uso de la curación de contenidos en periodismo

- (a) La curación puede ser una herramienta para mejorar la calidad periodística. Las *newsletters* que utilizan eficazmente la curación de contenidos ofrecen a los usuarios una experiencia informativa más completa y enriquecedora.
- (b) El periodista curador aparece como gestor y coproductor de información. La práctica de la curación de contenidos transforma el papel del periodista, que se convierte en gestor, coproductor, distribuidor de información y también creador de redes.
- (c) Una oportunidad alternativa para los periodistas *freelance*: las *newsletters* pueden ayudar a crearse una personalidad como experto en un nicho temático, publicar textos alternativos a los medios tradicionales y encontrar una ocupación provisional a falta de otras oportunidades profesionales (Zilberstein, 2022).

5. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

6. Referencias

- Andringa, P. (2022). *Email Newsletters and the Changing Journalist-Audience Relationship*. PhD diss., The Oxford Internet Institute at the University of Oxford. <https://peterandringa.com/oii-thesis.pdf>
- Carr, D. (2014, June 29). For Email Newsletters, A Death Greatly Exaggerated. *The New York Times*. <https://www.nytimes.com/2014/06/30/business/media/for-email-a-death-greatly-exaggerated.html>
- Cascón-Katchadourian, J.-D., Boté-Vericad, J.-J., & Guallar, J. (2022). Curación de información con marca de autor: análisis de newsletters independientes. *Ibersid: Revista de Sistemas de información y documentación*, 16(1), 41-53. <https://doi.org/10.54886/ibersid.v16i1.4799>
- Guallar, J., Anton, L., Pedraza-Jiménez, R., & Pérez-Montoro, M. (2021a). Curación de noticias en el correo electrónico: análisis de newsletters periodísticas españolas. *Revista Latina de Comunicación Social*, (79), 47-64. <https://doi.org/10.4185/RLCS-2020-1488>
- Guallar, J., Franch, P., Boté-Vericad, J.-J., & Anton, L. (2022). How do legacy and digital media curate coronavirus content. An assessment of newsletters from the USA and four European countries. *Profesional de la Información*, 31(3), e310318. <https://doi.org/10.3145/epi.2022.may.18>
- Guallar, J., Franch, P., Cascón-Katchadourian, J., & Boté-Vericad, J. J. (2025). The Rise of Curated Newsletters in Media: A Case Study of the New York Times. *Journalism Practice*, 1-22. <https://doi.org/10.1080/17512786.2025.2457332>
- Guallar, J., Pedraza-Jiménez, R., Pérez-Montoro, M., & Anton, L. (2021b). Curación de contenidos en periodismo. Indicadores y buenas prácticas. *Revista española de documentación científica*, 44(2). <https://doi.org/10.3989/redc.2021.2.1742>
- Isaac, M. (2019, March 19). The New Social Network That Isn't New At All. *The New York Times*. <https://nyti.ms/2AsoNcn>
- Kim, S. J., Kim, H., Choi, J. R., & Malthouse, E. C. (2023). Newly Subscribed! Effects of e-Mail Newsletters on News-Reading Habit and Subscriber Retention During Onboarding: Evidence from Clickstream and Subscription Data. *Journal of Media Economics*, 35(3-4), 87-107. <https://doi.org/10.1080/08997764.2024.2333368>
- Lopezosa, C., Cascón-Katchadourian, J., & Guallar, J. (2023). Newsletters latinoamericanas de curación: estudio de casos y visión experta. *TransInformação*, 35. <https://periodicos.puc-campinas.edu.br/transinfo/article/view/10190>
- Lopezosa, C., Guallar, J., & Franch, P. (2024). Independent newsletters and content curation: The editors' view. *Tripodos*, (56), 167-182. <https://doi.org/10.51698/tripodos.2024.56.02>
- Mondéjar, D. (2024). "Newsletters. The Renaissance of a Valuable Product to Reach the Audience." In *Innovations in Journalism*, edited by Klaus Meier, 241-248. London: Routledge. <http://doi.org/10.4324/9781032630410-30>
- Pell, D. (2021, November 10). The Internet's Unkillable App. *The Atlantic*. <https://www.theatlantic.com/ideas/archive/2021/11/internet-newsletters-dave-pell/620664>

Rojas-Torrijos, J.L. & González-Alba, J. A. (2018). La newsletter como producto periodístico en la búsqueda de nuevos lectores. Estudio de Boletines de Noticias de El País, El Español y El Independiente. *adComunica*, 15, 165–195. <https://doi.org/10.6035/2174-0992.2018.15.9>

Rourke, M. (2021). The newsletter boom: capitalizing on the intimacy of the inbox. In: *Digital Content Next*. <https://digitalcontentnext.org/blog/2021/10/21/the-newsletter-boom-capitalizing-on-the-intimacy-of-the-inbox>

Santos-Silva, D. & Granado, A. (2019). Old Formats, New Combinations: How Newsletters Represent Innovation and Blurring Boundaries in Digital Journalism". In *ECREA Journalism Studies Conference 2019*, Vienna. <https://novaresearch.unl.pt/en/publications/old-formats-new-combinations-how-newsletters-represent-innovation>

Seely, N. & Spillman, M. (2021). Email Newsletters: An Analysis of Content from Nine Top News Organizations. *Electronic News*, 15 (3-4), 123–138. <https://doi.org/10.1177/19312431211037681>

Silva-Rodríguez, A. (2021). Emergencia de newsletters especializadas en Covid-19: Información curada y actualizada en el email. *Profesional de la Información*, 30(4), e300410. <https://doi.org/10.3145/epi.2021.jul.10>

Suárez, E. (2020). How to build a successful subscription news business: lessons from Britain and Spain. *Journalist Fellowship paper*, Reuters Institute for the Study of Journalism. <https://bit.ly/3feaoPH>

Terren, L. & Borge, R. (2021). Echo Chambers on Social Media: A Systematic Review of the Literature. *Review of Communication Research*, 9, 99-118. <https://doi.org/10.12840/issn.2255-4165.028>

Upmanyu, K. (2024). Re-humanising News: Delineating the Trends and Research Scope for Curatorial Journalism. *Studies in Media and Communication*, 12(1), 33–41. <https://doi.org/10.11114/smc.v12i1.6255>

Weder, F., Weaver, C. K., & Rademacher, L. (2023). Curating conversations in times of transformation: Convergence in how public relations and journalism are “Doing” communication. *Public Relations Inquiry*, 12(2), 163-182. <https://doi.org/10.1177/2046147X231154550>

Zilberstein, S. (2022). Digital Platforms and Journalistic Careers: A Case Study of Substack Newsletters. *Columbia Journalism Review*. The Tow Center for Digital Journalism at Columbia’s Graduate School of Journalism. https://www.cjr.org/tow_center_reports/digital-platforms-and-journalistic-careers-a-case-study-of-substack-newsletters.php

Curación de contenidos en redes sociales de revistas científicas. Tendencias y buenas prácticas

Wileidys Artigas

High Rate Consulting, EUA

<https://orcid.org/0000-0001-6169-5297>

Jesús Cascón-Katchadourian

Universidad de Zaragoza, España

<https://orcid.org/0000-0002-3388-7862>

Javier Guallar

Universitat de Barcelona, España

<https://orcid.org/0000-0002-8601-3990>

Artigas, W., Cascón-Katchadourian, J., & Guallar, J. (2025). Curación de contenidos en redes sociales de revistas científicas. Tendencias y buenas prácticas. En J. Guallar, M. Váñez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 39-50). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuivcom.04.esp>

Resumen

Se presentan las principales conclusiones de diversos estudios realizados en el marco del proyecto Cuvicom sobre redes sociales de revistas científicas de las áreas de Comunicación y de Biblioteconomía y Ciencias de la Información, con especial atención a su uso de la curación de contenidos y a la red social X (antes Twitter), así como indicaciones prácticas en forma de análisis de tendencias, recomendaciones y buenas prácticas. Se analizan cuestiones como la presencia de la curación en el conjunto de las publicaciones en redes sociales, las temáticas de estos posts, las técnicas de curación empleadas o la relación entre las publicaciones y los tipos de editoriales según su tamaño. Se espera que las recomendaciones y buenas prácticas mostradas sean de utilidad para estudiosos del ámbito y para gestores de revistas científicas y de redes sociales.

Palabras clave

Redes sociales; Curación de contenidos; Revistas científicas; X; Editoriales; Comunicación; Ciencias de la Información.

1. Introducción

En los últimos años, las redes sociales se han consolidado como canales estratégicos para la comunicación científica, y en este contexto, las revistas académicas, conscientes de su potencial, las están empleando en mayor o menor medida dependiendo de las estrategias de cada editorial o editor. Por todo ello son objeto de un creciente interés de investigación.

En esta línea, el presente capítulo sintetiza las principales buenas prácticas y recomendaciones derivadas de una serie de estudios desarrollados entre 2022 y 2025 en el marco del proyecto Cuvicom, sobre revistas científicas de alto impacto en las áreas de Comunicación (en adelante, COM) y Ciencias de la Información (LIS), tanto en Iberoamérica como a nivel global, tomando como referencia los datos que ofrece Scimago Journal Rank, SJR (<https://www.scimagojr.com/journalrank.php>), de la base de datos Scopus. Se aborda la adopción de redes sociales, el uso de técnicas de curación, las temáticas más frecuentes y las estrategias que generan mayor impacto, así como la relación entre el uso de las redes sociales y el tipo de editorial según su tamaño, todo ello con el objetivo de ofrecer a editores, gestores de revistas y profesionales de la información, así como estudiosos del ámbito, una guía práctica, útil y aplicable para optimizar sus publicaciones en redes sociales.

2. Adopción y uso de redes sociales en revistas científicas

2.1. Presencia en redes de revistas Q1 SJR COM y LIS

Se han estudiado los datos de presencia de revistas científicas en redes sociales, que se ofrecen en porcentajes debido a que hay más revistas de COM que de LIS, tanto en Iberoamérica (43 COM vs 20 LIS) como en las revistas de alto impacto del primer cuartil Q1 (112 vs 61).

Como se puede ver en la tabla 1, hay sensibles diferencias entre las dos categorías en función de si ponemos el foco en las revistas iberoamericanas de todos los cuartiles o las globales de alto impacto y siempre observando los datos de perfiles propios y activos. En las primeras el desempeño de las revistas LIS es mejor que en COM, mientras que en las segundas las revistas COM se comportan bastante mejor que las de LIS.

En cualquier caso, aunque la evolución de las revistas de ambas categorías tiende a mejorar en el uso de las redes sociales, hay que decir que hay un amplio margen de mejora, sobre todo en las revistas Q1 de la categoría LIS. La percepción de esta investigación es que las revistas iberoamericanas (que suelen tener menor impacto y aparecen en cuartiles bajos en el SJR) no tienen demasiado interés en subir contenidos a las redes sociales generalistas. Mientras que entre las revistas de alto impacto globales, en su gran mayoría copadas por editoriales grandes y gigantes, algunas editoriales no favorecen el uso de redes sociales propias, sino que aglutinan las publicaciones en la red social de la editorial.

Tabla 1
Revistas y perfiles de redes sociales.

Categoría	Número revistas	Perfil red social	Perfil propio y activo
COM Iberoamérica	43	20 (46,5%)	16 (37,2%)
LIS Iberoamérica	20	12 (60%)	11 (55%)
COM Q1	112	100 (89,2%)	63 (56,2%)
LIS Q1	61	55 (90,16%)	16 (26,22%)

2.2. Rol de editoriales pequeñas vs. grandes

En cuanto al rol de las editoriales pequeñas frente a las grandes y las gigantes en el estudio de Cascón-Katchadourian et al. (2024) se analizan las revistas de mayor impacto en el SJR en ambas categorías. Las tendencias observadas son que la gran mayoría pertenecen a editoriales gigantes (75% COM, 68% LIS), entre las que destaca Taylor and Francis. Si a eso le unimos que del resto de revistas un buen número pertenece a editoriales grandes (19,6% LIS frente a 9,8% COM) nos encontramos con que la inmensa mayoría, en torno al 85%-90% de las revistas son de editoriales de gran tamaño. Estas editoriales tienen un peor comportamiento en cuanto a uso y adopción de redes sociales (salvo la excepción de las revistas de editoriales grandes de la categoría COM, donde un 63,6% posee al menos un perfil activo en redes sociales), que las revistas de editoriales pequeñas e independientes que tienen una mayor presencia en redes sociales respecto a las anteriores. Una hipótesis es que las revistas de editoriales pequeñas tienden a ser más activas en redes sociales porque dependen en mayor medida de la visibilidad para mantenerse en un entorno competitivo. En cambio, las grandes editoriales, al contar con un volumen elevado de publicaciones y mayor consolidación, podrían no requerir tanta presencia activa en estas plataformas.

2.3. Red social preferida

En cuanto a la plataforma social más usada por parte de las revistas científicas, los estudios muestran un predominio de X sobre el resto de las redes sociales generalistas (Facebook, LinkedIn, Instagram y YouTube), cuestión que se ha acentuado con el paso de los años. Si en el año 2022 las revistas COM de Iberoamérica tenían el mismo número de perfiles propios

activos de Twitter que de Facebook, en 2023 en las revistas LIS de Iberoamérica Twitter (ahora X) estaba en 9 revistas mientras Facebook solo en 4. En el año 2024 un estudio sobre las revistas de mayor impacto de LIS y COM reflejó un dominio todavía mayor de X sobre Facebook, tanto en número de perfiles, como en perfiles en exclusiva en la red social. Facebook aparecía en segundo lugar a gran distancia, pero esta red social también saca mucha ventaja a Instagram, LinkedIn y YouTube. Ante estos datos parece clara la recomendación de priorizar X, si bien la notable presencia de revistas en Facebook también se puede tener en cuenta.

3. Niveles de curación

En general todos los estudios realizados han mostrado que la curación de contenidos dentro de los posts de redes sociales por parte de las revistas científicas ha sido alta o muy alta. En las revistas de comunicación iberoamericanas, su nivel de posts curados es al menos del 80% en la mayoría de ellas (en 20 de los 24 perfiles que poseen las revistas), teniendo 9 perfiles el 100% de sus posts curados. Algo similar ocurre con las revistas LIS de Iberoamérica, donde 15 de los 18 perfiles tienen más del 86% de contenidos curados.

Por su parte, en las revistas Q1 COM y LIS globales, encontramos que las revistas COM tienen el 80% de sus publicaciones con contenidos curados, mientras que las de LIS les superan con el 93%.

Por otra parte, hay que destacar que, aunque los estudios sobre curación de contenidos generalmente versan sobre contenido curado externo, las revistas científicas, como también sucede con los medios de comunicación, suelen priorizar la curación de sus propios contenidos publicados, lo que también se ha denominado auto-curación (Guallar et al., 2022).

4. Técnicas de curación de contenido más empleadas

En el estudio de Artigas y Guallar (2022), para las revistas iberoamericanas COM estudiadas, se puede concluir que la técnica más empleada es "comentar", y en menor medida, "resumir", siendo más minoritarias las de "citar" y "retitular", mientras que otras tienen un uso inexistente. Las revistas suelen emplear de manera predominante una u otra técnica en su estilo de curación; así, por ejemplo, "comentar" es mayoritaria en Review of Communication Research, Comunicación y Sociedad o Revista Latina de Comunicación, "resumir" en Comunicar y Revista de Comunicación y "citar" en Profesional de la información (EPI).

Un año después, Cascón-Katchadourian et al. (2023) encontraron para el caso de revistas LIS iberoamericanas, que no se observa ningún caso de *storyboarding* y de paralelizar y que, salvo algunas excepciones, cada revista suele tener una técnica preferida que utiliza mayoritariamente (por ejemplo, EPI utiliza de manera destacada "citar" en sus cuatro redes sociales). En el análisis del perfil de red social se encuentra cierto equilibrio entre el uso de "comentar", "extraer", "resumir" y "citar", destacando "comentar" como técnica preferida en seis perfiles, respecto a los cuatro perfiles que usan las otras técnicas.

Es importante destacar que en referencia a las técnicas queda claro que cada revista elige y sostiene técnicas que continúa usando a lo largo del tiempo, sean estas individuales o combinadas, pues se observa la repetición del formato en el cual publican de forma continua. Las revistas que usan combinación de técnicas han podido sacar provecho al valor añadido de cada técnica usada y mostrar formas distintas de curación, por ejemplo, no simplemente copiando el título

del artículo y autores para recordar su publicación, sino también colocando citas o resumiendo a través de videos o muestras de extractos del contenido para atraer al lector.

Por su parte, en el estudio de revistas Q1 globales de ambas áreas (Cascón-Katchadourian et al., 2025), se observaron similitudes y diferencias entre las mismas. Por ejemplo, “extraer” y “resumir” eran las principales técnicas usadas, siendo la menos usada “retitular”; sin embargo, se observó que el área COM usa más “resumir”, mientras que LIS hace mayor uso de “extraer”, lo que destaca que en el área COM usan técnicas más complejas de curación.

En dicho estudio de 2025 también se analizó la forma de integración de contenido, observando el uso de las menciones (en promedio una por post en ambas categorías), así como el uso de hashtags (en el caso de LIS un promedio de 2,68 por post, doblando a COM con 1,46 por post), lo que muestra un mayor uso en las revistas LIS de los hashtags como herramienta de categorización del post en las redes sociales.

Se puede entonces concluir que para el caso de Iberoamérica la técnica de “comentar” es mayormente utilizada (LIS y COM), mientras que cuando se pasa al estudio de revistas Q1 de las mismas áreas se observa un mayor uso de “extraer” y “resumir”. En la tabla 2 se muestra un resumen de los principales hallazgos encontrados.

Tabla 2
Técnicas y su uso en las diferentes publicaciones.

	2022 (Iberoamericanas COM)	2023 (Iberoamericanas LIS)	2025 (Q1 LIS y COM)
Técnica mayormente usada	Comentar	Comentar	Extraer (LIS) y Resumir (COM)
Técnica menos usada	Retitular	Retitular	Retitular
Combinación frecuente	Resumir y comentar	Extraer y comentar	Extraer/Resumir con comentar
Nuevas inclusiones en los estudios	No aplica	No aplica	Integración de contenido, menciones y hashtags

Fuente: Artigas y Guallar (2022), Cascón-Katchadourian et al. (2023; 2025).

5. Temáticas frecuentes

Los temas de los posts en redes sociales de revistas científicas son bastante coincidentes entre los estudios publicados, con algunos matices. Estos datos se han recogido con diferencias metodológicas, puesto que en los dos primeros estudios se distinguen solo 3 categorías temáticas mientras que en el último estudio de 2025 la categoría Otros se subdivide en varias.

En cualquier caso, la curación prioriza el contenido propio de la revista, especialmente del último número publicado. Por ello, las temáticas más comunes giran en torno a la difusión y promoción de los artículos recientes, en su mayoría se trata de un post que presenta un artículo publicado por la revista. Además, encontramos otros tipos como, por ejemplo, *podcasts* con autores de los artículos, videos relativos a los artículos e invitaciones abiertas a revisar ciertos artículos. En el artículo de Artigas y Guallar (2022) hay un empate entre Último número y Anuncios y Otros debido a la gran amplitud temática de esta segunda categoría. En el estudio de Cascón-Katchadourian et al. (2023), donde todavía se mantuvo la categoría Anuncios y Otros sin dividir, a pesar de ello la categoría mayoritaria es Último número publicado en 12 de 18 perfiles estudiados. Además, cuantitativamente es el contenido más numeroso con diferencia, con 275 de los 454 contenidos analizados (60% del total). En el tercer estudio de 2025, tanto

en la categoría COM como en la LIS, Último número publicado es la categoría mayoritaria con el 64,97% y el 71,63% respectivamente de los posts analizados.

En segundo lugar, en los estudios de 2022 y 2023, la categoría Otros/anuncios u otras temáticas engloba diversos temas como: Call for Papers, Anuncios de la publicación de un número, Temas de futuros números, Difusión de actividades propias, Difusión de contenidos externos a la revista, Convocatorias a premios, Clasificación en rankings de revistas y Otros. Se trata de la segunda categoría temática en importancia por delante de Números anteriores que está compuesta por la difusión de publicaciones de números anteriores. En el primer estudio sobre las revistas iberoamericanas de comunicación es la categoría dominante en 11 perfiles de 8 revistas, similar a lo visto con Último número publicado. En el estudio de 2023 es la categoría predominante en 5 perfiles de 5 revistas y cuantitativamente es la segunda categoría más empleada con el 25,5% de los contenidos curados.

En el estudio de 2025 esta categoría se divide en otras varias con porcentajes de posts lógicamente más modestos, se trata de Call for Papers, Difusión del volumen actual, Contenidos externos, Actividades propias y Otros.

Tabla 3

Resto de temáticas excluyendo Último número publicado.

Temas curados	% de posts COM	% de posts LIS
Call for Papers	4,87%	2,02%
Difusión volumen actual	8,60%	2,77%
Contenido externo	2,33%	2,35%
Actividades propias	4,26%	5,40%
Otros	3,44%	0,09%
Volúmenes anteriores	11,59%	15,71%

Fuente: Cascón-Katchadourian et al. (2025).

Por tanto, Difusión de artículos de volúmenes anteriores, que en anteriores estudios era la categoría más modesta ahora emerge en el estudio de 2025 como segunda categoría temática en importancia con el 11,59% de los posts en COM y el 15,71 en LIS, solo detrás de Número actual. Hay que decir que esta categoría ha ganado en importancia con el paso de los años ya que en 2022 en solo dos perfiles era la predominante (uno de ellos con un solo post), y, en 2023 un solo perfil donde domina en el Twitter de Revista Española de Documentación Científica. Cuantitativamente también crece ligeramente del 13,8% de los posts en LIS en el año 2023 en las revistas iberoamericanas al 15,71% en LIS en 2025 en las revistas de alto impacto.

6. Buenas prácticas en temáticas y en técnicas de curación de contenido

Por último, se han extraído y presentado algunos casos de prácticas eficientes y notorias de curación presentadas en los artículos de Artigas y Guallar (2022) y Cascón Katchadourian et al. (2023). Para cada ejemplo se muestra la publicación en redes sociales con sus características esenciales (texto, descripción de la imagen, enlace al contenido curado), y un breve análisis.

6.1. Técnicas

Extractar

Revista Latina de Comunicación Social

Texto: “‘La Noticias falsas y desinformación sobre el Covid-19: análisis comparativo de seis países iberoamericanos’ escrito por Liliana Gutiérrez, Patricia Coba y Javier Andrés Gómez en el número 78 #Especial #Covid19 de #RevistaLatina”.

Imagen: De creación propia, es un gráfico de un mapa mundi con el área iberoamericana señalada y las palabras Covid-19.

<https://twitter.com/revistalatinacs/status/1337412573839446019>

Ejemplo de Extractar, una técnica caracterizada por su simplicidad y casi siempre por la automatización. Aunque ofrece escasa aportación de valor, puede permitir a una revista mantener una mínima presencia de contenidos en redes sociales con poca dedicación.

Retitular

Revista de Comunicación

Texto: “¿Netflix es entretenimiento, o hay otros elementos que van más allá del disfrute personal de los jóvenes? Lee el artículo de Benavides Almarza, C., & García-Béjar, L. <https://doi.org/10.26441/RC20.1-2021-A2> #rcom #udep”.

Imagen: fotografía de una familia viendo Netflix en el televisor.

<https://twitter.com/RevComUdep/status/1370393863878279172>

En este caso, el título original, que se inicia con un interrogante (“¿Por qué ven Netflix quienes ven Netflix?”), es sustituido en el tweet por otro interrogante (“¿Netflix es entretenimiento, o hay otros elementos que van más allá del disfrute personal de los jóvenes?”). Interesante uso de la retitulación para despertar el interés de la audiencia por el contenido del artículo.

Resumir

Revista Española de Documentación Científica

Texto: “Sergi Cortiñas y Marc Darriba Zaragoza estudian el porcentaje de obras pseudo-científicas presentes en #bibliotecas públicas. Resultados: mayor porcentaje #astrología | Más números de títulos sobre #homeopatía en proporción a #quimioterapia. <https://doi.org/10.3989/redc.2018.1.1474>”.

Imagen: no tiene.

https://twitter.com/Revista_REDOC/status/1529795736191344640

Un buen uso de la técnica de resumir se encuentra en este tweet sintético y directo, en el que todo el contenido se basa en el texto, sin incluir imagen. El tweet resume en pocas palabras el tema de un artículo publicado en la revista, así como dos de sus resultados más significativos, indicando también el nombre de los autores. Además, utiliza con precisión algunos hashtags para identificar determinadas palabras clave: bibliotecas, astrología, homeopatía, quimioterapia.

Comentar

Comunicar

Texto: "We invite you to read this interesting paper, which describes a language learning model that applies social media to foster contextualized and connected language learning in communities. Read more on: <https://doi.org/10.3916/C50-2017-01#AcademicTwitter#phdchat>".

Imagen: Texto con el título del artículo, los autores y su país junto a una fotografía que muestra una señora y dos niños mirando un libro ilustrado.

https://twitter.com/Rev_Comunicar/status/1445812347998773253

En este ejemplo, el tweet busca, más allá de la descripción del contenido del artículo, un acercamiento más personal al lector ("We invite you") y emplea un tono de opinión y recomendación ("this interesting paper").

Citar

Cuadernos.info

Texto: "#LecturaRecomendada @NekaneParejo analizó 11 películas de Almodóvar centrándose en la figura del fotógrafo. Concluyó una evolución hacia la caracterización de otros roles: voyerismo, fotografía doméstica e imágenes robadas. Lee el artículo completo en <https://doi.org/10.7764/cdi.52.36267>".

Imagen: consiste en una frase destacada entre comillas del artículo con indicación de la página concreta y el autor. También incluye el nombre de la revista, el título, el año de publicación, el número y las páginas.

<https://twitter.com/CuadernosInfo/status/1564650714395287554>

Un buen ejemplo del uso de la técnica de citar en este tweet. La cita se incorpora al post no en el texto, sino a través de una imagen mediante un gráfico de aspecto cuidado. La cita seleccionada permite resaltar uno de los resultados más destacados de la investigación, de modo que se condensa en pocas palabras el artículo que se difunde.

6.1. Temas

Número actual

Profesional de la información

Texto: "Thanks for joining our life: Intimacy as performativity on YouTube parenting vlogs. By @aVizcainoVerdu, @Patriciadcm and @DaniJaraDent #SocialMedia #Intimacy #Performativity #YouTube #Platforms #Vlogs".

Imagen: contiene el título del artículo, el nombre y los retratos de las tres autoras, una cita textual del artículo y los datos de referencia de este.

https://twitter.com/revista_EPI/status/1556520670766456832

https://twitter.com/revista_EPI/status/1559208973722583044

EPI en Twitter (no en el resto de las plataformas sociales) publica dos tweets para cada nuevo artículo de la revista: el primero de ellos, en cuanto aparece el artículo en la modalidad de publicación continua, y el segundo, unos días más tarde. El contenido de ambos tweets es mayoritariamente idéntico, pero traducido al inglés el segundo incluyendo la cita textual seleccionada, así como el formato gráfico de la imagen, que es diferente en las primeras y en las segundas publicaciones de cada artículo. Esta duplicidad se considera, técnicamente en el ámbito de los social media, una práctica de reciclaje de contenido, y es algo que se aconseja realizar especialmente en aquellas plataformas sociales en las que los contenidos tienen un menor tiempo de vida útil, como es el caso de Twitter (Sanagustín, 2017). De esta manera, aumentan las posibilidades de visibilidad y de acceso al nuevo artículo de la revista.

Números anteriores

Review of Communication Research

Texto: “‘Social Norms: A Review’, authored by Adrienne Chung & Rajiv Rimal has been cited 100 times in Google Scholar. Congrats to the authors! Congrats also to the reviewers, & Lijiang Shen, who served as editor, for helping the authors to publish an excellent literature review. We are happy to see that the articles we publish are useful to scholars worldwide! <https://buff.ly/2CH5Kbc>”.

Imagen: Grupo de personas de celebración.

<https://www.facebook.com/RCommunicationR/posts/2792781654090639>

Excelente ejemplo de post que hace referencia a un artículo publicado por la revista en números anteriores, del cual se informa que ha alcanzado un número de citas considerable en Google Scholar. El texto incluye referencias al contenido y a sus autores y editor (Nota: el enlace interno desde el post de Facebook al artículo quedó roto posteriormente a su selección para este estudio).

Otros temas

BiD

Texto: “Anàlisi de l’ISBN a Iberoamèrica <https://ub.edu/blokdebid/ca/content/analisi-de-lisbn-iberoamerica> amb Isabel-Cristina Arenas al #BlokdeBiD @revistaBiD”.

Imagen: muestra la imagen que se presenta en el blog.

<https://twitter.com/revistaBiD/status/1506627616173965320>

Un buen ejemplo de post en redes sociales no relacionado estrictamente con los contenidos de la propia revista son los tweets habituales en el perfil de BiD acerca del Blok de BiD, un blog con el que la revista guarda relación, y que publica reseñas de informes y estudios sobre las temáticas que aborda la revista. De esta manera, puede aumentar la visibilidad de su perfil de Twitter, y por extensión, de la marca de la revista.

7. Recomendaciones prácticas para revistas científicas

A continuación, se señalan algunas recomendaciones para editores y gestores de revistas académicas.

En cuanto a la adopción y el uso:

- En primer lugar, los editores deberían considerar la creación de perfiles propios de cada revista en redes sociales como una estrategia para aumentar la visibilidad e impacto de sus revistas.
- En segundo lugar, deberían considerar mantener una presencia activa al menos en una red social principal, idealmente X, ya que actualmente es la plataforma social preferida por la comunidad científica. No obstante, dadas las recientes transformaciones y según cómo evolucionen en el futuro cercano, también podría explorarse la incorporación a la nueva plataforma social Bluesky. Es también interesante tener otra red social complementaria, que puede ser Facebook o LinkedIn. En cualquier caso, lo más correcto sería la elección de al menos una red social en la que considere que tiene más presencia su público objetivo. Asimismo, se recomienda que los editores consideren el uso de otras plataformas como Instagram, LinkedIn o YouTube, siempre que cuenten con una estrategia clara y recursos para mantenerlas activas.
- Publicar con regularidad y evitar periodos prolongados de inactividad. Tener un perfil de red social para publicar dos veces al año no tiene sentido y da mala imagen a la publicación.
- Designar a una persona o equipo editorial responsable de las redes. Invertir en la difusión de contenidos es una estrategia que siempre da sus frutos.
- Otra recomendación es adoptar diferentes formatos de difusión que ayuden a que la presencia en las redes sea continua, por ejemplo, revistas como Mediterránea han logrado tener un ecosistema de medios de difusión entre los que destacan: redes sociales, audio resumen (podcast), webinars y grupo de WhatsApp (Guallar, 2024).

En cuanto a la temática:

- Es recomendable publicar contenido regularmente sobre una variedad de temas y, aunque se sigan priorizando los artículos más recientes, se puede volver también a publicaciones anteriores para darles nueva vida, curando todo el contenido, incluidos materiales externos, que pueda ser de interés para su comunidad. De hecho, reutilizar contenidos pasados con nuevos enfoques es una forma rentable y eficiente de poner en valor de nuevo los contenidos de la revista.
- En relación con lo anterior, aprovechar el perfil para fomentar comunidad, y no solo para difundir publicaciones propias. Que los usuarios visiten recurrentemente el perfil para ver las novedades de su campo específico.
- Un ejemplo de temática novedosa se ha visto en revistas que hacen alianzas o presentan concursos sobre los mejores papers en donde interactúan con otras revistas de su nivel y atraen al público en grupo.
- Además, se recomienda incluir algunos contenidos externos relevantes y de interés para el público objetivo, como la actualidad de la temática y noticias relevantes relacionadas.

En cuanto a las técnicas:

- Combinar diversas técnicas de curación para propiciar el interés del lector.
- Integrar contenido curado de distintas formas y usar estratégicamente los hashtags y menciones para lograr mayor visibilidad.

8. Conclusiones

La curación de contenidos en redes sociales, según los estudios realizados en los últimos años, es una práctica en auge por parte de las revistas científicas. Su capacidad para conectar con las audiencias, sintetizar el valor del conocimiento publicado y ampliar el impacto de los artículos, la posiciona como una práctica necesaria en el ecosistema digital académico actual.

Las evidencias recopiladas en las investigaciones recientes demuestran que, aunque aún existe un margen amplio de mejora, muchas revistas están haciendo uso de la curación dentro de las redes sociales. Desde la elección de plataformas adecuadas hasta la aplicación de técnicas editoriales que aporten valor, la curación de contenido se revela como una práctica adaptable, escalable y eficaz.

Las recomendaciones y casos analizados en este capítulo ofrecen una guía concreta para que otras revistas puedan replicar y adaptar estas prácticas a sus propias realidades. La clave radica en comprender que no se trata solo de difundir, sino de dialogar, interpretar y conectar. Esto implica asumir una mirada editorial activa, creativa y centrada en el lector.

A partir de lo expuesto, se puede afirmar que llevar a cabo una curación de contenidos de forma coherente y sostenida en el tiempo en redes sociales representa una estrategia clave para las revistas científicas. Esta práctica no solo les permite dar visibilidad y difusión a sus publicaciones específicas, sino también fortalecer su identidad de marca ante públicos amplios. Asimismo, ofrece a los autores una plataforma valiosa para promocionar sus trabajos, lo que incrementa sus oportunidades de ser leídos, evaluados y eventualmente citados. En este sentido, la curación de contenidos realizada por las revistas académicas en sus canales sociales favorece una mayor atracción de lectores mediante diversas formas de presentación, lo que se traduce en un impacto y visibilidad superiores tanto para las publicaciones como para quienes las firman.

Asimismo, sería importante explorar desde las diferentes perspectivas de las revistas los formatos que más pueden generar interacción con el público objetivo en función de sus propios contextos, realidades e intereses. Esto queda como tarea para futuras investigaciones.

9. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUCOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

10. Referencias

- Artigas, W., & Guallar, J. (2022). Curación de contenidos científicos en medios sociales de revistas Iberoamericanas de Comunicación. *Revista de Comunicación*, 21(2), 15–32. <https://doi.org/10.26441/RC21.2-2022-A1>
- Cascón-Katchadourian, J., Artigas, W., & Guallar, J. (2023). Curación de contenidos en las redes sociales de revistas de información y documentación de Iberoamérica. *BiD: textos universitaris de biblioteconomia i documentació*, 51 (diciembre). <https://doi.org/10.1344/BID2023.51.05>
- Cascón-Katchadourian, J., Artigas, W., Arroyo-Machado, W., & Guallar, J. (2024). Uso de las redes sociales por las revistas de alto impacto de Comunicación y Ciencias de la información. *Revista Mediterránea de Comunicación*, 15(1), 19–38. <https://doi.org/10.14198/MEDCOM.25487>
- Cascón-Katchadourian, J., Guallar, J., & Artigas, W. (2025). Social media analysis of high-impact Information and Communication journals: Adoption, use, and content curation. *Publications*, 13(1). <https://doi.org/10.3390/publications13010005>
- Guallar, J. (2024). 25 recomendaciones para revistas científicas en redes sociales. En *13º Congreso Internacional sobre Revistas Científicas CRECS*, Arequipa, Perú, 10 mayo 2024. <http://eprints.rclis.org/45758>
- Guallar, J., Franch, P., Boté-Vericad, J.-J., & Anton, L. (2022). How do legacy and digital media curate coronavirus content. An assessment of newsletters from the USA and four European countries. *Profesional de la información*, 31(3), e310318. <https://doi.org/10.3145/epi.2022.may.18>
- Sanagustín, E. (2017). *Plan de contenidos en medios sociales*. Editorial UOC.

Curación de fuentes y narrativas de igualdad en la agenda mediática

Anna Ventura-Cisquella

Universitat de Barcelona, España

<https://orcid.org/0000-0002-3677-8055>

Javier Guallar

Universitat de Barcelona, España

<https://orcid.org/0000-0002-8601-3990>

Ventura-Cisquella, A., & Guallar, J. (2025). Curación de fuentes y narrativas de igualdad en la agenda mediática. En J. Guallar, M. Vállez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 51-71). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicom.05.esp>

Resumen

Se analiza la labor de las periodistas de igualdad en la prensa digital española con énfasis en la curación de fuentes documentales y las temáticas de los contenidos. El estudio se sustenta en un corpus de 316 artículos publicados durante el primer semestre de 2024 en dos diarios tradicionales (*El País* y *El Periódico*) y cuatro nativos digitales (*eldiario.es*, *El Salto*, *Infolibre* y *Público*), medios que actualmente cuentan con periodistas especializadas en igualdad. La identificación de temáticas se realizó mediante lectura exploratoria de titulares y textos completos, seguida de un análisis de frecuencia de etiquetas y palabras clave en cada artículo, organizadas en un libro de códigos. Para el análisis de fuentes se aplicó el método CAS (*Curation Analysis System*) que permitió clasificar todos los enlaces curados de cada artículo según el tipo de organización: medios de comunicación, fuentes oficiales, corporativas o ciudadanas. Los resultados muestran que las periodistas realizan una rica curación de fuentes documentales para fundamentar reportajes, elaborar titulares y visibilizar desigualdades. Se identifican cuatro grandes temas: desigualdades y violencias de género; avances y tensiones en movimientos feministas y derechos reproductivos; seguimiento del Ministerio de Igualdad y de sus políticas; y progresos y desafíos en derechos y violencias hacia la comunidad LGBTQ+. Este estudio pone de relieve la función estratégica de estas profesionales en la consolidación de narrativas de derechos dentro de la agenda mediática y en el registro de las políticas de igualdad en España.

Palabras clave

Curación de contenidos; Documentación; Editoras de género; Prensa digital; Igualdad; Agenda mediática; Fuentes documentales.

1. Introducción

En las últimas décadas los medios tradicionales han ido perdiendo protagonismo frente a los grandes gigantes tecnológicos. Este fenómeno ha contribuido a la aparición de un ecosistema informativo marcado por el auge del “capitalismo de plataformas”, la proliferación de pseudomedios y la difusión de discursos de odio que niegan las desigualdades en redes sociales (Pan et al., 2025; García-García, 2025).

Ante esta realidad, resulta fundamental analizar cómo se producen y difunden los contenidos mediáticos que buscan promover la igualdad de género, entendida como la equidad en derechos, responsabilidades y oportunidades entre mujeres y hombres, y también entre niñas, niños y personas con identidades de género diversas (ONU Mujeres, 2025).

En este sentido, la labor de las periodistas especializadas en igualdad enfrenta en la actualidad múltiples desafíos. Estas profesionales son frecuentemente cuestionadas por curar y editar contenidos informativos con enfoque en derechos e igualdad de género, especialmente cuando lo hacen en medios de línea progresista, y, principalmente, por parte de sectores ultraconservadores (Bennett y Livingston, 2018; Reuning et al., 2022).

En el caso de España, además, se observa una creciente desconfianza hacia las fuentes oficiales. Esta duda es instigada y amplificada por determinados grupos políticos de la derecha radical, como Vox. La situación complica el ejercicio de un periodismo especializado en igualdad, en un contexto gobernado por un Ejecutivo progresista.

La erosión de la confianza en las instituciones debilita a su vez la credibilidad de la información oficial y de los medios tradicionales donde trabajan las periodistas de igualdad. Esto impulsa la búsqueda de fuentes “alternativas” que no han sido verificadas.

Estas fuentes alternativas suelen estar asociadas a estrategias de la derecha radical como la “manosfera”, un conjunto de comunidades en línea centradas en discusiones sobre masculinidad, a menudo críticas del feminismo y de los medios tradicionales (García-Mingo y Díaz-Fernández, 2023).

Finalmente, la autoridad de estas profesionales también se ve socavada debido a su condición de mujeres periodistas, como han señalado distintos organismos internacionales (UNESCO, 2019; Amnistía Internacional, 2024).

En este escenario, el periodismo de curación documental se ha consolidado como una estrategia clave para garantizar la calidad y el rigor informativo (Guerrini, 2013; Guallar, 2014; Cui y Liu, 2017; Lopezosa et al., 2023). Esta labor informativa se concibe como un conjunto de prácticas orientadas a seleccionar, enlazar y contextualizar fuentes verificadas, con el objetivo de ofrecer información precisa, accesible y socialmente relevante (Guallar y Codina, 2018).

En esta línea, la presente investigación analiza el enlazado de fuentes de los contenidos elaborados por las periodistas de igualdad en la prensa digital española. Se ha seleccionado este objeto de estudio, porque si bien el crecimiento de las redes sociales y otras plataformas digitales ha desplazado en parte el protagonismo de los medios tradicionales (Waisbord, 2018; Bailo, 2021), la prensa digital sigue desempeñando un papel clave como agente de configuración de la agenda pública (Virgili y Puche, 2019). Esto la convierte en un espacio fundamental para influir en la percepción social, así como en la formulación de políticas y marcos legislativos vinculados a la igualdad de género.

Además, los contenidos de la prensa digital no se limitan a sus propios entornos, sino que también circulan ampliamente a través de las redes sociales. Incluso entre quienes no recurren a estas plataformas con fines informativos, la exposición incidental a noticias es habitual y se asocia con un incremento en el consumo de información en línea (Feezell, 2017; Fletcher y Nielsen, 2017).

Por otra parte, aunque investigaciones anteriores han abordado el papel de las periodistas de igualdad en las redacciones —centrándose en aspectos como la incorporación de editoras de género, sus funciones y su capacidad de influir en las líneas editoriales de distintas secciones del medio (Spinetta, 2022; Irazo-Cabrera et al., 2023)—, aún se carece de un análisis profundo sobre los contenidos que estas profesionales generan. Mediante el estudio del enlazado de fuentes documentales que realizan, esta investigación examina cómo procuran posicionar titulares y materiales vinculados con las desigualdades y violencias de género, así como con los derechos de las mujeres y del colectivo LGBTIQ+, dentro de la agenda mediática española.

Este trabajo puede resultar de interés para periodistas e investigadoras que busquen estrategias orientadas a promover contenidos de igualdad y a contrarrestar discursos de odio en los medios de comunicación.

2. Material y métodos

El análisis de contenidos se sustenta en un corpus de 316 artículos publicados por periodistas de igualdad durante el primer semestre de 2024, en dos diarios tradicionales españoles (*El País* y *El Periódico*) y cuatro nativos digitales (*ElDiario.es*, *El Salto*, *Infolibre* y *Público*), que son los que actualmente cuentan con estas figuras.

De los seis medios analizados, tres —*ElDiario.es*, *El País* y *El Salto*— cuentan con editoras de género que ocupan cargos de decisión estratégica dentro del equipo directivo. Estas profesionales supervisan, gestionan y mejoran la representación del género en los contenidos de todas las secciones del medio. Los otros tres medios —*El Periódico*, *Público* e *Infolibre*— tienen responsables de la sección de igualdad, pero no ejercen en la actualidad como editoras de género en sentido estricto. Su labor se centra principalmente en la ideación y redacción de contenidos específicos en la sección de igualdad.

De estos medios se seleccionó una muestra de artículos de periodistas de igualdad publicados en la primera mitad de 2024, una etapa marcada por acontecimientos relevantes. Durante este tiempo, España se situó como el cuarto país europeo con mejor desempeño en materia de igualdad (EIGE, 2023) y se produjo la transición en el Ministerio de Igualdad de Irene Montero a Ana Redondo (Ministerio de Igualdad, 2023), ambas dentro del espectro político progresista, pero con enfoques distintos en la gestión de la cartera. En este periodo, se intensificaron los debates sobre derechos trans, LGTBIQ+ e igualdad de género, junto con el aumento de discursos negacionistas desde la política, la juventud y las redes sociales (CIS, 2024; Europa Press, 2024).

La tabla 1 recoge los seis medios analizados y las periodistas responsables de las secciones de igualdad. Incluye sus roles, enlaces a las fichas de autoría y el número total de artículos publicados entre enero y junio de 2024. Estos artículos conforman la muestra estudiada.

Tabla 1

Medios con periodistas de igualdad y sus contenidos publicados.

Periódico	Nombre	Rol	Ficha de autora	Artículos
El Diario.es	Anna Requena	Editora	https://www.eldiario.es/autores/ana_requena_aguilar/	53
El País	Isabel Valdés	Editora	https://elpais.com/autor/isabel-valdes-aragones/	48
El Periódico	Patricia Martín	Redactora	https://www.elperiodico.com/es/autor/patricia-martin	29
El Salto	Patricia Reguero	Editora	https://www.elsaltodiario.com/autor/patricia-reguero	31
Infolibre	Sabela Rodríguez	Redactora	https://www.infolibre.es/autores/sabela-rodriguez/	76
Público	Candela Barro	Redactora	https://www.publico.es/author/candela-barro	79
Total				316

En cuanto a la metodología empleada, para hacer el análisis de temáticas se desarrolló una lectura exploratoria de titulares y textos completos. Después se hizo un análisis de frecuencia de etiquetas y palabras clave empleadas por las periodistas de igualdad para clasificar y posicionar cada artículo. Los resultados se organizaron en un libro de códigos siguiendo a Braun y Clarke (2006) y Krippendorff (2018).

Para el análisis de fuentes documentales se aplicó el método CAS (Curation Analysis System), ya utilizado en estudios sobre curación periodística (Guallar et al., 2021; Lopezosa et al.,

2023; Guallar et al., 2025). Se clasificaron los enlaces salientes de cada artículo según el tipo de organización: medios de comunicación, fuentes oficiales, corporativas y ciudadanas.

Durante el periodo analizado se identificaron cuatro temas principales: (1) desigualdades, discriminaciones y violencias de género; (2) avances, retrocesos y tensiones en los movimientos feministas, con foco en derechos reproductivos; (3) seguimiento del Ministerio de Igualdad y de la aplicación de sus políticas y leyes; y (4) progresos y desafíos en derechos y violencias hacia la comunidad LGTBIQ+.

En el abordaje de estas temáticas, las fuentes documentales —medios de comunicación, documentos oficiales, informes corporativos y tweets o declaraciones ciudadanas— fueron centrales. Sirvieron tanto para elaborar titulares y fundamentar reportajes como para analizar la construcción de debates públicos y el seguimiento de políticas de igualdad. A continuación, se presenta el análisis realizado, estructurado en cinco epígrafes:

- Algunas características de la curación de fuentes
- Cómo curar las desigualdades y violencias de género
- Cómo curar fuentes y recursos de los movimientos feministas
- Cómo curar leyes y políticas de igualdad
- Cómo curar avances y retrocesos en derechos LGTBIQ+

3. Algunas características de la curación de fuentes

Durante el primer semestre de 2024, las periodistas de igualdad enlazaron más de dos mil documentos (enlaces externos), concentrándose principalmente en medios como *El País* y *Público*, seguidos de *Infolibre*, *ElDiario.es*, *El Periódico* y *El Salto*, lo que evidencia que la curación con perspectiva de género se centra en unos pocos medios que buscan influir en la agenda informativa.

En cuanto al tipo de fuentes enlazadas, *El País* e *Infolibre* recurren sobre todo a fuentes oficiales para cubrir violencias, desigualdades y políticas públicas, mientras *El Salto* prioriza perspectivas comunitarias y de movimientos feministas, curando fuentes corporativas (entidades, colectivos y ONG). *El Periódico* y *Público* combinan fuentes oficiales, corporativas y ciudadanas (personas individuales), aunque estas últimas se emplean de manera limitada. Asimismo, estos medios adjuntan enlaces a contenido de terceros proveniente de diversas fuentes informativas para cubrir casos destapados por otros, denunciar malas prácticas periodísticas o abordar temas locales e internacionales relacionados con el feminismo.

La temporalidad de los contenidos es mayormente retrospectiva, con curación de enlaces a materiales y fuentes de más de un mes de antigüedad y, en ocasiones, de hasta 10 años atrás. Dado que se trata de enlaces retrospectivos, la preservación digital de los documentos es especialmente importante.

Las técnicas de curación empleadas incluyen resumen y contextualización, lo que permite a las editoras trazar relaciones, construir cronologías y definir conceptos clave, así como analizar estructuras de violencia y desigualdad y avances en derechos feministas y LGTBI, integrando también la agenda del Ministerio de Igualdad.

Respecto a la morfología de las fuentes, predominan los sitios web, seguidos de fuentes secundarias y redes sociales, mientras que los blogs tienen mínima presencia, reflejando un enfoque centrado en contenidos verificados.

4. Cómo curar las desigualdades y violencias de género

Al abordar la violencia de género —machista, feminicidios, sexual, vicaria, institucional y económica—, las periodistas hacen curación documental, enlazando y analizando fuentes, sobre cómo estas violencias se sostienen en estructuras sociales. También elaboran titulares, dan cobertura y enlazan fuentes para visibilizar desigualdades en el trabajo, la salud, las relaciones afectivas, los cuidados, la maternidad, la conciliación laboral y familiar, y la participación política y liderazgo de mujeres y colectivos disidentes.

Figura 1

(15/02/2024). Noticia de Anna Requena, editora de género en *ElDiario.es* sobre el coste económico de la violencia de género en España. El titular se basa en los datos de un estudio del Ministerio de Igualdad del año 2022.



Estas coberturas señalan los discursos antifeministas que circulan desde la ultraderecha y la “manosfera” en redes sociales, y que intentan deslegitimar las políticas de igualdad presentándolas como un ‘invento ideológico’. Un 30 % de los artículos analiza cómo estos discursos se construyen y disputan, ganando presencia en medios, plataformas digitales y en la política institucional, especialmente con el auge de Vox en España.

Las fuentes oficiales predominan en la documentación utilizada para cubrir las violencias de género y las desigualdades. Instituciones como el Ministerio de Igualdad y su Delegación del Gobierno contra la Violencia de Género publican informes estadísticos clave en el [Portal Estadístico de Violencia de Género](#). Estos informes sirven de referencia y sus enlaces se con-

sultan y publican al elaborar informaciones sobre feminicidios, denuncias, órdenes de protección y violencia de género, sexual y vicaria.

En estos artículos también se incluyen fuentes curadas relevantes para la prevención y el seguimiento de la violencia de género y sexual. Entre ellas, se proporciona el correo electrónico de contacto para consultas y denuncias: 0016-online@igualdad.gob.es. Además, también se da visibilidad a enlaces a documentos que explican los programas de atención a mujeres maltratadas, como el programa [VioPet](#), que facilita la acogida temporal de animales de mujeres víctimas de violencia de género.

A partir de estas piezas se destacan encuentros estatales en violencia, como el [Encuentro Estatal sobre Violencia Vicaria y Violencia de Género](#), que propician el intercambio de experiencias y la coordinación de políticas públicas en la materia.

Otra fuente oficial prioritaria es el [BOE](#) (Boletín Oficial del Estado) que permite mediante el uso de su [Biblioteca Jurídica Digital](#), consultar normas actualizadas en igualdad de género. Reformas como la del artículo 709 de la *Ley de Enjuiciamiento Criminal*, que prohíbe preguntas revictimizantes a víctimas de violencia sexual, y el artículo 681, que autoriza al tribunal a vetar preguntas vejatorias, entre otros artículos. Estos artículos son un ejemplo de fuente documental que facilita analizar avances legales y fiscalizar la aplicación real de esta normativa en los tribunales.

También los informes del Observatorio del [Consejo General del Poder Judicial \(CGPJ\)](#) son una fuente documental de valor; el análisis de sentencias ofrece a las periodistas la posibilidad de elaborar reportajes sobre la impunidad o las barreras en el acceso a la justicia.

Un ejemplo de ello fue el seguimiento periodístico del “caso Dani Alves”. El exfutbolista brasileño fue acusado en 2023 de violar a una joven en Barcelona, condenado en 2024 a cuatro años y medio de cárcel y finalmente absuelto en 2025 por el Tribunal Superior de Justicia de Cataluña, que consideró que no existían pruebas suficientes. El caso generó una fuerte controversia en España.

En este proceso resultó clave otra fuente documental: el Protocolo “[No Callem](#)” del Ayuntamiento de Barcelona, impulsado durante el mandato de Ada Colau para prevenir agresiones sexuales en espacios de ocio nocturno. Su aplicación en este caso tuvo relevancia internacional y mostró cómo una fuente documental oficial puede orientar actuaciones inmediatas de seguridad y servicios de emergencia, además de convertirse en noticia y ocupar diversos titulares.

Al cubrir los discursos antifeministas, se consultan informes de otras fuentes oficiales, como los del [Centro de Investigaciones Sociológicas \(CIS\)](#). Sus resultados se reflejan en noticias y reportajes, como el sondeo que mostró que un 44% de los hombres considera que las políticas de igualdad han avanzado tanto que ahora ellos se sienten discriminados.

Figura 2

(15/01/2024). Titular elaborado por Ana Requena a partir de una fuente documental oficial: encuesta del Centro de Investigaciones Sociológicas (CIS) sobre la percepción de los hombres respecto a las políticas de igualdad en España.

El 44% de los hombres cree que las políticas de igualdad han ido demasiado lejos y que son ellos los discriminados

La encuesta sobre la percepción de la igualdad publicada por el CIS muestra que las mujeres dedican tres horas al día a las tareas del hogar, frente a las dos de los hombres. La brecha se agranda cuando hay hijos: ellas emplean 6,7 horas; ellos 3,7

—El 44% de los hombres cree que las políticas de igualdad han ido demasiado lejos y que son ellos los discriminados



Foco | IGUALDAD



Miles de personas protestan con carteles durante la manifestación encabezada por el Moviment Feminista de Mallorca por el BM. Tomas Moyà / Europa Press

Ana Requena Aguilar
15 de enero de 2024 -13:14 h Actualizado el 08/03/2024 -22:55 h 119

SEGUIR AL AUTOR/A

PUBLICIDAD

elDiario.es
depende de ti
Apoya el periodismo independiente desde 3,33 €/mes



Hazte socio/a

A nivel internacional, destacan fuentes como enlaces curados al [Convenio de Estambul](#), que obliga a los Estados a prevenir y sancionar la violencia contra las mujeres, y a las campañas y recursos de [ONU Mujeres](#), como las “[16 formas de enfrentar la cultura de la violación](#)”, que buscan sensibilizar y prevenir la violencia de género a nivel global. Los informes y sesiones de organismos internacionales, como el documento de la [74ª sesión de la CEDAW](#), sirven como marco para políticas y reportajes sobre violencias.

Junto a las fuentes oficiales, las periodistas curan enlaces a fuentes corporativas, como los informes de Aspacia, entidad creada en España en 2020 para coordinar políticas sociales y de igualdad. Estas organizaciones, sostenidas en algunos casos con una parte de fondos públicos, ofrecen recursos como [violencia-sexual.info](#), con guías de apoyo psicológico, legal y social para víctimas de agresiones sexuales, a menudo enlazados en las piezas de periodismo de género.

Asimismo, las periodistas incorporan informes sociológicos —como las encuestas de la consultora de investigación social 40 dB para [El País](#) y la [SER](#)— que ayudan a contextualizar las brechas generacionales en torno al feminismo y la igualdad, junto con estudios de entidades como [Save the Children](#), que advierte de la alta exposición de jóvenes a la violencia sexual digital.

También se curan tweets y recursos procedentes de fuentes ciudadanas en X (anteriormente, Twitter) y en la web, donde se visibilizan voces expertas en la materia. Entre ellas destaca la curación de tweets de [Victoria Rosell](#), exdelegada del Gobierno contra la Violencia de Género y jurista especializada en derechos de las mujeres, o [Julia Humet](#), abogada y politóloga especializadas en violencias de género; y documentos e informes elaborados por [Sonia Vaccaro](#), psicóloga clínica y forense conocida por su trabajo sobre violencia de género y por acuñar el concepto de “violencia vicaria”.

En cuanto a fuentes documentales provenientes de medios, las periodistas enlazan estudios académicos que aportan evidencia sobre sexualidad y justicia. Por ejemplo, [Measuring female genital functions](#) de Roy J Levin (2004) examina métodos para evaluar funciones genitales femeninas. Este estudio ayuda a desmontar mitos sobre consentimiento y revictimización. Otro estudio, [Does the sex of judges matter?](#) de Arantegui, Romero y Tamarit (2024), analiza cómo el género de los jueces influye en las sentencias por delitos sexuales en España. Los resultados muestran que las juezas tienden a imponer penas más severas.

A la vez, se seleccionan y curan artículos de otros medios de comunicación tradicionales para abordar temas de interés local, internacional o complementar informaciones, como los titulares que señalan la mayor rigurosidad de las juezas en penas por delitos sexuales (El País), la primera ley de Castilla-La Mancha contra la brecha salarial (Periódico CLM / Público), el video El violador eres tú (La Nación, digital argentino) y los consejos de Interior sobre violaciones (20 Minutos).

Aparte de las violencias de género, las periodistas especializadas en igualdad recurren a una amplia gama de fuentes documentales para cubrir y elaborar titulares sobre las múltiples formas de desigualdad que afectan a las mujeres en distintos ámbitos:

En educación y empleo, se curan enlaces a fuentes documentales oficiales como el Instituto Nacional de Estadística (INE), Eurostat y el Sistema Estatal de Indicadores de la Educación (2023), que aportan datos sobre segregación formativa, [tasas de abandono escolar](#) y [acceso desigual al empleo](#). Del mismo modo, la Encuesta de Población Activa (EPA) y las Encuestas de Estructura Salarial del INE resultan esenciales para analizar las brechas laborales y económicas. De manera similar, la [Encuesta de Fecundidad del INE](#) (2018) ofrece claves sobre maternidad, conciliación y retraso en la edad de tener hijos, factores cruciales para comprender las desigualdades demográficas y laborales.

A nivel de fuentes corporativas, el informe de [GWL Voices](#), organización integrada por casi 80 mujeres líderes de todo el mundo, sirvió a las periodistas de igualdad para elaborar diferentes titulares revelando un dato clave, que solo 67 de los 523 líderes en organismos globales en los últimos 79 años han sido mujeres, evidenciando el techo de cristal en la representación internacional.

En maternidad, las periodistas recurren a recursos elaborados por entidades como *Yo no renuncio* ([yonorenuncio.com](#)) y *Familias Monomarentales* ([familiasmonomarentales.es](#)) para visibilizar desigualdades de género en conciliación, derechos de madres solteras y cargas de cuidado, respaldando sus artículos con datos y testimonios confiables.

Este trabajo de documentación y análisis de datos —combinado con herramientas de visualización de datos y microdatos accesibles— fortalece el periodismo de igualdad, permitiendo transformar cifras complejas en relatos comprensibles que visibilizan desigualdades persistentes.

Para continuar profundizado en la práctica del periodismo de igualdad, es fundamental considerar cómo se seleccionan y organizan las fuentes y recursos producidos por los movimientos feministas.

5. Cómo curar fuentes y recursos de los movimientos feministas

Las noticias sobre feminismos se centran en las luchas y avances de sus movimientos, con especial énfasis en los derechos reproductivos y sociales en España, y en menor medida, en otros países de Europa y América Latina. Asimismo, se abordan debates internos de los feminismos españoles y europeos, como la regulación de la prostitución, las implicaciones de la llamada "ley trans", la inclusión LGBTQ+ y la interpretación del consentimiento sexual tras la ley "Solo sí es sí", que redefine legalmente este concepto. También se incluyen entrevistas a feministas de la cultura, la política y el activismo, aportando perspectivas y reflexiones sobre los retos actuales de los movimientos feministas.

Figura 3

(04/04/2024). Titular elaborado por Sabela Rodríguez Álvarez para InfoLibre sobre la iniciativa de activistas feministas de ocho países europeos, incluida España, para garantizar el derecho al aborto libre, seguro y gratuito en toda la UE.

IGUALDAD

Abortar en España si en Polonia no se puede: mujeres europeas se organizan para blindar sus derechos

- ▶ Feministas de ocho países lanzan una iniciativa para que la Unión Europea sufrague el desplazamiento entre estados miembros y garantizar la libre interrupción del embarazo
- ▶ "En vez de actuar cuando nos quiten nuestros derechos, la idea es decirles que los vamos a blindar", dice la escritora Cristina Fallarás, una de las impulsoras de 'Mi voz, mi decisión'
- ▶ Exclusivo para socias y socios.

3 COMENTARIOS



Manifestación en 2022 contra la reforma del aborto en Polonia. EP



Sabela Rodríguez Álvarez
[X@SabelaRuiha](#)

4 de abril de 2024 - 20:57h Actualizado el 05/04/2024 - 06:00h



Estas piezas parten de una genealogía feminista que contextualiza luchas y debates actuales en un marco histórico más amplio. Más allá del marco feminista más punitivista o del seguimiento de las violencias de género, también destacan iniciativas propositivas que abren nuevos marcos en positivo.

Figura 4

(02/03/2024). Titular elaborado por Isabel Valdés para El País sobre una década de movilización feminista en España, desde el Tren de la Libertad hasta la influencia en redes sociales.



Se recopilan diversas fuentes. Entre ellas se incluyen fuentes oficiales, donde se curan hipervínculos, como [informes del Ministerio de Sanidad](#) sobre interrupciones del embarazo, así como enlaces a iniciativas europeas como [My Voice, My Choice](#) (ICE), que busca garantizar un acceso seguro y accesible al aborto en toda la Unión Europea. No obstante, se otorga una visibilidad destacada a las fuentes corporativas y ciudadanas que respaldan estos derechos.

Con este fin, se da visibilidad a enlaces a materiales elaborados por colectivos feministas. Algunos ejemplos son: [FemBloc](#), que ofrece [recursos](#) de apoyo a mujeres y personas LGTBQ+ frente a la violencia digital; los estudios de la [Asociación CATS](#), que defiende los derechos de las trabajadoras sexuales y ofrece servicios sociales y de salud; y las publicaciones de [Punt6](#), que promueven un urbanismo feminista más inclusivo y equitativo.

Por otro lado, también se observa una amplia publicación de artículos sobre los movimientos feministas en España que destacan cómo los debates sobre la abolición de la prostitución y los derechos de las personas trans los han dividido. En 2024, por tercer año consecutivo, las manifestaciones se celebraron por separado, incluida Madrid, con dos marchas simultáneas:

[El Movimiento Feminista de Madrid](#) pidió la abolición de la prostitución y rechazó la autodeterminación de género. Por su parte, la [Comisión 8M](#) defendió los derechos trans y la

erradicación de la trata. También apoyó la participación de las trabajadoras sexuales en los debates que las afectan, aunque no definió una postura sobre la prostitución. En este sentido, las periodistas especializadas en igualdad recopilaban y curaban enlaces a los argumentarios de ambos movimientos para facilitar su consulta y análisis.

En cuanto a los enlaces curados a otros medios, se incorporan entrevistas a feministas difundidas en diversos periódicos, junto con libros de distintas editoriales y reportajes externos.

Estos enlaces se emplean para desarrollar una genealogía feminista, es decir, para explorar cómo se han construido históricamente las ideas sobre las mujeres y el género. Con este propósito, se seleccionan enlaces a libros de editoriales feministas como [La Bella Varsovia](#), [Virus Editorial](#) o [Sudakasa](#), destacando la literatura escrita por mujeres. Se recopilan referencias a proyectos web relacionados con la memoria histórica y la visibilidad de las mujeres, como [Nietas de la Memoria](#), perfiles de figuras feministas y expertas en Wikipedia, y plataformas colaborativas como [Mujeres en Red](#), que buscan visibilizar a las mujeres y reducir la brecha de género en la enciclopedia digital. También se incluyen iniciativas periodísticas que reconstruyen la vida y obra de mujeres relevantes, como el artículo de [Vanitatis](#) en *El Confidencial* sobre Emilia Pardo Bazán, escritora y defensora de los derechos de las mujeres, o reportajes de memoria histórica a nivel local, como los elaborados por Newtral sobre [Ana Orantes, víctima de violencia de género](#) cuyo caso impulsó cambios legislativos en España.

De igual manera, se incorporan artículos, reportajes y entrevistas de otros medios que visibilizan a activistas proaborto, como [Justyna Wydrzynska](#), activista en Polonia, un país con leyes muy restrictivas sobre el aborto. Su labor fue reportada por medios como *El País* y posteriormente citada por otros medios incluidos en la muestra.

Por otro lado, estos medios progresistas incluyen de manera puntual declaraciones de medios del espectro conservador, como [COPE](#), con el fin de ilustrar posturas contrarias al aborto y contextualizar su origen.

Finalmente, y de manera más minoritaria se curan fuentes ciudadanas, con enlaces a expertas y divulgadoras como el perfil en Twitter de [Isabel Mellén](#), filósofa, historiadora del arte y escritora o el perfil en Instagram de [Júlia Salander](#), activista feminista y politóloga, ambas son divulgadoras feministas.

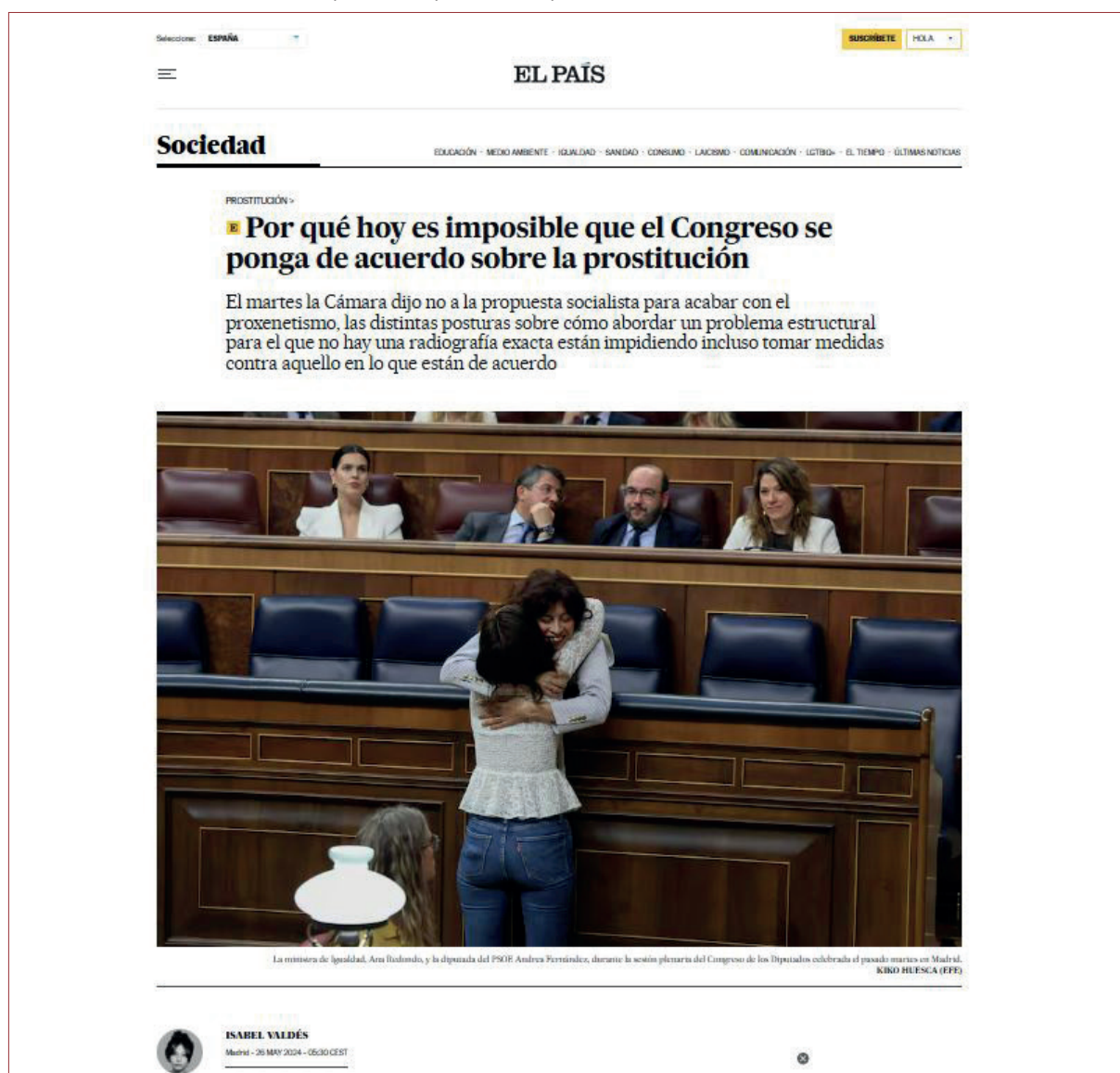
Además de estas voces, la curación se extiende al ámbito normativo y de las políticas públicas. Estos contenidos se introducen en el siguiente epígrafe.

6. Cómo curar leyes y políticas de igualdad

Las piezas de este bloque temático informan sobre la tramitación institucional de leyes clave como la normativa en torno a la prostitución, el aborto y la Ley Trans, además de la actividad del Ministerio de Igualdad y las posturas políticas en torno a la legislación en materia de igualdad. Analizan cómo se están implementando estas leyes y hacen seguimiento para comprobar si se aplican realmente o si hay retrasos. También cubren los debates sociales y políticos que generan, mostrando qué partidos las apoyan, cuáles proponen cambios y quiénes se oponen, evaluando así su impacto político y social.

Figura 5

Titular elaborado por Isabel Valdés para El País sobre las divisiones políticas y feministas que dificultan un consenso del Congreso español respecto a la prostitución.



A la hora de abordar estas temáticas se recopilan principalmente fuentes oficiales, especialmente los documentos publicados en las Salas de Prensa de los sitios web de los partidos políticos y del Ministerio de Igualdad y la propia normativa del BOE. Entre estas fuentes se incluyen iniciativas legislativas recientes, como [la propuesta del PSOE](#) para prohibir el proxenetismo (2024) y el [Plan Camino](#) (2022), aprobado por el Ministerio de Irene Montero (Unidas Podemos), que busca ofrecer alternativas económicas, laborales y sociales a las víctimas de trata y explotación sexual.

También se incluyen enlaces a documentos de fuentes corporativas, como la carta de [Human Rights Watch](#) al Congreso o las publicaciones de [Amnistía Internacional](#), que advierten sobre el impacto de las leyes punitivas en los derechos de las trabajadoras sexuales y aportan a las coberturas enfoques centrados en los derechos humanos. En especial, destacan los relacionados con la legislación sexual, los derechos de las trabajadoras sexuales y los informes de

la [International Planned Parenthood Federation](#), una organización internacional dedicada a la salud sexual y reproductiva.

Figura 6

(11/06/2024). Titular elaborado por Ana Requena Aguilar para eldiario.es sobre las comunidades autónomas que no dispensan gratis la píldora anticonceptiva de emergencia, incumpliendo la ley vigente.



Además de dar seguimiento a los debates parlamentarios en torno a las políticas del Ministerio de Igualdad, se analiza la implementación real de dichas políticas mediante curación documental. Con este fin, se recopilan enlaces a fuentes oficiales como el [BOE](#), pero también a fuentes corporativas, como el informe de la [Federación de Planificación Familiar \(SEDRA\)](#), que muestra cómo varía el acceso a la anticoncepción en las comunidades autónomas de España o Mujeres Juristas Themis, que aporta informes y recursos legales para visibilizar y combatir la desigualdad de género en el ámbito jurídico. Además, se propone incluir informes sindicales, como los de [CCOO](#), que analizan la igualdad de género en el ámbito laboral y la brecha salarial, como complemento para evaluar de manera más completa la implementación y el impacto de las políticas de igualdad.

Asimismo, se recopilan y curan algunas fuentes ciudadanas y de acceso público, como la cuenta de Twitter de [Elizabeth Duval](#), así como las entradas de Wikipedia sobre [Marina Subirats](#) y [Marisa Soleto](#), que aportan información contextual y biográfica relevante para el análisis de los debates y actores en el ámbito de la igualdad de género.

Esta misma lógica de selección y cuidado de fuentes resulta clave también para analizar los avances y retrocesos en derechos LGTBIQ+.

7. Cómo curar avances y retrocesos en derechos LGBTQ+

Las noticias de este bloque tratan sobre la situación de los derechos y las violencias que afectan a las personas LGBTQ+. Informan, por ejemplo, sobre retrocesos en la atención y protección de esta comunidad, especialmente tras el cambio de cartera en el Ministerio de Igualdad.

Figura 7

(09/02/2024). Titular elaborado por Patricia Reguero Ríos para El Salto sobre las dificultades de las personas trans extranjeras para actualizar sus documentos un año después de la Ley Trans en España.



En estos artículos también se incluyen enlaces a fuentes oficiales, como la [Ley 4/2023, de 28 de febrero, para la igualdad real y efectiva de las personas trans](#) y la garantía de los derechos de la comunidad LGBTQ+. Sin embargo, se da especial visibilidad a fuentes corporativas y ciudadanas que defienden estos derechos. Por ejemplo, las editoras de medios como Público o El Salto incorporan enlaces a [peticiones de activistas](#) LGBTQ+ que solicitaban revocar a Isabel García como directora del Instituto de las Mujeres, debido a sus declaraciones consideradas transfóbicas y su oposición a la Ley Trans.

Entre las fuentes corporativas, se incluyen enlaces a información publicada por entidades LGBTQ+, como la Federación Estatal de Lesbianas, Gais, Trans y Bisexuales (FELGTBI+) sobre el compromiso del Ministerio de Igualdad con el Pacto de Estado contra los discursos de odio hacia grupos vulnerables ([felgtbi.org](#)) y el portal Euforia, que ofrece documentación sobre el reconocimiento legal de identidades no binarias ([euforia.org.es](#)).

Figura 8

(25/01/2024). Titular elaborado por Candela Barro para Público sobre las campañas de odio contra mujeres trans en Twitter, según un estudio de la FELGTBI+ y 40dB.

Menú **Público** Inicia sesión Únete a Público

SOCIEDAD Educación Sanidad Migración Medio Ambiente Sucesos Gente

Las campañas de odio se ceban con las mujeres trans en Twitter a pesar de que el 86% de los usuarios apoya el Orgullo

Los usuarios lgtbifóbicos, la mayoría asociados a cuentas falsas, muestran más su odio en comentarios y lo hacen, sobre todo, contra las mujeres trans, según un estudio sobre la red social X de la FELGTBI+ junto a 40dB.

Una captura de pantalla de un tuit de la directora del Instituto de las Mujeres, Isabel García, durante una concentración convocada por la Federación Plataforma Trans, frente al Ministerio de Igualdad, a 9 de enero de 2024, en Madrid (España). **Alberto Ortega / Europa Press**

Candela Barro
Madrid - 25/01/2024 15:28 - Actualizado a 25/01/2024 16:06

Anuncio

¡Prueba Domino's Burger!
¡Pide la nueva pizza con sabor a burger!
¡Recoge en tienda dos o más desde 5,99 € c/u!

Domino's Pizza ¡Pruébala!

Por otro lado, se da visibilidad a fuentes ciudadanas del colectivo como la inclusión de enlaces a entrevistas con activistas LGTBQ+ como Alan Moreno, activista LGTBQ+ colombiano exiliado en España después de sufrir amenazas por ser un chico trans y por levantar la voz contra el acoso sexual y el machismo.

También se curan enlaces a otros medios fuera del ecosistema, seleccionados por las editoras de género. Por ejemplo, entre medios locales se incluyen diarios como *El Diario de Cantabria*, que reporta sobre la denuncia del Consejo de Estudiantes de la Universidad de Cantabria ante mensajes de LGTBfobia, machismo y odio dirigidos a estudiantes universitarios (eldiariocantabria.publico.es), así como *El Periódico*, que informa sobre el aumento de incidencias LGTBfóbicas durante 2023 (elperiodico.com).

Igualmente se enlazan medios internacionales, como *The New York Times*, que cubren casos de discriminación en universidades estadounidenses, por ejemplo, la situación de un estudiante no binario en Oklahoma (nytimes.com) y otros reportajes relacionados (nytimes.com).

La información de medios ajenos, tanto locales como internacionales, se cura cuidadosamente para complementar la cobertura propia, ofrecer diferentes perspectivas y aportar contexto

confiable sobre situaciones de discriminación y LGTBIfobia. Esto permite a las editoras de género proporcionar a los lectores un panorama más amplio y documentado de los problemas que afectan a la comunidad LGTBQ+ en distintos entornos y geografías.

8. Importancia de la labor documental de las periodistas de igualdad

Para concluir, es importante destacar que esta investigación evidencia la escasez de roles especializados en periodismo de género en España, con la presencia de editoras de género limitada a tres redacciones de prensa digital y solo seis periodistas con cargos directivos dedicados a las secciones de igualdad en prensa estatal, lo que refleja la limitada incorporación de estas profesionales en los medios.

La falta de recursos y reconocimiento se agrava en un contexto hostil, donde las periodistas de igualdad sufren hostigamiento, censura y cuestionamientos a su labor, especialmente desde sectores ultraconservadores, algunos medios digitales y la llamada “manosfera” en redes sociales. Por otro lado, la desconfianza hacia las fuentes oficiales y su condición de mujeres periodistas dificultan aún más su trabajo.

A pesar de ello, las periodistas españolas continúan realizando una curación rigurosa de fuentes, visibilizando temas de igualdad y cuestionando los discursos que niegan la desigualdad. Se trata de un periodismo documental que evidencia que las violencias de género siguen siendo centrales en la agenda mediática, a la vez que un 30 % de los artículos elaborados por las periodistas de igualdad en el periodo analizado abordan el crecimiento de discursos que niegan estas desigualdades y violencias.

En cuanto a la curación de fuentes, durante el primer semestre de 2024, se curaron más de dos mil documentos. Las profesionales recurrieron principalmente a fuentes oficiales, complementadas con fuentes corporativas, ciudadanos y medios. No obstante, un eventual gobierno ultraconservador, como los del PP o Vox, podría limitar la producción y confiabilidad de estas fuentes, lo que hace necesaria la diversificación de la curaduría para garantizar la continuidad futura de la cobertura sobre igualdad y violencias de género.

En este sentido, este estudio analiza las prácticas actuales del periodismo de género, señalando que, aunque se han consolidado avances, se trata de dinámicas coyunturales y persisten desafíos a futuro. Entre ellos destaca la necesidad de fortalecer redes documentales alternativas que aseguren un periodismo crítico y con enfoque de derechos, incluso en escenarios adversos, así como la creación de espacios seguros para trabajar y redes de apoyo con colegas, abogadas y profesionales de la salud, fomentando alianzas y rechazando la autocensura.

Aunque todos los medios seleccionados curan diversidad de fuentes documentales, se observan variaciones en sus enfoques: mientras las periodistas de *El País* e *Infolibre* priorizan fuentes oficiales sobre violencia, desigualdades y políticas públicas, la editora de *El Salto* da protagonismo a perspectivas propositivas, comunitarias y feministas, mientras que las profesionales de *El Periódico* y *Público* combinan fuentes oficiales, corporativas y ciudadanas, aunque con uso limitado de estas últimas.

La curación de estos contenidos se centra en enlaces retrospectivos, algunos con hasta diez años de antigüedad, lo que subraya la relevancia de la preservación digital. Esta práctica asegura el acceso a fuentes verificables y permite mantener la continuidad e integridad de los

análisis sobre desigualdades y discursos negacionistas. No obstante, un desafío recurrente es la mala conservación de los documentos digitales, ya que con el tiempo muchos enlaces se rompen, dificultando la verificación y la reconstrucción de los debates.

El enlazado de estas fuentes documentales se realiza mediante técnicas de resumen y contextualización, que facilitan la construcción de cronologías, la definición de conceptos clave y el análisis de estructuras de violencia, desigualdad y avances en derechos feministas y LGTBIQ+, incorporando también la agenda del Ministerio de Igualdad.

Asimismo, el seguimiento de la actividad del Ministerio, la implementación de sus políticas y normativas, y la correcta gestión de los fondos asignados —a nivel estatal, autonómico y municipal—, junto con el monitoreo de los avances y desafíos en derechos y violencias hacia la comunidad LGTBIQ+, contribuye a la rendición de cuentas institucional. Esto permite visibilizar logros, señalar carencias y exigir responsabilidades en la promoción de la igualdad.

En cuanto a la morfología de estas fuentes, predominan los sitios web, seguidos de fuentes secundarias y redes sociales, mientras que los blogs tienen una presencia mínima, reflejando un enfoque orientado a contenidos verificados.

Aunque todos los medios analizados se dirigen a audiencias progresistas, cada editora desarrolla un enfoque propio sobre feminismo y desigualdad de género. Mientras *Público* y *El Salto* recopilan más materiales sobre derechos trans, de las personas migrantes y participación de trabajadoras sexuales, otros medios incluyen también debates sobre la abolición de la prostitución y contrarios a la autodeterminación de género, manteniendo en todos los casos un enfoque proderechos y respetuoso con la diversidad de corrientes feministas.

Este periodismo de curación de fuentes cumple funciones informativa, educativa y estratégica, proporcionando herramientas y argumentos a quienes promueven la igualdad. Aunque no busca convencer a los negacionistas, fortalece las narrativas de derechos y respalda políticas, contenidos y reflexiones críticas. De este modo, las periodistas sitúan la agenda de igualdad en la esfera pública, consolidando una cobertura que visibiliza violencias y desigualdades, pero también propone soluciones y nuevos marcos de pensamiento.

A partir de esta investigación se abren nuevas líneas de estudio, entre las que destacan el análisis de las temáticas y la curación de fuentes de las periodistas en distintos periodos, países o contextos políticos, lo que permitiría enriquecer la comparación y comprender mejor las dinámicas globales del periodismo de género.

También sería relevante examinar la curación y el uso de fuentes documentales en medios que carecen de figuras especializadas en igualdad, así como en aquellos cercanos a posturas de derecha radical, para entender cómo estas condiciones afectan la cobertura, la selección de fuentes y el tratamiento de los temas. Cabe señalar que esta investigación se ha centrado únicamente en los contenidos publicados por seis periodistas con cargos de dirección dentro de la estructura editorial; por ello, sería interesante ampliar el estudio a otros trabajos elaborados por profesionales de las secciones de igualdad y, además, entrevistarlas para profundizar en su trabajo de selección de temas y curación de fuentes, complementando así la información presentada en este capítulo.

9. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

10. Referencias

- Amnistía Internacional. (2024). *Muteadas: el impacto de la violencia digital contra las periodistas*. Amnistía Internacional Argentina. https://amnistia.org.ar/wp-content/uploads/delightful-downloads/2024/10/Muteadas_InformeCompleto.pdf
- Arantegui, L., Romero, P., & Tamarit, J. M. (2024). Does the sex of judges matter? Analysis of sentencing in sex crimes in Spain. *Criminology & Criminal Justice*, First published online. <https://doi.org/10.1177/17488958241239785>
- Bailo, F., Meese, J., & Hurcombe, E. (2021). The institutional impacts of algorithmic distribution: Facebook and the Australian news media. *Social Media + Society*, 7(2). <https://doi.org/10.1177/20563051211024963>
- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Centro de Investigaciones Sociológicas. (2024, 15 de enero). *Avance de resultados del estudio 3428: Percepciones sobre la igualdad entre hombres y mujeres y estereotipos de género*. <https://www.cis.es/es/-/avance-de-resultados-del-estudio-3428-percepciones-sobre-la-igualdad-entre-hombres-y-mujeres-y-estereotipos-de-genero>
- 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>
- Europa Press. (2024, 4 de septiembre). Plataforma Trans insta a Redondo a desarrollar instrucciones y protocolos de la Ley Trans para “hacerla efectiva”. <https://www.europapress.es/sociedad/noticia-plataforma-trans-insta-redondo-desarrollar-instrucciones-protocolos-ley-trans-hacerla-efectiva-20240904131041.html>
- European Institute for Gender Equality. (2023). *Gender Equality Index 2023*. Publications Office of the European Union. <https://op.europa.eu/en/publication-detail/-/publication/73a85107-9b0b-11ee-b164-01aa75ed71a1>
- Feezell, J. T. (2017). Agenda setting through social media: The importance of incidental news exposure and social filtering in the digital era. *Political Research Quarterly*, 71(2), 482–494. <https://doi.org/10.1177/1065912917744895>
- Fletcher, R., & Nielsen, R. K. (2017). Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society*, 20(7), 2450–2468. <https://doi.org/10.1177/1461444817724170>

- García-García, J. (2025). Respuestas normativas para enfrentar el discurso del odio desde la experiencia europea. *European Public & Social Innovation Review*, 10, 1–12. <https://doi.org/10.31637/epsir-2025-1904>
- García-Mingo, E., & Díaz Fernández, S. (2023). Wounded men of feminism: Exploring regimes of male victimhood in the Spanish manosphere. *European Journal of Cultural Studies*, 26(6), 897–915. <https://doi.org/10.1177/13675494221140586>
- Guallar, J. (2014). Content curation en periodismo (y en documentación periodística). *Hipertext.net*, (12). <https://raco.cat/index.php/Hipertext/article/view/275781>
- 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., Franch, P., Cascón-Katchadourian, J., & Boté-Vericad, J.-J. (2025). The rise of curated newsletters in media: A case study of the New York Times. *Journalism Practice*. <https://doi.org/10.1080/17512786.2025.245733>
- Guallar, J., Pedraza, R., Pérez-Montoro, M., & Antón, L. (2021). Curación de contenidos en periodismo. Indicadores y buenas prácticas. *Revista Española de Documentación Científica*, 44(2). <https://doi.org/10.3989/redc.2021.2.1742>
- Guerrini, F. (2013). *Newsroom curators & independent storytellers: Content curation as a new form of journalism*. Reuters Institute.
- Iranzo-Cabrera, M., Figueras-Maz, M., & Mauri-Ríos, M. (2023). Journalistic self-regulation for equality: The role of gender editing in Spain. *Journal of Media Ethics*, 38(1), 2–15. <https://doi.org/10.1080/23736992.2022.2158336>
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.
- Levin, R. J. (2004). Measuring female genital functions—a research essential but still a clinical luxury? *Sexual and Relationship Therapy*, 19(2), 191–200. <https://doi.org/10.1080/14681990410001691406>
- Lopezosa, C., Guallar, J., Codina, L., & Pérez-Montoro, M. (2023). Curación de contenido y periodismo: revisión sistematizada exploratoria y visión experta. *Revista Mediterránea de Comunicación*, 14(1), 205–223. <https://doi.org/10.14198/MEDCOM.22810>
- Ministerio de Igualdad. (2023, 21 de noviembre). *Ana Redondo toma posesión como ministra de Igualdad*. <https://www.igualdad.gob.es/home/ana-redondo-toma-posesion-como-ministra-de-igualdad/>
- ONU Mujeres. (2025). *Los derechos de las mujeres bajo examen: 30 años después de Beijing*. <https://www.unwomen.org/es/digital-library/progress-of-the-worlds-women>
- Pan, R., García-Díaz, J. A., & Valencia-García, R. (2025). Spanish MTLHateCorpus 2023: Multi-task learning for hate speech detection to identify speech type, target, target group and intensity. *Computer Standards & Interfaces*, 94, 103990. <https://doi.org/10.1016/j.csi.2025.103990>
- Reuning, K., Whitesell, A., & Hannah, A. L. (2022). Facebook algorithm changes may have amplified local Republican parties. *Research & Politics*, 9(2). <https://doi.org/10.1177/20531680221103809>

- Spinetta, A. B. (2022). Editoras de género en Argentina. Avances y desafíos de un rol en construcción. *BiD: Textos Universitaris de Biblioteconomia i Documentació*, (49). <https://doi.org/10.1344/BiD2022.49.09>
- UNESCO. (2019). *The Chilling: Tendencias mundiales en la violencia en línea contra mujeres periodistas*. <https://www.unesco.org/es/articles/the-chilling-tendencias-mundiales-en-la-violencia-en-linea-contra-mujeres-periodistas>
- Virgili, J. R., & Puche, J. S. (2019). Medios de comunicación y opinión pública en España: Una aproximación desde la teoría de agenda setting. *Panorama Social*, 30, 27–39. <https://dialnet.unirioja.es/servlet/articulo?codigo=7207334>
- Waisbord, S. (2018). Truth is what happens to news. *Journalism Studies*, 19(13), 1866–1878. <https://doi.org/10.1080/1461670x.2018.1492881>

SEOVisFrame: buenas prácticas y método para evaluar la optimización SEO de las visualizaciones de datos narrativas

Rubén Alcaraz-Martínez

Universitat de Barcelona, España

<https://orcid.org/0000-0002-7185-0227>

Mari Vállez

Universitat de Barcelona, España

<https://orcid.org/0000-0002-3284-2590>

Raquel Escandell-Poveda

Universidad de Alicante, España

<https://orcid.org/0000-0002-8398-1873>

Diana Bajaña-Cedeño

Universitat de Barcelona, España

<https://orcid.org/0009-0008-7395-3028>

Alcaraz-Martínez, R., Vállez, M., Escandell-Poveda, R., & Bajaña-Cedeño, D. (2025). SEOVisFrame: buenas prácticas y método para evaluar la optimización SEO de las visualizaciones de datos narrativas. In J. Guallar, M. Vállez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 72-88). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicom.06.esp>

Resumen

Las visualizaciones de datos narrativas han adquirido un papel destacado en los cibermedios, facilitando la comunicación de información compleja de manera comprensible y atractiva. Sin embargo, su impacto y repercusión en buscadores depende en gran medida de su capacidad para aparecer en los resultados de búsqueda. Bajo esta premisa, en este capítulo se examinan los factores de posicionamiento clave para estos productos informativos y se propone un conjunto de buenas prácticas, así como una metodología para evaluar la optimización SEO a partir de una lista de indicadores a la que denominamos SEOVisFrame.

Palabras clave

Visualizaciones de datos narrativas; Periodismo de datos; Cibermedios; Search Engine Optimisation (SEO); Visibilidad web.

1. Introducción

En la era de la información, el periodismo de datos ha emergido como una potente herramienta para estructurar, analizar y narrar historias basadas en grandes volúmenes de información. Esta especialidad periodística, que combina técnicas de análisis de datos con narrativas visuales, permite no solo desentrañar la complejidad de los datos sino también hacerlos accesibles y comprensibles para la ciudadanía (Córdoba-Cabús, 2020). Su evolución ha sido impulsada por el auge de la digitalización y la disponibilidad de datos abiertos, marcando una transformación en la práctica del periodismo tradicional cada vez más datificado (Radcliffe y Lewis, 2019).

Si bien su origen se remonta a la segunda mitad del siglo XX, la práctica moderna se consolida a finales de la década del 2000 y principios de 2010 (Ferrerías Rodríguez, 2015), impulsada por la proliferación de datos abiertos disponibles, especialmente, procedentes de las administraciones públicas, la disponibilidad de nuevas herramientas tecnológicas centradas en el análisis de datos masivos y la creación de visualizaciones interactivas, junto a un contexto mediático en el que las audiencias demandaban nuevas narrativas innovadoras. Algunos hitos en este sentido son el resultado de la investigación de los denominados *Papeles de Panamá* (2016) desarrollada por el Consorcio Internacional de Periodistas de Investigación o aplicaciones derivadas de esta modalidad como las publicadas por la Fundación Civio (*Dónde van mis impuestos* o el *Indultómetro*) (Ferrerías Rodríguez, 2015).

La integración de departamentos o unidades especializadas en el periodismo de datos dentro de los cibermedios es una realidad desde hace más de una década (Peiró y Guallar, 2013). Durante la pandemia de COVID-19, el periodismo de datos desempeñó un papel crucial, convirtiéndose en un intermediario esencial entre las instituciones y la sociedad. La abundancia de cifras y su falta de coherencia destacaron la importancia de este enfoque para organizar, interpretar y comunicar información relevante (Córdoba-Cabús et al., 2020).

Entre los trabajos previos que se han centrado en el análisis y evaluación de estos productos informativos, encontramos la propuesta de Alexander y Vetere (2011), quienes plantean una lista de parámetros para asegurar la calidad del contenido en el periodismo de datos tales

como la confianza en los datos presentados, la integración del factor sorpresa, la transparencia en los procesos, la identificación de un líder, el uso de un lenguaje accesible y la posibilidad de acceso a los datos originales. El trabajo de Córdoba-Cabús (2020) se centra en el análisis de algunas de las visualizaciones de datos con mayor reconocimiento internacional surgidas del sector periodístico, con el objetivo de sintetizar una lista de elementos comunes en esta práctica periodística. Sus variables para el análisis son la temática, el estilo narrativo, el tipo de historia, la dimensión, los elementos de foco, el propósito, la manera de comunicar, la fuente, el tipo de análisis realizado, las formas y la ratio de visualizaciones. La calidad y la efectividad de estos productos narrativos dependen de parámetros como la transparencia, la selección y el tratamiento de las fuentes, así como la adecuación e innovación en las formas seleccionadas para presentar la información (Córdoba-Cabús, 2020).

Desde finales de la década de 1990 y principios de los 2000, cuando tanto los principales periódicos internacionales como los nacionales lanzaron sus respectivas versiones digitales, la optimización para motores de búsqueda (SEO) se ha convertido en una herramienta indispensable para el periodismo, ya que influye directamente en la visibilidad y accesibilidad de los contenidos informativos. Dado que los motores de búsqueda son el principal punto de acceso para encontrar información en la web, la capacidad del contenido periodístico para posicionarse en los primeros resultados es crucial para alcanzar a una audiencia más amplia. Las prácticas de SEO no solo mejoran la capacidad de descubrimiento de los artículos, sino que también garantizan que el contenido se presente de una forma que responda a las demandas cambiantes tanto de los algoritmos como de los usuarios. Dado el carácter competitivo del panorama mediático en línea, contar con estrategias de SEO eficaces resulta fundamental para que las organizaciones periodísticas mantengan su relevancia y logren implicar a las audiencias en un entorno digital cada vez más saturado.

El número, factores concretos de posicionamiento y el peso que presenta cada uno de ellos en los algoritmos de los buscadores son, en gran medida, incógnitas con las que los profesionales del sector se enfrentan a diario. Este conocimiento tácito en el sector se ha difundido fundamentalmente a través de algunas monografías especializadas (Maciá Domene, 2019; Enge et al., 2023; Lewandowski, 2023; Vicente et al., 2024), así como a través de circuitos menos académicos (Semrush, 2024; Dean, 2024) y, en menor medida, en publicaciones científicas (Chotikitpat et al., 2015; García-Carretero et al., 2016; Lopezosa et al., 2018; Ziakis et al., 2019; Almukhtar et al., 2021), algunas de ellos centrados específicamente en cibermedios (Dick, 2011; Giomelakis y Veglis, 2015; Lopezosa et al., 2019; 2020; Giomelakis, 2023). Al conocimiento anterior, es imprescindible sumarle la documentación técnica publicada por los buscadores, entre los cuales, Google es el más prolífico y abarca desde cuestiones generales relacionadas con los fundamentos de los buscadores (Google, 2024a), como con cuestiones mucho más específicas como la optimización de los sitios en el marco de las directrices EEAT (Experiencia, Conocimiento, Autoridad y Confianza) (Google, 2024b) o recomendaciones para la optimización de imágenes desde un punto de vista SEO (Google, 2024c).

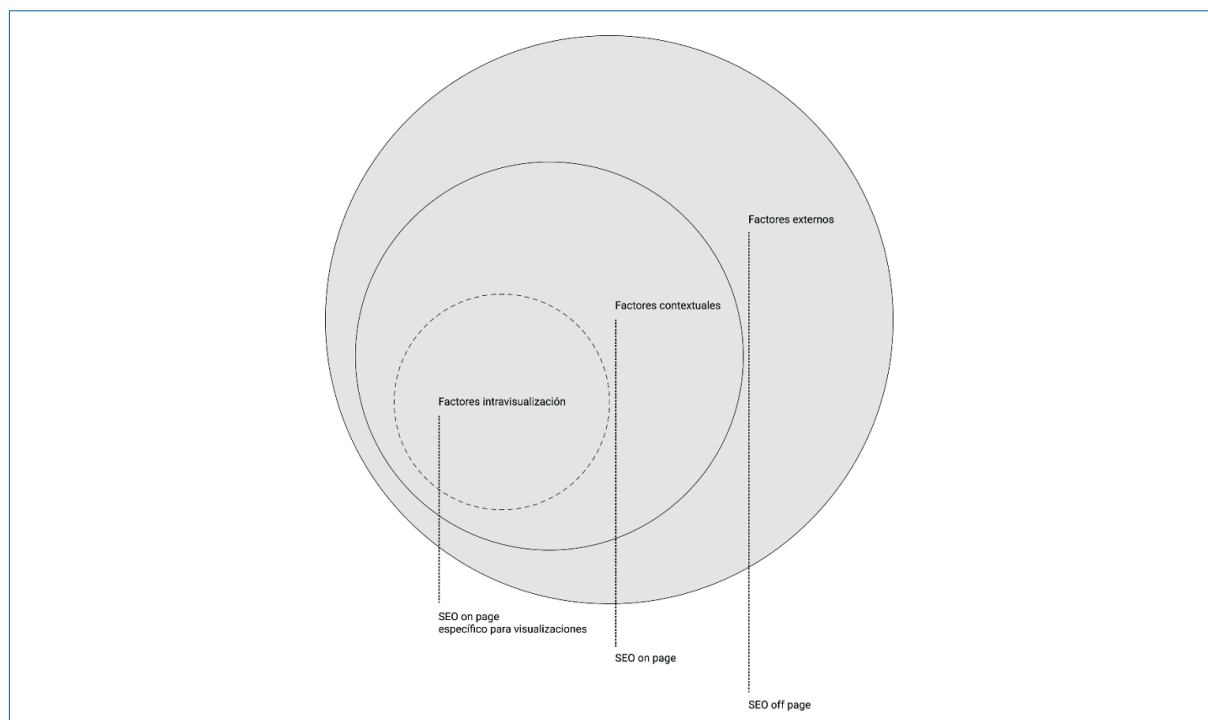
Los artículos publicados bajo el paraguas del periodismo de datos son productos complejos que incorporan múltiples formas de visualización de datos, como tablas, gráficos estadísticos interactivos, mapas, infografías y animaciones multimedia enriquecidas, acompañados de narrativas que suelen estructurarse en forma de relato (Domínguez, 2016). Este enfoque ha dado lugar a formatos como el *long-form journalism* y a técnicas como el *scrollytelling* (Seyser y Zeiller, 2018). Por tanto, abordar la optimización SEO de este tipo de productos requiere el análisis de diversas dimensiones (elementos textuales y gráficos), cada una de las cuales

está vinculada a diferentes factores de posicionamiento, así como a su visibilidad en distintos motores de búsqueda (por ejemplo, Google Search o Google Imágenes).

La forma más habitual de clasificación de los factores SEO es la que responde a la naturaleza del control y la influencia que se puede ejercer sobre ellos. De acuerdo con esta clasificación, es posible distinguir entre factores internos (SEO *on page* u *on site*) y factores externos (SEO *off page* u *off site*) (Lewandowski, 2023). Por su parte, el SEO *on page* puede subdividirse en SEO técnico y SEO de contenidos (Escandell-Poveda et al., 2021), dos áreas fundamentales, pero distintas que afectan el rendimiento y la visibilidad de un sitio web.

Figura 1

Clasificación de los factores de posicionamiento para productos de periodismo de datos narrativos.



El SEO técnico se centra en la optimización de ciertos factores internos relacionados con el rastreo y la indexación de las páginas de un sitio web, así como en la compatibilidad técnica del sitio con dispositivos móviles, la gestión de redirecciones o la optimización de su rendimiento y velocidad de carga (Krstić, 2019; Vicente et al., 2024).

Por otro lado, el uso del protocolo HTTPS es un factor de posicionamiento desde 2014 (Google, 2014). Al implementar HTTPS en un dominio, es importante asegurarse de que todas las páginas, así como los recursos necesarios para mostrar su contenido —como imágenes y otros elementos— se carguen utilizando este protocolo. Google (2014) recomienda el uso de URLs relativas para los recursos que se encuentran en el mismo dominio seguro, con el fin de garantizar el uso del protocolo correcto. Esto también puede mejorar el rendimiento en caso de que las URLs absolutas estén mal configuradas (por ejemplo, si se utiliza HTTP en lugar de HTTPS). En estos casos, los navegadores pueden resolver las URLs relativas más rápidamente, ya que se evita la posibilidad de redireccionamientos.

La integración de datos estructurados de Schema.org ha sido otro elemento clave del SEO técnico en muchos sitios web desde su introducción por parte de Google en 2011. Los datos estructurados constituyen un formato estandarizado para proporcionar información sobre una página y su contenido. Se implementan mediante marcado dentro del propio código de la página a la que se refiere la metainformación. La inclusión de este marcado mejora la indexación de noticias y aumenta la precisión de los resultados de búsqueda, además de reducir la ambigüedad en la interpretación del contenido por parte de los algoritmos de búsqueda (Salem et al., 2025). Cuando una noticia incluye datos estructurados, Google también puede utilizar esa información para mostrar un fragmento enriquecido (*rich snippet*) del contenido. Aunque la literatura especializada indica que la posición en el ranking sigue siendo el factor más determinante en la percepción de relevancia, en ciertos casos la riqueza de los fragmentos puede captar la atención del usuario (Marcos et al., 2015). Es decir, estos elementos pueden influir de manera significativa en el comportamiento y la toma de decisiones de los usuarios (Rodas et al., 2016). Los datos estructurados deben representar fielmente el contenido principal de la página, ser sintácticamente correctos, asegurar que el contenido referenciado sea visible para los usuarios y cumplir con las directrices establecidas para cada tipo de dato estructurado específico (Google, 2025a). En este caso, el tipo de datos estructurados más relevante es *NewsArticle*. No obstante, los datos estructurados para imágenes también pueden ser útiles, ya que permiten que Google Imágenes muestre detalles adicionales como la autoría, los derechos o la información de atribución.

El buen rendimiento de las páginas resulta también crucial en un contexto cada vez más mediado por dispositivos móviles. Para valorar este extremo, Google presentó en 2020 e incorporó como factor en su algoritmo en 2021, las denominadas *Core Web Vitals* (CWV). Se trata de un conjunto de métricas diseñadas para medir el rendimiento de las páginas web (Maciá Domene, 2019; Alcaraz Martínez, 2022; Vicente et al., 2024), entre las que se incluyen el LCP (*Largest Contentful Paint*), el INP (*Interaction to Next Paint*) y el CLS (*Cumulative Layout Shift*).

El LCP, o tiempo de carga del mayor elemento de texto o imagen visible en la ventana gráfica, es un indicador del tiempo de carga percibido. Por su parte, el INP es un indicador que mide la capacidad de respuesta de las páginas frente a las interacciones de los usuarios (clics o pulsaciones de teclas). Finalmente, el CLS es un indicador de la estabilidad visual y describe el desplazamiento máximo al que se somete la disposición de los elementos visibles en la ventana gráfica durante la carga de la página. La narración mediante visualizaciones de datos presenta importantes desafíos en relación con las CWV, ya que a menudo integra múltiples imágenes complejas en una sola página. La optimización de estos archivos —mediante la selección del formato de mapa de bits adecuado, la compresión, las dimensiones, etc.— o la creación de imágenes vectoriales rastreables y compatibles con los motores de búsqueda puede resultar compleja.

Por su parte, el SEO de contenidos intenta dar respuesta a las necesidades de los usuarios mediante la creación de textos y otros tipos de contenidos (imágenes, audiovisuales, interactivos...) optimizados para ser relevantes de acuerdo con las palabras clave y la intención de búsqueda planteada por los usuarios (Alcaraz Martínez, 2022).

Los factores *on page* más importantes para cualquier página son igualmente importantes para las visualizaciones narrativas objeto de estudio. En este sentido, la aparición de la palabra clave en la metaetiqueta *title*, la metadescripción, el URL, los encabezados de niveles h1-h3

o en el mismo cuerpo de la página con una densidad adecuada, son factores fundamentales (Chotikitpat et al., 2015; Almukhtar et al., 2021; Maciá Domene, 2019; Lewandowski, 2023; Dean, 2024; Google, 2024a; Semrush, 2024).

Aunque ya no presenta la importancia que tuvo en el pasado, la incorporación de palabras clave en la etiqueta <title> sigue siendo una señal relevante en el SEO *on page*. Por otro lado, si bien las meta descripciones fueron históricamente un factor importante de posicionamiento, su influencia ha disminuido. En los últimos años, Google ha generado cada vez más los fragmentos directamente a partir del contenido de la página. No obstante, puede seguir utilizando este metaelemento si ofrece a los usuarios un resumen más preciso de la página que el contenido extraído directamente del sitio.

Todas las páginas deben incluir una etiqueta de encabezado <h1> con un contenido estrechamente alineado con la etiqueta <title>. A diferencia de la etiqueta <title>, la etiqueta <h1> no contribuye a la generación de fragmentos en las páginas de resultados de los motores de búsqueda (SERP). Por ello, estrategias como la incorporación de la palabra clave de marca (por ejemplo, el nombre del medio de comunicación) se aplican comúnmente en los metatítulos, pero no en los encabezados principales. Optimizar la etiqueta <h1> implica redactar títulos de las historias que se correspondan con los términos de búsqueda para los cuales se busca visibilidad. Esto no excluye la creatividad, ya que la viralidad mediática y las tendencias informativas pueden hacer que títulos como *The Panama Papers* resulten altamente efectivos.

Dentro del elemento <main> de una página, se pueden incluir varios niveles de encabezados, que van desde <h2> hasta <h6>. Desde una perspectiva SEO, los que se encuentran más arriba en la jerarquía, especialmente las etiquetas <h2> y <h3>, tienen mayor relevancia. Según Mueller (2015), coordinador de Google Search Relations, "*heading tags in HTML help us to understand the structure of the page*". En tanto que una misma página puede intentar posicionarse para múltiples palabras clave, estructurar el contenido en bloques temáticos, cada uno con sus propios encabezados específicos, es una práctica SEO ampliamente adoptada (Maciá Domene, 2019).

Las imágenes deben marcarse en el código fuente de forma que puedan ser fácilmente indexadas por los motores de búsqueda. Las imágenes en formato de mapa de bits que se desea indexar se incluyen en el código HTML mediante elementos , <figure> o <picture>, pero nunca como imágenes de fondo en CSS, lo cual es bastante común al implementar efectos *Parallax* típicos en narrativas visuales. En términos de relevancia, deben incluirse palabras clave en el texto alternativo, el nombre del fichero, el pie de imagen y en el contenido contextual (Chotikitpat et al., 2015; Maciá Domene, 2019; Lewandowski, 2023; Semrush, 2024). El texto alternativo (*alt*) es el atributo más importante a través del cual proporcionar información sobre las imágenes a los buscadores (Google, 2024c), al mismo tiempo que mejora la accesibilidad para los usuarios que no pueden verlas. No es necesario añadir la palabra clave principal de la página en todas las imágenes, y debe evitarse la práctica de repetir de forma poco natural y excesiva el mismo término (*keyword stuffing*). La selección de imágenes apropiadas para ilustrar el contenido es un primer paso esencial. Cuando se hace correctamente, todas las imágenes estarán alineadas con el tema del artículo, lo que refuerza su relevancia general y permite la creación de textos alternativos naturales y bien optimizados que cumplan tanto con fines SEO como de accesibilidad (Alcaraz Martínez, 2024). Google recomienda utilizar nombres de fichero descriptivos en lugar de genéricos, incluso localizándolos para las diferentes versiones

idiomáticas del contenido (Google, 2024c). Aunque se desconoce el impacto exacto de este factor, su consideración como una práctica recomendada resulta beneficiosa por otros motivos (mantenimiento, localización de los ficheros, etc.). Cada imagen debe utilizarse en páginas que se correspondan con su contenido temático, y deben situarse dentro de la página cerca del texto relacionado con ellas (Lopezosa et al., 2018). Aunque no existen directrices oficiales sobre la proximidad del contenido textual relacionado con las imágenes, la lógica sugiere tener en cuenta los párrafos inmediatamente anteriores y posteriores a cada imagen. Otro elemento importante, dado su vínculo semántico directo con la imagen, es el elemento `<figcaption>`. El elemento `<figcaption>` dentro de un agrupamiento `<figure>` asocia una o más imágenes con una descripción visible (pie de imagen), de la siguiente manera:

Figura 2

Ejemplo de uso del elemento `<figcaption>`.

```
<figure>
  
  <figcaption>Figure 1: Figcaption.</figcaption>
</figure>
```

También puede resultar de interés añadir metadatos IPTC a las imágenes, una técnica que Google utiliza para extraer ciertos datos de éstas para su uso en Google Imágenes (Google, 2025b). Los metadatos IPTC se incrustan directamente en la imagen, lo que permite que tanto la imagen como sus metadatos se mantengan intactos al utilizarse en diferentes páginas. A diferencia de los datos estructurados, los metadatos IPTC solo necesitan añadirse una vez por imagen, independientemente del número de páginas en las que aparezca. Las propiedades que Google extrae incluyen información sobre derechos de autor (*Copyright Notice*), la autoría (*Creator*), la forma de proporcionar citar la imagen (*Credit Line*), la forma cómo se generó la imagen digital (*Digital Source Type*), el enlace al propietario de los derechos de autor (*Licensor URL*) y un URL en el que se especifiquen los usos previstos (*Web Statement of Rights*).

Según Google (2024c), *"high-quality photos appeal to users more than blurry, unclear images"*. Además, las imágenes nítidas resultan más atractivas en las miniaturas de los resultados y pueden aumentar la probabilidad de recibir visitas por parte de los usuarios. Es importante encontrar un equilibrio entre calidad y peso para cumplir con las métricas CWV mencionadas anteriormente y satisfacer tanto las expectativas como la experiencia del usuario.

En 2020, Google introdujo la metaetiqueta *max-image-preview*, que permite a los editores tener un mayor control sobre cómo se muestran las imágenes de sus sitios en Google Discover. Cuando se añade la metaetiqueta *max-image-preview:large* en el encabezado de una página web, se indica a Google que puede mostrar las imágenes del sitio en formatos más grandes. Esta función mejora la presentación visual del contenido en plataformas como Discover, ofreciendo una experiencia de usuario más atractiva.

Las imágenes SVG pueden integrarse en una página HTML de dos formas: como un fichero externo referenciado mediante el elemento `<img>`, o incorporando directamente el código SVG dentro del propio documento HTML. El primer método no difiere del uso habitual de las imágenes en formato de mapa de bits dentro del mismo elemento y requiere la inclusión de un texto alternativo, junto con el resto de los elementos mencionados anteriormente.

Incorporar el código SVG directamente en la página ofrece varias ventajas, como la posibilidad de manipular los elementos SVG mediante *scripts* y modificar su apariencia a través de CSS. Además, ciertos elementos y atributos SVG contribuyen a la accesibilidad, como `<desc>`, `<title>` y `<text>`, ya que proporcionan descripciones textuales útiles para personas con discapacidad visual. Estos elementos también pueden ser indexados por los motores de búsqueda (Ferraz, 2017).

Los elementos `<desc>` y `<title>` están específicamente diseñados para describir imágenes. Aunque no se muestran visualmente en la página. Por otro lado, el elemento `<title>` puede aparecer como un mensaje emergente (*tooltip*) al pasar el cursor por encima, aunque este comportamiento varía según el navegador. El elemento `<text>` define un componente de texto gráfico dentro del SVG. Puede incluir atributos que controlen propiedades visuales como la dirección del texto, su posicionamiento o el relleno (*fill*). A diferencia de `<desc>` y `<title>`, `<text>` se representa de forma visible en la página, y su ubicación depende tanto de su posición en el código como de las coordenadas X e Y definidas dentro del contenedor SVG.

Figura 3

Gráfico de barras vectorial en formato SVG con Código XML que incluye los elementos `<title>`, `<description>`, y `<text>`.

```
<svg width="450" height="300" viewBox="0 0 450 300" role="img" aria-labelledby="chartTitle chartDesc">
  <title id="chartTitle">Children at risk of poverty or social exclusion in Spain</title>
  <desc id="chartDesc">This bar chart displays the percentage of children at risk of poverty or social exclusion in Spain from 2019 to 2022. The x-axis represents the years, and the y-axis represents the percentage.</desc>
  <text x="225" y="290">Year</text>
  <text x="20" y="150">Percentage (%)</text>
  <rect class="bar" x="70" y="120" width="60" height="150"></rect>
  <rect class="bar" x="160" y="100" width="60" height="170"></rect>
  <rect class="bar" x="250" y="130" width="60" height="140"></rect>
  <rect class="bar" x="340" y="110" width="60" height="160"></rect>
  <text class="value-label" x="100" y="115">33.0%</text>
  <text class="value-label" x="190" y="95">35.8%</text>
  <text class="value-label" x="280" y="125">31.1%</text>
  <text class="value-label" x="370" y="105">34.5%</text>
  <text x="100" y="280">2019</text>
  <text x="190" y="280">2020</text>
  <text x="280" y="280">2021</text>
  <text x="370" y="280">2022</text>
</svg>
```

En el ejemplo anterior, también se incluye el atributo `aria-labelledby`, recomendado por Google (2024c) y recogido en la especificación W3C WAI-ARIA. Este atributo permite relacionar un elemento —en este caso, la visualización a través de su etiqueta `<svg>`— con una descripción textual. El valor del atributo `aria-labelledby` hace referencia a los elementos HTML (en este caso, `<desc>` y `<title>`) que contiene dicha descripción.

Publicadas originalmente en 2013, pero actualizadas periódicamente hasta el día de hoy (Google, 2024b), las *Search quality rating guidelines* son el documento de referencia para entender qué busca Google al evaluar la calidad del contenido en una página web a partir de los criterios que emplean sus revisores. Estas directrices introducen el concepto de EEAT (*Experience, Expertise, Authority, and Trustworthiness*), ligado no solo a cuestiones como la capacidad de las páginas para satisfacer la intención de búsqueda o su compatibilidad con dispositivos móviles, sino también con diversos elementos asociados a la experiencia, conocimiento, autoridad y confianza que transmite un sitio mediante su contenido, especialmente en medios de comunicación y sitios web de categorías sensibles que implican cuestiones económicas o relacionadas con la salud.

La experiencia (*experience*) directa del autor y su familiaridad con el tema sobre el que escribe constituye el primer factor clave. Ser un medio de comunicación prestigioso, vinculado al pe-

periodismo de datos o de investigación, es una forma de cumplir este criterio. El conocimiento (*expertise*) hace referencia al nivel de especialización tanto del autor como de la publicación en un tema específico, y refleja la trayectoria acumulada que los califica como especialistas. La autoridad (*authority*), en este contexto, se refiere a la reputación tanto del medio (y su sitio web) como del autor del contenido. Esta autoridad se evalúa a partir de factores como el volumen de búsquedas asociadas a palabras clave de marca (relacionadas con el medio y el autor), los enlaces entrantes y las menciones en redes sociales. Finalmente, la fiabilidad (*trustworthiness*) determina el grado de confianza, honestidad y precisión de la información publicada. Se valora a través de aspectos técnicos como el uso de HTTPS, la existencia de páginas de información empresarial y legal sobre el medio, páginas de autor dedicadas con una biografía relacionada con su área de especialización y enlaces a sus contenidos publicados y perfiles profesionales en redes sociales. En el contexto del periodismo de datos, también es fundamental proporcionar información sobre el origen o fuente de los datos, así como la metodología empleada.

Figura 4

El Nacional.cat tiene una sección (/firmas/) dedicada a sus editores. Cada artículo enlaza a sus respectivos autores. Estos perfiles incluyen una biografía, enlaces a sus perfiles sociales, etiquetas relacionadas con sus publicaciones y enlaces a páginas de otros autores del medio relacionados.

The image shows a screenshot of a webpage from El Nacional.cat. The top part features a green header with the word "VINO" and a large title: "¿Cuál es la bebida alcohólica más consumida? Las tendencias del vino, la cerveza y los licores". Below the title, it says "Los datos sobre consumo de alcohol en el Estado" and "Victor Antich". A photo of a man speaking is on the left. Below the article, there is a profile for "Víctor Antich" with a photo and a bio. The bio mentions his background as a baker and his current role as a coordinator at La Gourmeteria. To the right of the bio are tags: "GOURMETERIA", "OPINIÓN", "RECOMENDACIONES", and "RESTAURANTES". Below the tags is a section "Otros autores relacionados" listing Pep Antoni Roig, Ada Parellada, Rosa Molinero Trias, and Jordi Tubella. A red arrow points from the article title to the author's name in the profile.

VINO

¿Cuál es la bebida alcohólica más consumida? Las tendencias del vino, la cerveza y los licores

Los datos sobre consumo de alcohol en el Estado

Victor Antich
Foto: Carlos Baglietto
Barcelona, Martes, 10 de diciembre de 2024, 10:34
Actualizado: Jueves, 27 de febrero de 2025, 13:05
Tiempo de lectura: 2 minutos

Víctor Antich

De La Seu d'Urgell, cosecha de 1965. Con formación de panadero y un pasado en la tribu mod catalana. Vengo de padres librerías y desde siempre he estado muy vinculado al mundo de la cultura, formando parte, durante un buen montón de años, del equipo del Gran Teatre del Liceu. Actualmente, soy el coordinador editorial de *La Gourmeteria*, la sección gastronómica de *ElNacional.cat*, donde también me encontraréis escribiendo y descubriéndolos algunos de los mejores restaurantes del país. También...

MÁS

¿Sobre qué escribo?

GOURMETERIA OPINIÓN
RECOMENDACIONES RESTAURANTES

Otros autores relacionados
Pep Antoni Roig
Ada Parellada
Rosa Molinero Trias
Jordi Tubella

Finalmente, el SEO *off page* se centra en factores que no es posible gestionar a través de los contenidos o la configuración técnica del sitio, pero que influyen en su autoridad y posicionamiento. Entre estos factores destacan la cantidad de enlaces entrantes recibidos (Chotikitpat et al., 2015; Lopezosa et al., 2019), ciertas métricas de autoridad aplicables a nivel de dominio o página (*domain authority* y *page authority*) (García Carretero et al., 2016), el volumen de tráfico (Krstić, 2019) o algunas otras generadas como respuesta a la interacción con los usuarios (porcentaje de clics por impresión en la página de resultados de los buscadores, tiempo de permanencia en página, *pogo sticking*, rebote...) (Maciá Domene, 2019).

Es fundamental tener en cuenta la influencia de los enlaces externos y la diversidad de dominios de referencia al evaluar el potencial de posicionamiento de una página, dado su impacto en los algoritmos de los buscadores. Estos factores actúan como indicadores de autoridad y relevancia, y contribuyen de forma significativa a la visibilidad en los resultados de búsqueda. La autoridad del dominio, la autoridad de la página y el tráfico mensual estimado se han identificado como métricas clave que ayudan a explicar por qué algunas páginas logran buenos posicionamientos a pesar de no destacar en ciertos aspectos de optimización *on page*. Aunque métricas como la autoridad del dominio o de la página no son factores de posicionamiento directos, sí que ofrecen información valiosa sobre el potencial de rendimiento de un sitio. Un tráfico estimado elevado puede ser indicativo de la capacidad de una página para atraer usuarios, lo que a su vez puede correlacionarse con un buen posicionamiento del medio. Al comparar el posicionamiento de dos recursos periodísticos que compiten por las mismas palabras clave, estas métricas pueden ayudar a explicar el mayor potencial de posicionamiento de ciertos contenidos o medios frente a otros.

2. Propuesta

SEOVisFrame es una propuesta basada en un conjunto de indicadores para evaluar visualizaciones de datos narrativas. Estos indicadores abarcan tanto factores de posicionamiento SEO como otros elementos que pueden mejorar de forma indirecta la visibilidad y el potencial de posicionamiento de este tipo de productos informativos. Pueden entenderse tanto como un conjunto de buenas prácticas, así como una serie de criterios para evaluar contenidos ya existentes.

Como resultado de esta propuesta, se ha identificado un total de 34 factores que se han clasificado de la siguiente manera: SEO de contenido (7 factores); SEO técnico (7 factores); EEAT (5 factores); SEO para componentes gráficos (9 factores); y SEO *off page* (6 factores). Para evaluar cada uno de ellos se utiliza una escala Likert cualitativa de tres puntos, codificada de 0 a 2 para facilitar el cálculo de medias y permitir la comparación entre medios o productos de un mismo medio.

Table 1

Indicadores para la evaluación de la optimización SEO de las visualizaciones de datos narrativas.

ID	Factor	Descripción	Ámbito de aplicación	Método de análisis / Puntuación
1	Palabra clave en el título	La palabra clave objetivo aparece en la metaetiqueta <title>.	Todas	Revisión del <title> mediante el informe de auditoría de SEOquake. Puntuaciones 0: la página no tiene <title> o no aparecen las palabras clave objetivo 1: el <title> es parcialmente relevante 2: el <title> es completamente relevante
2	Palabra clave en la meta-descripción	La palabra clave objetivo aparece en la metadescripción.	Todas	Revisión de la metadescripción mediante el informe de auditoría de SEOquake. Puntuaciones 0: la página no tiene metadescripción o no aparecen las palabras clave objetivo 1: la metadescripción es parcialmente relevante 2: la metadescripción es completamente relevante

ID	Factor	Descripción	Ámbito de aplicación	Método de análisis / Puntuación
3	Palabra clave en el URL	La palabra clave objetivo aparece en el URL.	Todas	Revisión manual del URL. Puntuaciones 0: En el URL no aparecen las palabras clave objetivo. 1: El URL es parcialmente relevante. 2: El URL es completamente relevante.
4	Palabra clave en la etiqueta <h1>	La palabra clave objetivo aparece en la etiqueta <h1>.	Todas	Revisión de la metadescripción mediante el informe de auditoría de SEOquake. Puntuaciones 0: La página no tiene una etiqueta <h1> o no aparecen las palabras clave objetivo. 1: La etiqueta <h1> es parcialmente relevante. 2: La etiqueta <h1> es completamente relevante.
5	Palabra clave en las etiquetas <h2> / <h3>	La palabra clave objetivo aparece en las etiquetas <h2> / <h3>.	Todas	Revisión de los encabezados mediante el informe de auditoría de SEOquake. Puntuaciones 0: La página no tiene etiquetas <h2> / <h3> o no aparecen las palabras clave objetivo. 1: Las etiquetas <h2> / <h3> son parcialmente relevantes. 2: Las etiquetas <h2> / <h3> son totalmente relevantes.
6	Uso de términos semánticos latentes en el contenido	En el contenido de la página se incorporan variantes de la palabra clave (sinónimos, términos relacionados semánticamente...	Todas	Revisión manual del contenido. Puntuaciones 0: en la página no aparecen variantes de la palabra clave o términos relacionados. 1: En la página aparecen algunos términos relacionados, pero son poco relevantes. 2: En la página aparecen múltiples variantes y términos relacionados con la palabra clave.
7	Uso de HTTPS	La página y sus recursos se encuentran bajo el protocolo HTTPS.	Todas	Revisar manualmente la dirección URL. Puntuaciones 0: No usa https. 1: Usa https parcialmente. 2: Usa https totalmente.
8	Uso de Schema.org	La página cuenta con marcado Schema.org.	Todas	Revisión el uso de Schema.org mediante el informe de auditoría de SEOquake. Puntuaciones 0: No usa schema.org 1: Usa schema .org pero el tipo utilizado no es .relevante o presenta algún error de código. 2: Usa un tipo schema.org adecuado y válido.
9	Fuente de los datos	La página incorpora una referencia a la fuente de datos.	Todas	Revisión manual del contenido. Puntuaciones 0: No proporciona la fuente de los datos. 1: Sí proporciona la fuente de los datos.
10	Enlace a la fuente de los datos	La página incorpora un enlace a la fuente de datos.	Todas	Revisión manual del contenido. Puntuaciones 0: No proporciona un enlace a la fuente de los datos. 1: Sí proporciona un enlace a la fuente de los datos.
11	Autoría del contenido	La página incorpora una referencia a los autores de la pieza.	Todas	Revisión manual del contenido. Puntuaciones 0: No proporciona información de autoría. 1: Sí proporciona información de autoría.

ID	Factor	Descripción	Ámbito de aplicación	Método de análisis / Puntuación
12	Enlace a la página del autor	La página incorpora enlaces a páginas con información de los autores.	Todas	Revisión manual del contenido. Puntuaciones 0: No proporciona un enlace a una página con información del autor. 1: Sí proporciona un enlace a una página con información del autor.
13	Metodología	La página incorpora información sobre la metodología utilizada en la investigación.	Todas	Revisión manual del contenido. Puntuaciones 0: No proporciona información sobre la metodología. 1: Sí proporciona información sobre la metodología.
14	Uso de tecnologías compatibles	Las visualizaciones de datos y los textos que las acompañan se implementan con tecnologías fáciles de rastrear e indexar.	Todas	Revisión manual del código fuente de la página. Puntuaciones 0: Las tecnologías implementadas no son compatibles con los buscadores. 1: Las tecnologías implementadas son parcialmente compatibles con los buscadores. 2: Todas las tecnologías implementadas son compatibles con los buscadores.
15	Palabra clave en el texto alternativo	Las imágenes cuentan con un texto alternativo que contiene alguna de las palabras clave objetivo.	Imágenes en formato de mapa de bits	Revisión manual de los textos alternativos. Puntuaciones 0: Ninguna o solo algunas de las imágenes cuentan con un texto alternativo o es un texto alternativo que no tiene ninguna relación con el contenido. 1: Las imágenes cuentan con un texto alternativo, pero es poco relevante. 2: Las imágenes cuentan con un texto alternativo muy relevante.
16	Palabra clave en la descripción larga	Las imágenes cuentan con una descripción larga que contiene alguna de las palabras clave objetivo.	Imágenes en formato de mapa de bits	Revisión manual del contenido. Puntuaciones 0: Ninguna o solo algunas de las imágenes cuentan con una descripción larga o es una descripción larga que no tiene ninguna relación con el contenido. 1: Las imágenes cuentan con una descripción larga, pero es poco relevante. 2: Las imágenes cuentan con una descripción larga muy relevante.
17	Palabra clave en la etiqueta <figcaption>	Las imágenes cuentan con una etiqueta <figcaption> en la que se incluyen las palabras clave objetivo.	Todas	Revisión manual del código fuente de la página. Puntuaciones 0: Ninguna o solo algunas de las imágenes cuentan con un <i>figcaption</i> . 1: Las imágenes cuentan con un <i>figcaption</i> pero es poco relevante. 2: Las imágenes cuentan con un <i>figcaption</i> muy relevante.
18	Texto contextual	Las visualizaciones cuentan con un texto contextual cercano (inmediatamente antes o después).	Todas	Revisión manual del contenido. Puntuaciones 0: Ninguna o solo algunas de las imágenes cuentan con una descripción larga o es una descripción larga que no tiene ninguna relación con el contenido. 1: Las imágenes cuentan con una descripción larga, pero es poco relevante. 2: Las imágenes cuentan con una descripción larga muy relevante.

ID	Factor	Descripción	Ámbito de aplicación	Método de análisis / Puntuación
19	Metadatos IPTC	Las imágenes cuentan con propiedades IPTC que Google es capaz de extraer.	Imágenes en formato de mapa de bits	Revisión mediante un extractor de metadatos IPTC. Puntuaciones 0: Las imágenes no cuentan con metadatos IPTC. 1: Las imágenes cuentan con algunos de los metadatos IPTC recomendados. 2: Las imágenes cuentan con todos los metadatos IPTC recomendados.
20	Etiqueta max-image-preview:large	La página cuenta con una etiqueta <code>max-image-preview:large</code> para indicar a los buscadores que imagen pueden mostrar como vista previa del contenido.	Imágenes en formato de mapa de bits	Revisión manual del código fuente. Puntuaciones 0: No utiliza la etiqueta <code>max-image-preview:large</code> . 1: Sí que utiliza la etiqueta <code>max-image-preview:large</code> .
21	Palabra clave en el nombre del fichero	La palabra clave objetivo forma parte del nombre del fichero.	Imágenes en formato de mapa de bits	Revisión manual del código fuente. Puntuaciones 0: El nombre del fichero no contiene ninguna de las palabras clave relacionadas con la imagen. 1: El nombre del fichero es parcialmente relevante. 2: El nombre del fichero es totalmente relevante.
22	Palabra clave en los elementos <title>, <desc> o <text> de las imágenes vectoriales	Las imágenes en formato vectorial incluyen una etiqueta <title>, <desc> o <text> con el título, descripciones o contenido de cada objeto gráfico.	Imágenes vectoriales	Revisión manual del código fuente. Puntuaciones 0: Las imágenes vectoriales no cuentan con un <title> o no es relevante 1: El <title> de las imágenes vectoriales es parcialmente relevante 2: El <title> de las imágenes vectoriales es totalmente relevante
23	Calidad de las imágenes	Las imágenes en formato de mapa de bits cuentan con una calidad adecuada para visualizarse correctamente.	Imágenes en formato de mapa de bits	Revisión manual de las imágenes. Puntuaciones 0: La calidad de la imagen es tan pobre que resulta imposible o es muy complicado apreciar su contenido. 1: La imagen se presenta pixelada o borrosa, aunque puede apreciarse el contenido. 2: La calidad es óptima.
24	Largest Contentful Paint (LCP)	El tiempo de carga de la página se encuentra por debajo de 2,5 segundos.	Todas	Revisión mediante Page Speed Insights. Puntuaciones 0: Más de 4 segundos. 1: Entre 2,5 y 4 segundos. 2: Menos de 2,5 segundos.
25	Interaction to Next Paint (INP)	La interacción hasta el siguiente renderizado en la página es inferior a 200 ms.	Todas	Revisión mediante Page Speed Insights. Puntuaciones 0: Más de 500 ms. 1: Entre 200 ms y 500 ms. 2: Menos de 200 ms.
26	Cumulative Layout Shift (CLS)	Los cambios de diseño acumulados en la página son inferiores a 0,1.	Todas	Revisión mediante Page Speed Insights. Puntuaciones 0: Más de 0,25. 1: Entre 0,1 y 0,25. 2: Menos de 0,1.

ID	Factor	Descripción	Ámbito de aplicación	Método de análisis / Puntuación
27	Optimización móvil	La página se encuentra optimizada para dispositivos móviles.	Todas	Revisión manual cualitativa. Puntuaciones 0: La página no se encuentra optimizada para dispositivos móviles. 1: La página se encuentra parcialmente optimizada para dispositivos móviles. 2: La página se encuentra totalmente optimizada para dispositivos móviles.
28	Enlaces entrantes internos	Número de enlaces entrantes internos de la página.	Todas	Revisión con el informe de SEOquake u otras herramientas similares. Puntuaciones 0: Sin enlaces entrantes internos. 1: Hasta 5 enlaces entrantes internos. 2: Más de 5 enlaces entrantes internos.
29	Backlinks	Número de enlaces entrantes externos de la página.	Todas	Revisión de la métrica con Semrush u otras herramientas similares. Puntuaciones 0: Menos de 10000. 1: Entre 10001 y 100000. 2: Más de 100000.
30	Dominios de referencia	Número de dominios de referencia que enlazan a la página.	Todas	Revisión de la métrica con Semrush. Puntuaciones 0: Menos de 50. 1: Entre 51 y 100. 2: Más de 100.
31	Authority score del dominio	Autoridad del dominio.	Todas	Revisión de la métrica con Semrush. Puntuaciones 0: Menos de 50. 1: Entre 51 y 100. 2: Más de 100.
32	Authority score de la página	Autoridad de la página.	Todas	Revisión de la métrica con Semrush. Puntuaciones 0: Menos de 20. 1: Entre 21 y 50. 2: Más de 50.
33	Visitas mensuales del dominio	Número de visitas mensuales recibidas a nivel de dominio.	Todas	Revisión de la métrica con Semrush u otras herramientas similares. Puntuaciones 0: Menos de 1000000. 1: Entre 1000001 y 10000000. 2: Más de 10000000.
34	Tráfico orgánico del dominio	Tráfico mensual estimado a nivel de dominio.	Todas	Revisión de la métrica con Semrush u otras herramientas similares. Puntuaciones Revisión de la métrica con Semrush. 0: Menos de 500000 1: Entre 500001 y 5000000 2: Más de 5000000

3. Conclusiones

El SEO y el periodismo han ido de la mano durante años, asegurando que el contenido de los medios de comunicación alcance la visibilidad necesaria para seguir siendo relevante en la era de Internet. Desde los medios más poderosos y con mayores recursos hasta los más locales

y modestos, han incorporado personal y prácticas dedicadas a la optimización para motores de búsqueda. Las visualizaciones de datos narrativas representan un producto particular que requiere una atención especial, ya que su optimización implica la alineación de diversos factores. Esto es especialmente imperativo teniendo en cuenta la integración de imágenes, gráficos, infografías y otros componentes gráficos, pero también debiendo satisfacer aspectos más amplios que influyen en el ámbito periodístico, como las directrices EEAT. La propuesta SEOvisFrame presenta un conjunto de indicadores para evaluar la efectividad de las prácticas SEO en este contexto, pero también puede servir como una guía de referencia o lista de verificación que utilizar durante la conceptualización y el diseño del producto comunicativo.

4. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

5. Referencias

- Alcaraz Martínez, R. (2022). SEO centrado en el usuario: un cambio de paradigma orientado a satisfacer mejor las necesidades de las personas en el buscador de Google. *Cuadernos de Documentación Multimedia*, 32. <https://doi.org/10.5209/cdmu.81156>
- Alcaraz Martínez, R. (2024). Integración de la optimización de la experiencia de búsqueda (SXO), la usabilidad, la arquitectura de la información y la accesibilidad web: explorando un terreno común para mejorar la visibilidad en buscadores. *Ibersid*. 18(1). <https://doi.org/10.54886/ibersid.v18i1.4966>
- Alexander, S., & Vetere, C. (2011). Telling the data story the right way. *Healthcare Financial Management*, 65(10), 104-110.
- Almukhtar, F., Nawzad, M., & Shahab, K. (2021). Search engine optimization: A review. *Applied Computer Science*, 17(1), 70-80. <https://doi.org/10.35784/acs2021-07>
- Chotikitpat, K., Nilsook, P., & Sodsee, S. (2015). Techniques for improving website rankings with search engine optimization (SEO). *Advanced Science Letters*, 21(10), 3219-3224. <https://doi.org/10.1166/asl.2015.6503>
- Córdoba-Cabús, A. (2020). Estándares de calidad en el periodismo de datos: fuentes, narrativas y visualizaciones en los Data Journalism Awards 2019. *Profesional de la Información*, 29(3), e290328. <https://doi.org/10.3145/epi.2020.may.28>
- Córdoba-Cabús, A., García-Borrego, M., & López-Martín, Á. (2020). El periodismo de datos durante la crisis sanitaria del Covid-19 en la prensa española. *Revista Ibérica de Sistemas e Tecnologías de Informação*, 35, 325-337. <https://hdl.handle.net/10630/20031>
- Dean, B. (2024). Google's 200 ranking factors: the complete list. <https://backlinko.com/google-ranking-factors>
- Dick, M. (2011). Search engine optimisation in UK news production. *Journalism Practice*, 5(4), 462-477. <https://doi.org/10.1080/17512786.2010.551020>

Domínguez, E. (2016). El reto narrativo de integrar datos en los formatos audiovisuales interactivos. In E. Blanco Castilla & M. Quesada (Coords.). *Periodismo de Datos* (pp. 87-105). Cuadernos Artesanos de Comunicación. <http://www.cuadernosartesanos.org/2016/cac112.pdf>

Enge, E.; Spencer, S.; Stricchiola, J. (2023). *The art of SEO: Mastering search engine optimization*. O'Reilly.

Escandell-Poveda, R., Iglesias-García, M., & Papí-Gálvez, N. (2021). Who does SEO in Spain? a cybermetric methodology for the construction of company universes. *Profesional de la Información*, 30(4). <https://doi.org/10.3145/epi.2021.jul.19>

Ferraz, R. (2017). Accessibility and search engine optimization on scalable vector graphics. In *2017 IEEE 4th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, (pp. 94-98). <https://doi.org/10.1109/ISCMI.2017.8279605>

Ferreras Rodríguez, E.M. (2016). El periodismo de datos en España. *Estudios sobre el Mensaje Periodístico*, 22(1) 255-272. https://doi.org/10.5209/rev_ESMP.2016.v22.n1.52594

García-Carretero, L., Codina, L., Díaz-Noci, J., & Iglesias-García, M. (2016). Herramientas e indicadores SEO: características y aplicación para análisis de cibermedios. *Profesional de la Información*, 25(3), 497-504. <https://doi.org/10.3145/epi.2016.may.19>

Giomelakis, D. (2023). Semantic search engine optimization in the news media industry: Challenges and impact on media outlets and journalism practice in Greece. *Social media+ society*, 9(3), 20563051231195545.

Giomelakis, D., & Veglis, A. (2015). Employing search engine optimization techniques in online news articles. *Studies in Media and Communication*, 3(1), 22-33. <https://doi.org/10.11114/smc.v3i1.683>

Google (2014). *HTTPS as ranking signal*. Google search central. <https://developers.google.com/search/blog/2014/08/https-as-ranking-signal>

Google (2024a). *Google search essentials*. <https://developers.google.com/search/docs/essentials>

Google (2024b). *General guidelines*. <https://static.googleusercontent.com/media/guidelines.raterhub.com/es//searchqualityevaluatorguidelines.pdf>

Google (2024c). *Google image SEO best practices*. <https://developers.google.com/search/docs/appearance/google-images>

Google (2025a). *General structured data guidelines*. <https://developers.google.com/search/docs/appearance/structured-data/sd-policies>

Google (2025b). *Image metadata in Google images*. <https://developers.google.com/search/docs/appearance/structured-data/image-license-metadata>

Krstić, N. (2019). The relationship between website management and search engine optimisation. *Megatrend Revija*, 16(2), 155-168. <https://scindeks-clanci.ceon.rs/data/pdf/1820-3159/2019/1820-31591902155K.pdf>

Lewandowski, D. (2023). Search Engine Optimization (SEO). In *Understanding search engines* (pp. 175-190). Springer. https://doi.org/10.1007/978-3-031-22789-9_9

- Lopezosa, C., Codina, L., & Caldera Serrano, J. (2018). SEO semántico: Framework ISS para la optimización de sitios intensivos en contenidos. *Cuadernos de Documentación Multimedia*, 29, 97-123. <https://doi.org/10.5209/CDMU.60607>
- Lopezosa, C., Codina, L., & Gonzalo-Penela, C. (2019). SEO off page y construcción de enlaces: estrategias generales y transmisión de autoridad en cibermedios. *Profesional de la Información*, 28(1). <https://doi.org/10.3145/epi.2019.ene.07>
- Lopezosa, C., Codina, L., Gonzalo Penela, C., & Pont Sorribes, C. (2020). SEO y periodismo: Marco de optimización global como parte del emprendimiento en cibermedios. <http://hdl.handle.net/10230/44923>
- Maciá Domene, F. (2019). *SEO avanzado: casi todo lo que sé sobre posicionamiento web*. Anaya Multimedia.
- Marcos, M. C., Gavin, F., & Arapakis, I. (2015, September). Effect of snippets on user experience in web search. In *Proceedings of the XVI international conference on human computer interaction* (pp. 1-8). <https://doi.org/10.1145/2829875.2829916>
- Mueller, J. (2015). *English Google webmaster central office-hours hangout*. https://youtu.be/i_xnKznRNCc?t=24m5s
- Peiró, K., & Guallar, J. (2013). Introducció al periodisme de dades. Característiques, panoràmica i exemples, *Ítem: Revista de Biblioteconomia i Documentació*, 57, 23-37. <https://raco.cat/index.php/Item/article/view/269701>
- Radcliffe, D., & Lewis, S. C. (2019). The datafication of journalism: strategies for datadriven storytelling and industry-academy collaboration. In *The data journalism handbook 2*. <https://datajournalism.com/read/handbook/two/training-datajournalists/the-datafication-of-journalism>
- Rodas, C. M., Vidotti, S. A. B. G., & Monteiro, S. D. (2016). Interface de busca do Google e Yahoo: a experiência do usuário sob o olhar do eye tracking. *Informação & Sociedade*, 26(2). <https://doi.org/10.22478/ufpb.1981-0695.2017v12n1.34043>
- Salem, H., Salloum, H., Orabi, O., Sabbagh, K., & Mazzara, M. (2025). Enhancing News Articles: Automatic SEO Linked Data Injection for Semantic Web Integration. *Applied Sciences*, 15(3), 1262. <https://doi.org/10.3390/app15031262>
- Seyser, D., & Zeiller, M. (2018). Scrollytelling: An analysis of visual storytelling in online journalism. In *22nd International Conference Information Visualisation* (pp. 401-406). <https://doi.org/10.1109/IV.2018.00075>
- Semrush (2024). *Google ranking factors*. <https://go.semrush.com/Ranking-Factors.html>
- Vicente, J.E., Saquete, R., Rodríguez, R., & Maciá Domene, F. (2024). *SEO técnico: aprende a optimizar tu web como un profesional*. Anaya Multimedia.
- Ziakis, C., Vlachopoulou, M., Kyrkoudis, T., Karagkiozidou, M. (2019). Important factors for improving Google search rank. *Future Internet*, 11(32). <https://doi.org/10.3390/fi11020032>

Contenido semántico y visualización de información: una propuesta de narrativa comunicacional basada en datos

Mario Pérez-Montoro

Universitat de Barcelona, España

<https://orcid.org/0000-0003-2426-8119>

Pérez-Montoro, M. (2025). Contenido semántico y visualización de información: una propuesta de narrativa comunicacional basada en datos. En J. Guallar, M. Váñez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 89-109). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvi-com.07.esp>

Resumen

Las gráficas son uno de los principales productos visuales que se utilizan para representar y transmitir la información. Una parte importante de nuestro trabajo diario consiste en la búsqueda, obtención y refinado de un conjunto de datos que nos permita narrar una historia. Pero los datos no hablan por sí solos. Necesitan ser visualizados para mostrar un análisis derivado y permitir que afloren los patrones que encierran. En este trabajo vamos a intentar caracterizar los diferentes y principales contextos y contenidos comunicacionales que podemos transmitir utilizando una gráfica y veremos cuál es la propuesta visual más adecuada para cubrir con eficiencia y eficacia esas intenciones comunicativas. O, dicho en otros términos, en este capítulo queremos desplegar de forma ordenada la metodología habitual que se utiliza para implementar de forma adecuada una correcta visualización de información: en primer lugar, se identifica el tipo de información a codificar, en segundo lugar, se elige el tipo de gráfica que mejor representa ese tipo de información y, por último, se diseñan los elementos visuales que conforman esa gráfica.

Palabras clave

Visualización de información; Visualización de datos; Comunicación visual de la información; Diseño de la información; Gráficas; Representación visual de la información.

1. Introducción

Las gráficas son uno de los principales productos visuales que se utilizan para representar y transmitir la información.

En este capítulo vamos a intentar caracterizar los diferentes y principales contextos y contenidos comunicacionales que podemos transmitir utilizando una gráfica y veremos cuál es la propuesta visual más adecuada para cubrir con eficiencia y eficacia esas intenciones comunicativas. O, dicho en otros términos, en este capítulo queremos desplegar de forma ordenada la metodología habitual que se utiliza para implementar de forma adecuada una correcta visualización de información: en primer lugar, se identifica el tipo de información a codificar, en segundo lugar, se elige el tipo de gráfica que mejor representa ese tipo de información y, por último, se diseñan los elementos visuales que conforman esa gráfica.

Para cubrir ese objetivo, y siguiendo ese esquema, vamos a tratar de desarrollar los siguientes puntos. En el primero de los apartados (apartado 2), abordaremos los principales tipos básicos de información o contenidos semánticos que podemos representar en una gráfica, ilustrándolos a partir de ejemplos que codificaremos en forma de tabla para que puedan ser entendidos de una forma más intuitiva. Y, en el apartado 3, examinaremos cuál es tipo de visualización que se adecua de forma más conveniente para comunicar cada uno de esos tipos de contenidos informativos.

2. Contenidos semánticos

De forma intuitiva podemos afirmar que cualquier tipo de estrategia de comunicación se fundamenta sobre tres elementos básicos: el mensaje o contenido semántico que queremos transmitir, el receptor al cuál va dirigido y el contexto que rodea todo el acto comunicativo.

El caso de la comunicación a través del uso de gráficas no es diferente. Esa comunicación depende del mensaje o contenido informativo que pretendemos hacer llegar a nuestra audiencia, el público al cuál queremos que llegue ese mensaje y el contexto (la unidad documental en la que se inserta la gráfica, el texto que acompaña a la gráfica, el canal de transmisión, nuestra intención comunicativa o el momento de su consumo, entre otros muchos factores) que arroja todo ese intento de comunicar con graficas. Centrémonos ahora exclusivamente en esos contenidos que queremos transmitir.

Normalmente, el contenido semántico que intentamos transmitir con una gráfica es una proposición (una idea, en un sentido no técnico) en la que se afirma la existencia de una relación entre dos o más tipos de información. Concretamente, se afirma la relación entre dos (o más) ítems de información cuantitativa y/o cualitativa. O, dicho de otra manera, en esos contenidos se expresa algún tipo de relación entre valores que pertenecen a dos o más variables cuantitativas, entre las diferentes categorías o atributos que conforman dos o más variables cualitativas, o entre categorías o atributos y valores cuantitativos.

En este sentido, es posible establecer una clasificación de las principales relaciones expresadas en los contenidos semánticos representados por una gráfica (Zelazny, 2001; Few, 2012; Shneiderman, 1996; Pérez-Montoro, 2022). Entre estas, podríamos destacar, por ejemplo, la relación de comparación nominal, la de series temporales, la de ranking, la de parte-todo, la de desviación, la de distribución y la relación de correlación.

2.1. Relación de comparación nominal

La relación de comparación nominal podría ser considerada como el contenido semántico más básico y habitual que solemos representar utilizando una gráfica.

En términos técnicos este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la comparación entre esos valores numéricos relacionados con esos ítems.

Figura 1

Cargo y sueldo de los empleados de la empresa ACME.

Cargo	Sueldo
Gerente	150.000
Analista	40.000
Documentalista	40.000
Administrativo	32.000
Auxiliar	25.000
Becario	8.000

(Elaboración propia, datos simulados)

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 1. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (gerente, analista, documentalista, administrativo, auxiliar y becario) de una variable cualitativa (los diferentes empleados (cargos o perfiles laborales) que podemos encontrar en la empresa ACME) con un valor numérico o cuantitativo (el sueldo bruto que reciben cada uno de esos empleados según el cargo que desempeñan en la empresa).

2.2. Relación de series temporales

La relación de series temporales es otro de los contenidos semánticos habituales que acostumbramos a comunicar mediante el uso de gráficas.

De nuevo, desde una perspectiva conceptual, este contenido semántico se identifica como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable temporal para facilitar la comparación entre esos valores numéricos asociados a esas unidades temporales.

Figura 2

Ventas por trimestre en el ejercicio anterior.

Trimestre	Ventas
Primero	456.876
Segundo	189.387
Tercero	63.829
Cuarto	122.345

(Elaboración propia, datos simulados)

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 2. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (primero, segundo, tercero y cuarto) de una variable temporal (los diferentes trimestres del año pasado) con un valor numérico o cuantitativo (las ventas que ha generado la empresa en esas unidades temporales).

2.3. Relación de ranking

Este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se señala que esta asignación de valores numéricos establece una ordenación (de mayor a menor o de menor a mayor) entre los ítems de la variable cualitativa.

Figura 3

Ventas por provincias en Cataluña.

Ranking	Provincia	Ventas
1	Barcelona	456.876
2	Tarragona	189.387
3	Gerona	122.345
4	Lérida	63.829

(Elaboración propia, datos simulados)

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 3. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (Barcelona, Tarragona, Gerona y Lérida) de una variable cualitativa (las provincias que conforman la región española de Cataluña) con un valor numérico o cuantitativo (las ventas que una empresa ha conseguido en esa provincia). Esta

asignación de valores numéricos establece una ordenación (de mayor a menor o de menor a mayor) entre esas provincias catalanas.

2.4. Relación de parte-todo

La relación de parte-todo podría ser considerada como otro de los contenidos más habituales que solemos representar utilizando una gráfica.

Desde una perspectiva conceptual, este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se señala que la agregación de la totalidad de esos ítems cualitativos conforma una entidad o un todo y, por tanto, cada una de esas categorías se identifica como una de las partes de ese todo. Es importante destacar también que la unidad común de medida más utilizada en este tipo de relaciones que expresan la contribución de las partes a un todo es el porcentaje. En ese contexto, el todo es el 100% y cada una de sus partes es una porción de ese 100%.

Figura 4

Actividades desarrolladas a lo largo de un día.

Actividad	Porcentaje
Comer	7,40%
TV/Internet	8,20%
Socializar	9,40%
Dormir	29,20%
Trabajar	45,80%
Total	100,00%

(Elaboración propia, datos simulados)

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 4. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (trabajar, comer, socializar, ver TV/internet y dormir) de una variable cualitativa (las diferentes actividades realizadas a lo largo del día) con un valor numérico o cuantitativo (el porcentaje de tiempo del día que le dedicamos a esa actividad). Esta asignación de valores numéricos expresa la relación cuantitativa que mantienen cada una de las partes (su peso, por así decirlo) respecto al todo y facilita también la comparación entre esos valores numéricos relacionados con esos ítems.

2.5. Relación de desviación

De forma intuitiva, la relación de desviación expresa como varía un conjunto de datos cuantitativos respecto a un valor numérico concreto.

En un sentido más técnico, es un contenido semántico que se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se comparan esos valores cuantitativos asociados respecto a un valor numérico concreto y de referencia. Normalmente, el valor numérico de referencia se encuentra también

asignado a un ítem cualitativo (de esa misma u otra variable cualitativa). En algunas circunstancias pueden existir dos o más valores de referencia respecto a los cuales se quiere realizar la comparación.

Figura 5

Comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa.

Empresa	Ventas
Competidor A	280%
Competidor B	220%
Competidor C	115%
Competidor D	60%
Competidor E	20%

(Elaboración propia, datos simulados)

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 5. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (competidor A, competidor B, competidor C, competidor D y competidor E) de una variable cualitativa (las diferentes empresas que compiten sectorialmente con la nuestra) con un valor numérico o cuantitativo (el porcentaje de ventas de esa firma respecto a las ventas de nuestra empresa). Esta asignación de valores numéricos expresa la relación cuantitativa que mantienen las ventas de cada empresa respecto a las de la nuestra y facilita también la comparación entre esos valores numéricos relacionados con esos ítems. En este caso, el valor de referencia (las ventas de nuestra empresa) se identifica con el 100% y el resto de los valores (las ventas de la competencia) son expresados como una proporción o porcentaje respecto a ese valor primario o de referencia.

2.6. Relación de distribución

La relación de distribución podría ser considerada como un contenido semántico que se corresponde con una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable de intervalo para facilitar la comparación entre esos valores numéricos asociados a esas unidades de intervalo. Ese contenido expresa cómo los valores cuantitativos de un conjunto se encuentran repartidos o diseminados a lo largo de un rango (intervalo entre el valor máximo y el valor mínimo) de ese conjunto, desde el valor más bajo hasta el más alto, a través de las categorías que conforman esa variable de intervalo.

Figura 6

Pacientes por franja de edad en una población de enfermos de epilepsia.

Franja edad	Pacientes
0-15	15
16-30	30
31-45	55
46-60	75
61-75	25

(Elaboración propia, datos simulados)

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 6. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (0 a 15, 16 a 30, 31 a 45, 46 a 60 y 61 a 70) de una variable de intervalo (las franjas de edad de una población de pacientes que sufren episodios de epilepsia) con un valor numérico o cuantitativo (el número de pacientes). Esta asignación de valores numéricos expresa la distribución del número de pacientes (cómo están repartidos o diseminados) por franja de edad de una población de 200 enfermos de epilepsia.

2.7. Relación de correlación

En un sentido intuitivo, la relación de correlación intenta expresar si dos conjuntos de valores cuantitativos se encuentran relacionados entre sí y, si es el caso, cómo varía uno respecto al otro.

Desde una perspectiva más conceptual y en su formulación más básica, este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asignan o conectan dos valores cuantitativos (cada uno perteneciente a una variable cuantitativa distinta) a cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la identificación de algún patrón en la variación entre esos dos tipos de valores numéricos relacionados con esos ítems. Intenta expresar si existe o no una situación en la que los valores de una de las dos variables cuantitativas se modifican o cambian de manera sistemática con respecto a los valores de la otra.

Figura 7

Inversión y beneficios por producto fabricado.

Producto	Inversión	Beneficios
Producto 01	22.864	678.543
Producto 02	45.789	778.765
Producto 03	50.678	783.213
Total	119.331	2.240.521

(Elaboración propia, datos simulados)

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 7. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (producto 1, producto 2 y producto 3) de una variable cualitativa (los diferentes productos que fabrica la empresa ACME) con dos valores numéricos o cuantitativos que pertenecen respectivamente a dos variables cuantitativas (la inversión realizada en la fabricación y distribución de un producto, por un lado; y los beneficios totales obtenidos en la vida comercial de cada producto, por otro).

3. Visualización de contenidos

Una vez revisados los principales contenidos semánticos que queremos comunicar utilizando gráficas, nos queda ahora abordar el proceso de visualización de esos contenidos o unidades proposicionales a través de la representación gráfica. O dicho en otros términos, queremos analizar cuáles son las mejores soluciones gráficas para codificar adecuadamente cada uno de esos tipos de contenidos semánticos.

En un sentido práctico, una de las acciones críticas a la hora de generar una buena visualización es, aunque no sea la única, la elección correcta de la gráfica que mejor comunica (de forma eficaz y eficiente) la información que queremos representar. No se trata de una tarea sencilla. Depende de muchos factores, pero especialmente del tipo de información a representar y de nuestra intencionalidad comunicativa. Y es que cada tipo de contenido semántico reclama una representación gráfica determinada dependiendo también de nuestros objetivos comunicativos.

La mayoría de los programas de gestión de datos no nos asisten ni ayudan en la elección de la mejor gráfica para realizar esa visualización (Hugues y van Dam, 2013; Hearn, 2011; Shirley, 2009). Sólo algunas herramientas informáticas nos ofrecen algunas pequeñas orientaciones parciales (incompletas, en muchos casos) para hacerlo en función de la estructura de los datos a representar (Cherven, 2015; Jones, 2014; Khan, 2016).

En este apartado vamos a intentar afrontar ese problema. Vamos a intentar ofrecer una serie de recursos y consejos para poder elegir la gráfica comunicativamente más adecuada para representar cada uno de los contenidos semánticos que solemos codificar. Concretamente, introduciremos los mejores recursos para visualizar de forma adecuada la relación de comparación nominal, la de series temporales, la de ranking, la de parte-todo, la de desviación, la de distribución y la relación de correlación.

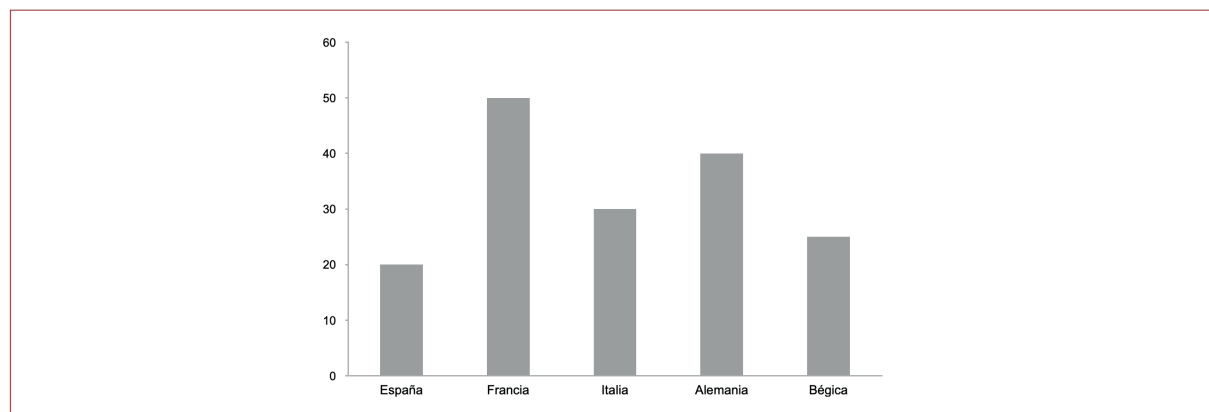
3.1. Visualización de la relación de comparación nominal

Una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifique simultáneamente el valor numérico y el ítem cualitativo con el que se encuentra asociado. Y deben codificar esta relación mostrando que cada una de las duplas resultante formadas por un valor cuantitativo y una categoría cualitativa son independientes entre sí, que no existe una conexión entre estas duplas.

Sin perder de vista esta restricción semántica, los objetos visuales que mejor codifican esta relación son las barras (o columnas) (figura 8) y los puntos. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo) mostrando que visualmente son independientes entre ellas y que, por tanto, no existe una conexión intrínseca entre las mismas. Al hacerlo de ese modo, facilitan así la comparación visual entre esos valores numéricos relacionados con esos ítems.

Figura 8

Goles en contra recibidos por las cinco selecciones nacionales de fútbol mejor clasificadas en la última Eurocopa de fútbol.

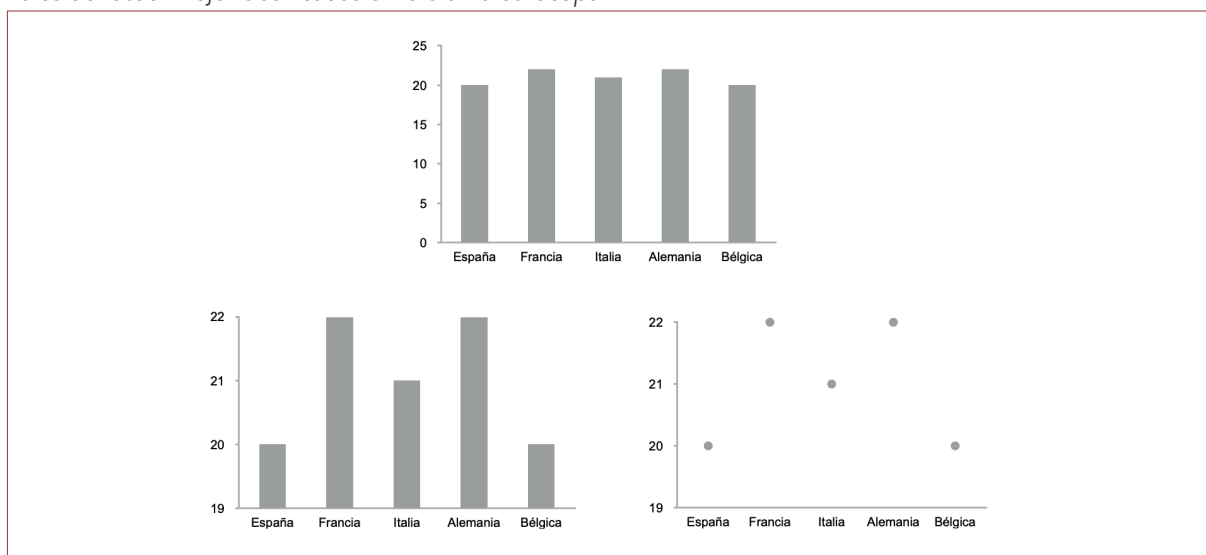


(Elaboración propia, datos simulados)

En ciertos contextos, la gráfica de puntos puede considerarse como una buena alternativa a la gráfica de columnas (o barras) en la representación de este tipo de contenido semántico. Concretamente, cuando la diferencia entre los valores cuantitativos que se representan es muy escasa. O, dicho en otros términos, cuando el rango en el que fluctúan esos valores es pequeño, cuando existe un grado de dispersión bajo entre ese conjunto de valores.

Figura 9

Gráficas de columnas y de puntos para representar los goles en contra recibidos por las cinco selecciones nacionales de fútbol mejor clasificadas en la última eurocopa.



(Elaboración propia, datos simulados)

En esos casos, si representamos esos datos mediante una gráfica de columnas (o barras), puede costar apreciar la diferencia entre la longitud de esas figuras geométricas y, dificultar, de esa manera, la comparación visual entre las mismas y los valores numéricos que codifican. Para representarlos de una forma alternativa con columnas (o barras) y hacer evidente las diferencias entre los valores, podemos tener la tentación de comenzar el eje vertical Y (o X, en el caso de la gráfica de barras) en un valor superior a 0, en un valor poco menor que el valor numérico más pequeño representado y hacer visualmente más grandes la diferencia entre la longitud de esas figuras geométricas y de los valores que codifican. Pero esa práctica es incorrecta. Las columnas (o barras), al codificar el valor cuantitativo mediante su longitud, necesitan comenzar de 0 en el eje vertical (u horizontal, en el caso de la gráfica de barras). Si no lo hacen, no codifican correctamente el valor numérico.

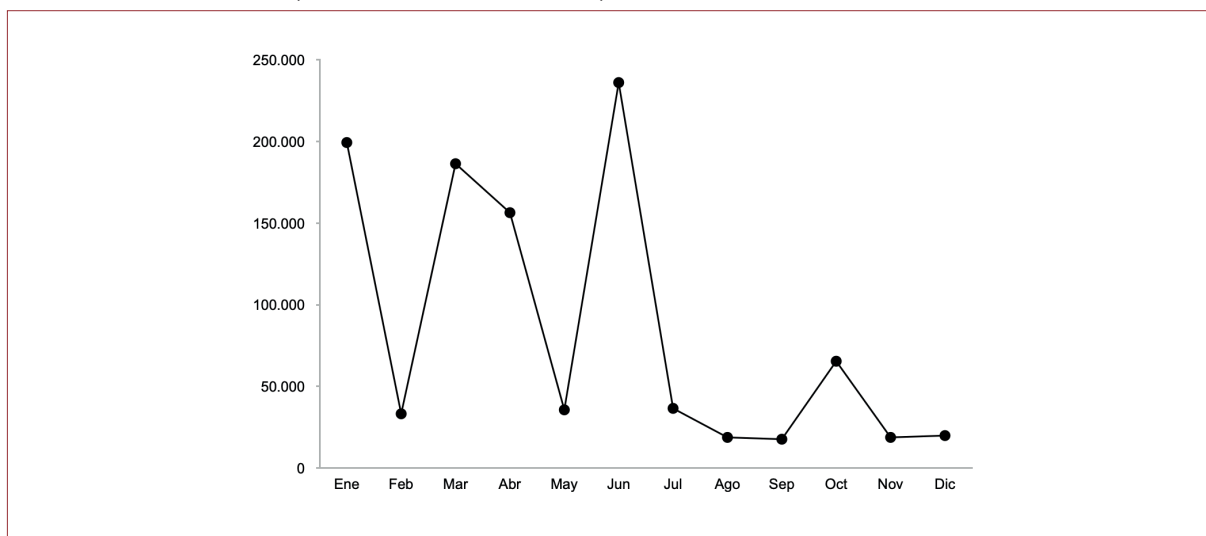
En cambio, si utilizamos una gráfica de puntos como alternativa, esa comprensión errónea no debería producirse ya que los puntos reciben su valor cuantitativo y cualitativo asociado a través de la posición que ocupan en el plano generado por los ejes X e Y (figura 9).

3.2. visualización de la relación de series temporales

Como ya destacamos, la relación de series temporales nominal asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable temporal para facilitarnos la comparación entre esos valores numéricos asociados a esas unidades temporales.

Figura 10

Ventas mensuales de la empresa ACME durante el año pasado.

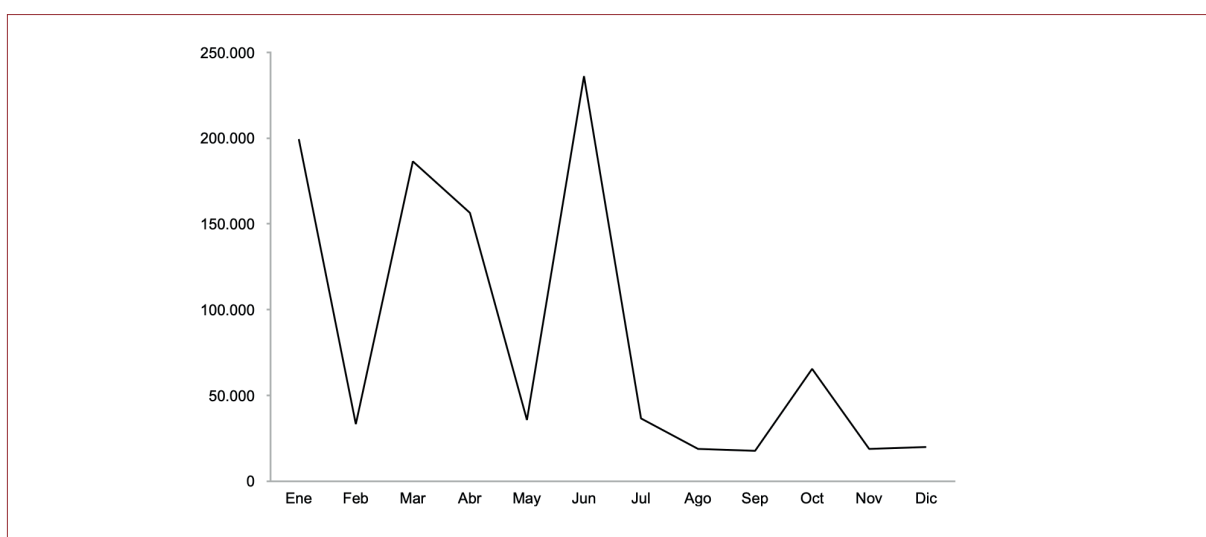


(Elaboración propia, datos simulados)

A la luz de esa descripción, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente el valor numérico y la categoría temporal con la que se encuentra asociado. Para que estos objetos puedan codificar estas duplas semánticas de forma adecuada, en uno de los ejes debe ubicarse esa variable temporal con cada una de las etiquetas que representan un momento de tiempo (año, mes, etc.). Y, como los valores temporales poseen un orden natural, ese orden debe ser también respetado en la visualización. Por convención, en nuestra cultura occidental este tipo de orden se suele visualizar horizontalmente, de izquierda a derecha, a través del eje X.

Figura 11

Ventas mensuales de la empresa ACME durante el año pasado.



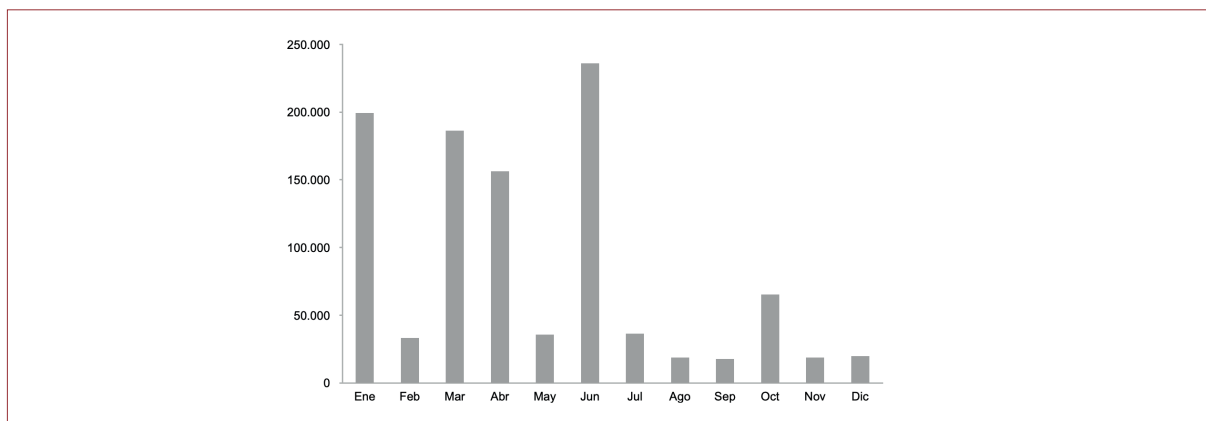
(Elaboración propia, datos simulados)

Teniendo en cuenta todo esto, los objetos visuales que mejor codifican esta relación son la combinación de puntos y líneas (figura 10), las líneas utilizadas de forma no combinada (figura 11) y las columnas. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo)

mostrando que visualmente están conectadas con la siguiente, que son el eslabón o estadio dentro de una cadena temporal. Al hacerlo de ese modo, representan esa continuidad temporal entre esas duplas.

Figura 12

Ventas mensuales de la empresa ACME durante el año pasado.



(Elaboración propia, datos simulados)

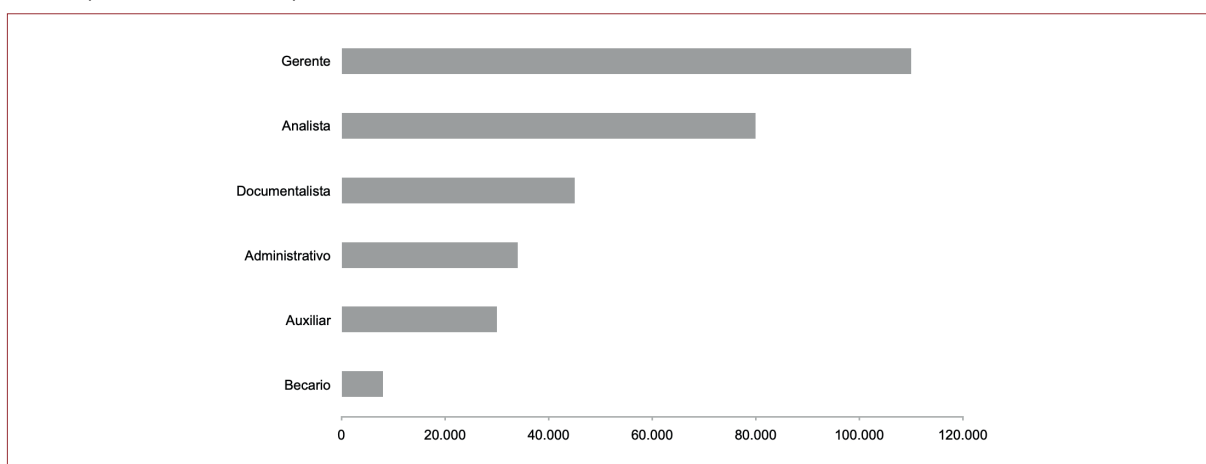
La gráfica de columnas, en cambio, la utilizaremos cuando queramos representar una relación de series temporales, pero queramos poner nuestro foco comunicativo en la transmisión de los valores cuantitativos individuales asociados a los momentos de tiempo de una variable temporal (cuando busquemos facilitar la comparación entre esos valores) y no en la evolución global de la serie (figura 12).

3.3. Visualización de la relación de *ranking*

Anteriormente pudimos ver que la relación de *ranking* asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se señala que esta asignación de valores numéricos establece una ordenación (de mayor a menor o de menor a mayor) entre los ítems de esa variable cualitativa.

Figura 13

Sueldo por cargo de la empresa ACME.

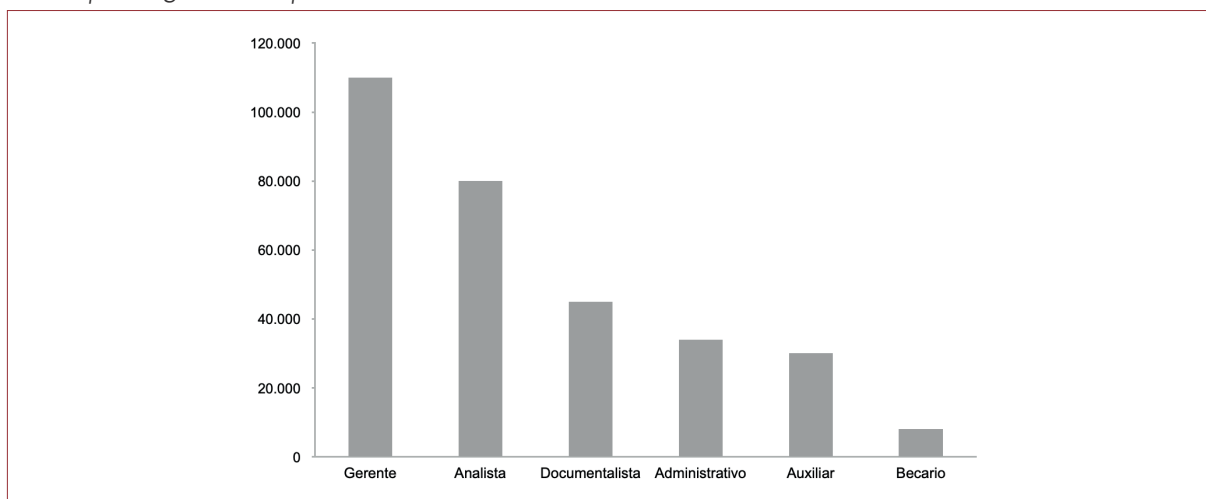


(Elaboración propia, datos simulados)

Teniendo en cuenta esto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifique simultáneamente el valor numérico y la categoría con la que se encuentra asociado y que muestren ese orden intrínseco entre las categorías. Para que estos objetos puedan codificar estas duplas semánticas de forma adecuada, en uno de los ejes debe ubicarse esa variable cualitativa respetando su ordenación interna y, en el otro, la escala cuantitativa.

Figura 14

Sueldo por cargo en la empresa ACME.

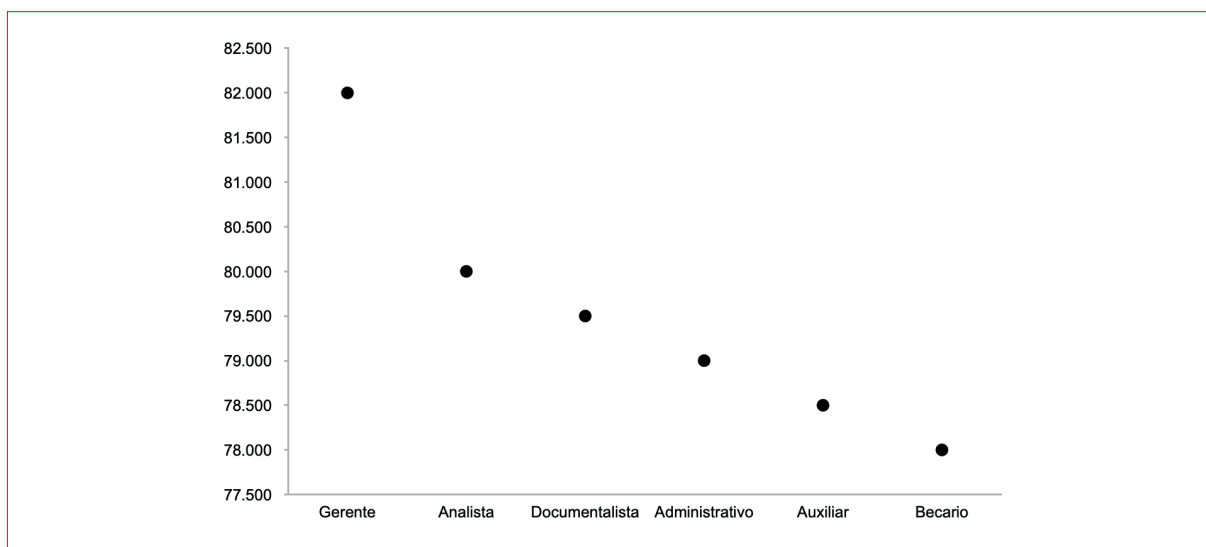


(Elaboración propia, datos simulados)

A la luz de estas condiciones, los objetos visuales que mejor codifican esta relación son las barras (figura 13), las columnas (figura 14) y los puntos. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo) mostrando visualmente el orden interno establecido entre los ítems de la variable cualitativa.

Figura 15

Sueldo por cargo en la empresa ACME.



(Elaboración propia, datos simulados)

Frente a la de barras o a la de columnas, la gráfica de puntos la utilizaremos en aquellos casos en los que la diferencia entre los valores cuantitativos que se representan sea muy escasa y queramos visualizar de forma cómoda la relación de ranking (figura 15).

3.4. Visualización de la relación de parte-todo

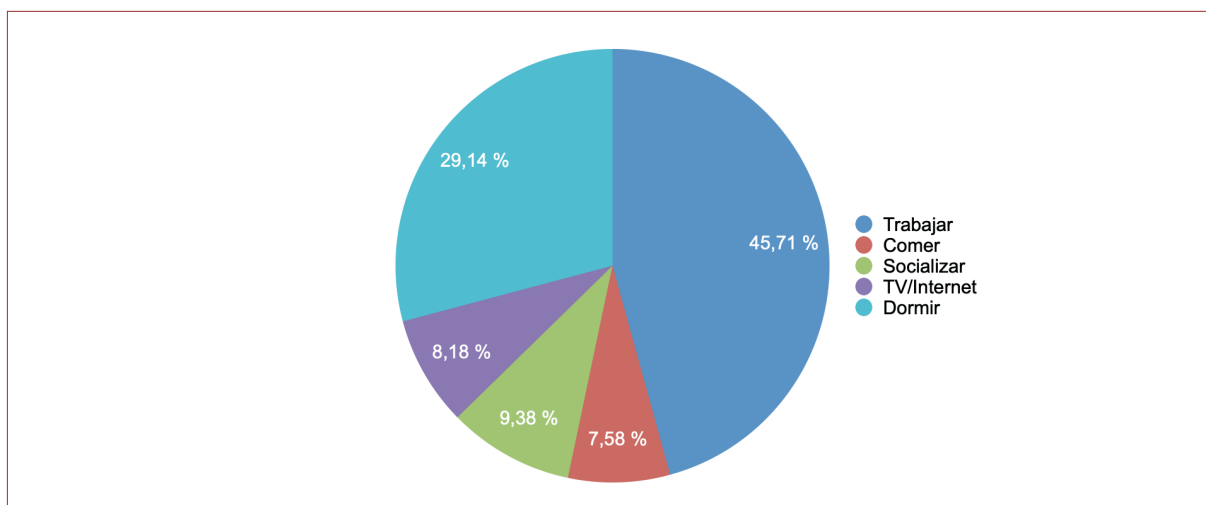
Como ya señalamos, la relación de parte-todo asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa que, a su vez, conforman una entidad o un todo.

Teniendo en cuenta esto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente el valor numérico y la categoría con la que se encuentra asociado y que expresen también esa relación cuantitativa que mantienen cada una de las partes (su peso, por así decirlo) respecto al todo y poder facilitar así la comparación entre esos valores numéricos relacionados con esos ítems.

En muchos contextos comunicacionales, las propuestas visuales que habitualmente se utilizan para codificar esta relación son las gráficas circulares y las gráficas de barras (o columnas) apiladas.

Figura 16

Actividades desarrolladas a lo largo de un día.

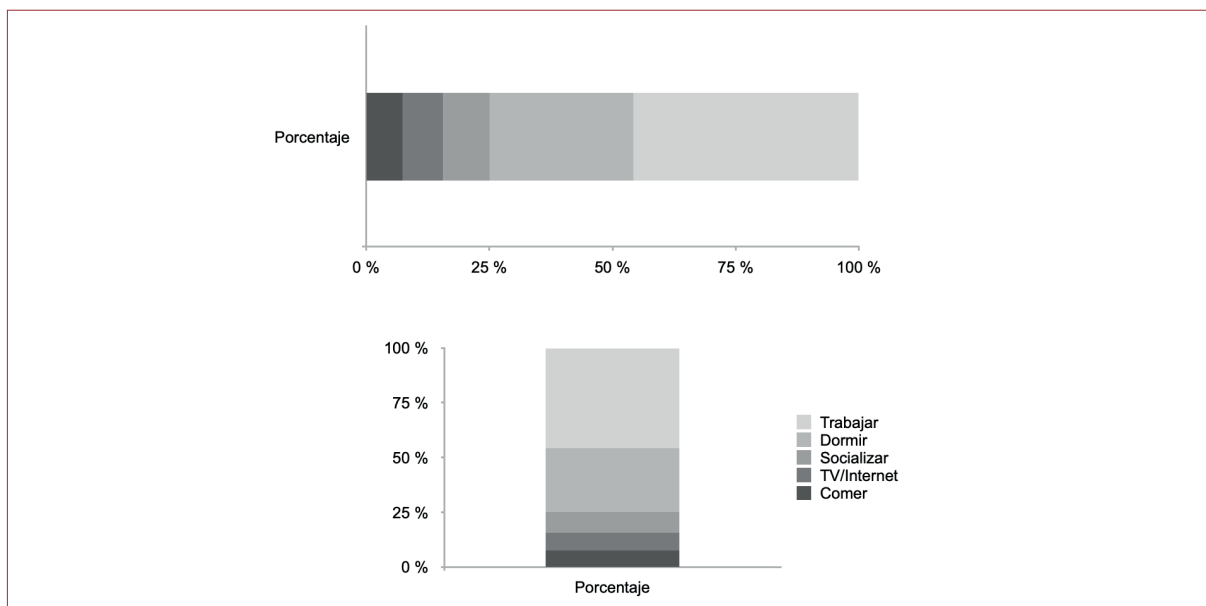


(Elaboración propia, datos simulados)

En una gráfica circular, cada sector nos codifica un valor cuantitativo (en función del ángulo de los radios que los limitan) asociado a un ítem cualitativo (determinado por el color que le asigna la leyenda) y nos muestra las relaciones cuantitativas que mantienen cada una de esas partes o sectores entre sí y con el todo (figura 16).

Figura 17

Actividades desarrolladas a lo largo de un día.

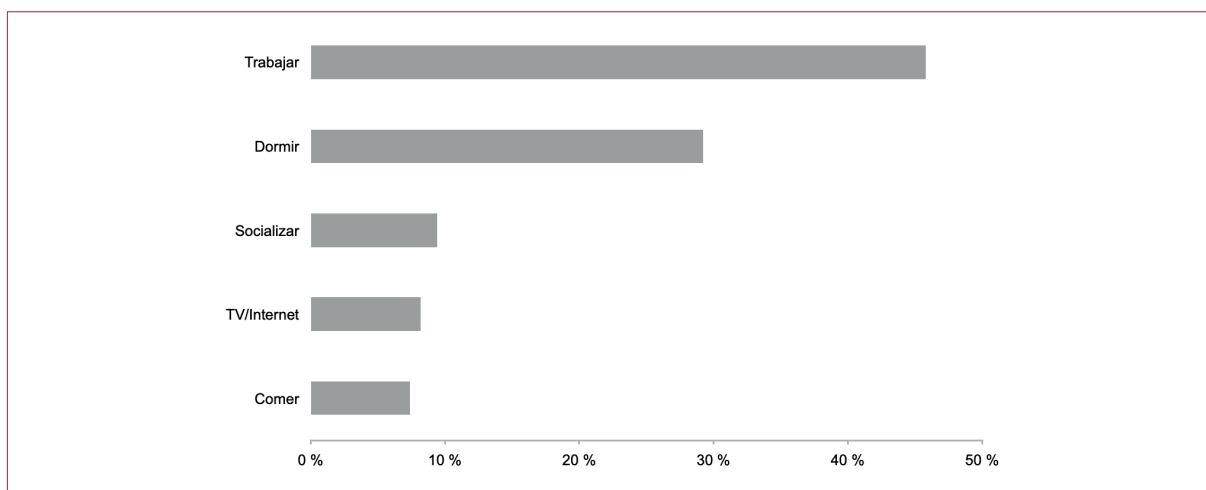


(Elaboración propia, datos simulados)

La otra propuesta habitual que se utiliza para representar la relación de parte-todo es la gráfica de barras (o columnas) apiladas. Estas gráficas representan mediante una barra o columna (una forma rectangular) un valor (un todo) que se obtiene a partir de la agregación de la representación, en forma de cuadrilátero, de los valores de sus partes. Igual que en el caso de la gráfica circular, nos muestra las relaciones cuantitativas que mantienen cada una de esas partes entre sí y con el todo (figura 17).

Figura 18

Actividades desarrolladas a lo largo de un día .

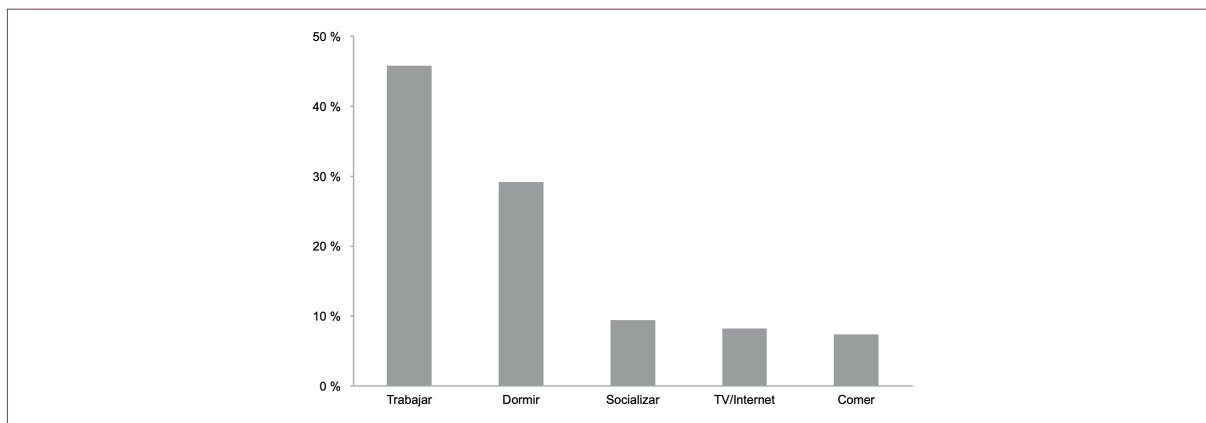


(Elaboración propia, datos simulados)

Frente a estos dos tipos de gráfica que normalmente se utilizan para representar la relación de parte-todo, los objetos visuales que mejor codifican esta relación son las barras (figura 18) y las columnas (figura 19) no apiladas.

Figura 19

Actividades desarrolladas a lo largo de un día.



(Elaboración propia, datos simulados)

Es importante señalar que, frente a las barras y columnas, los elementos visuales que no utilizaremos para representar la relación de parte-todo son las líneas y los puntos (o la combinación de líneas y puntos).

3.5. Visualización de la relación de desviación

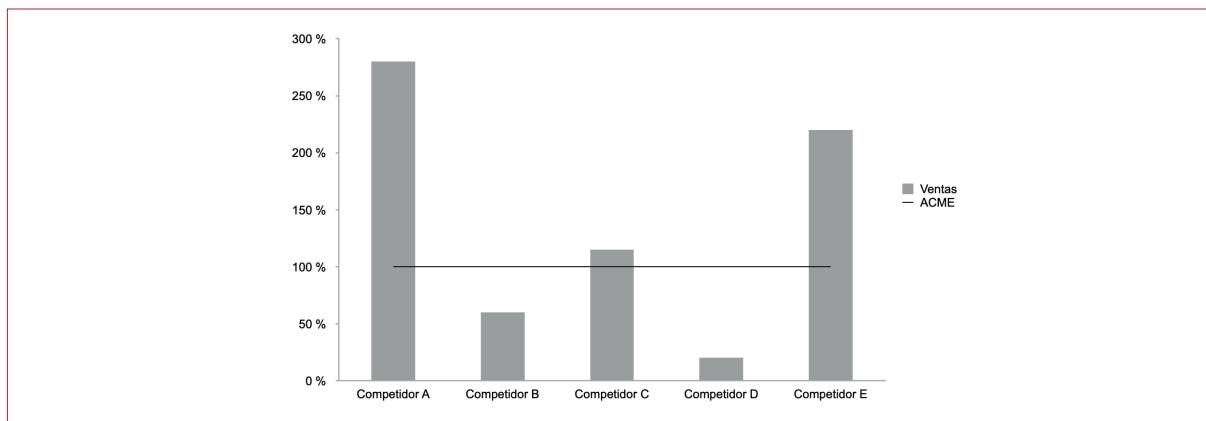
Como ya planteamos, la relación de desviación asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se comparan esos valores cuantitativos asociados respecto a un valor numérico concreto y de referencia.

Teniendo en cuenta esto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente el valor numérico y la categoría con la que se encuentra asociado y que permitan la comparación cuantitativa de esos datos numéricos respecto al valor de referencia.

Contemplando esa restricción, los objetos visuales que mejor codifican esta relación son las barras y las columnas. Este tipo de elementos representan adecuadamente esas duplas (valor cuantitativo e ítem cualitativo) mostrando visualmente qué relación cuantitativa mantienen con el valor de referencia.

Figura 20

Comparación de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa.

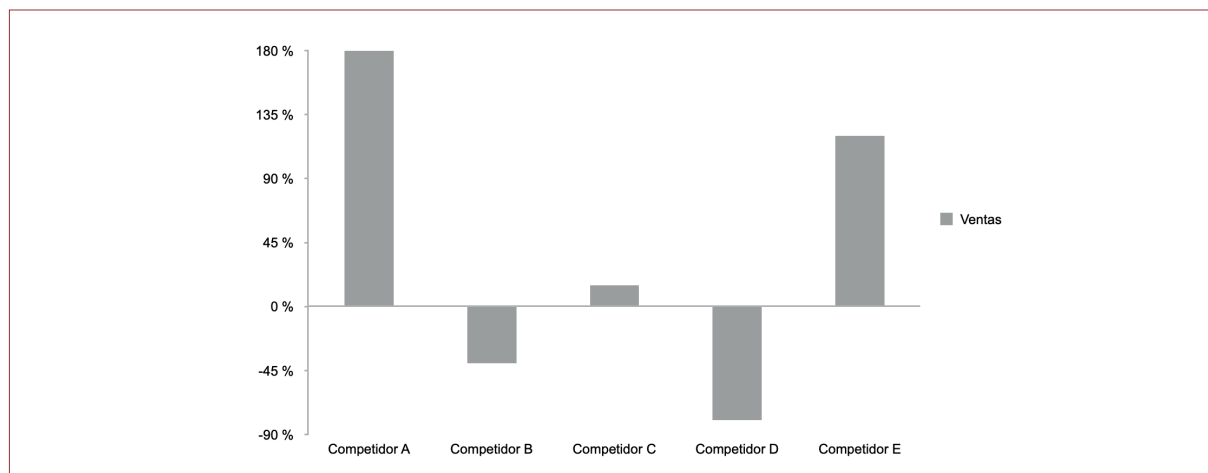


(Elaboración propia, datos simulados)

En una gráfica de barras (o columnas), cada figura rectangular nos codifica un valor cuantitativo (respecto al eje X en el caso de las barras, y respecto al eje Y en el caso de la columnas) asociado a ítem cualitativo (respecto al eje Y en el caso de las barras y respecto al eje X en el caso de la columnas), y, si representamos el valor de referencia mediante una línea o haciéndolo coincidir con uno de los ejes (con el eje Y en el caso de las barras y con el eje X en el caso de la columnas), nos expresa también la relación cuantitativa que mantienen esos valores numéricos respecto al de referencia (figura 20).

Figura 21

Comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa.



(Own elaboration, simulated data)

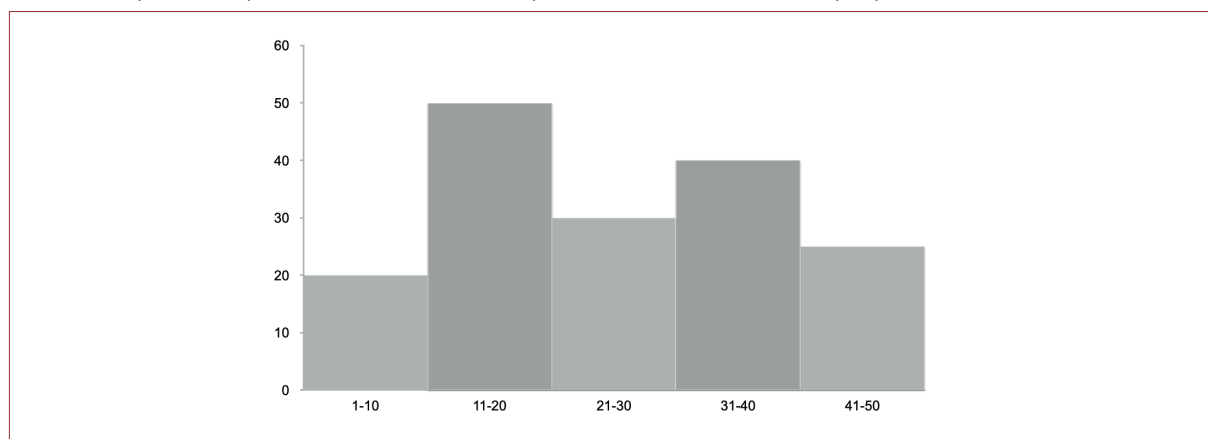
Frente a esta gráfica mixta, podemos ofrecer una versión alternativa sin línea, que incluya las columnas como únicos objetos visuales. En este caso, para no tener que utilizar la línea, haremos coincidir el valor de referencia (las ventas de ACME) con el eje X (figura 21).

3.6. Visualización de la relación de distribución

Anteriormente definimos que la relación de distribución asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable de intervalo para facilitar la comparación entre esos valores numéricos asociados a esas unidades de intervalo.

Figura 22

Número de pacientes por franja de edad en una población de enfermos de epilepsia .

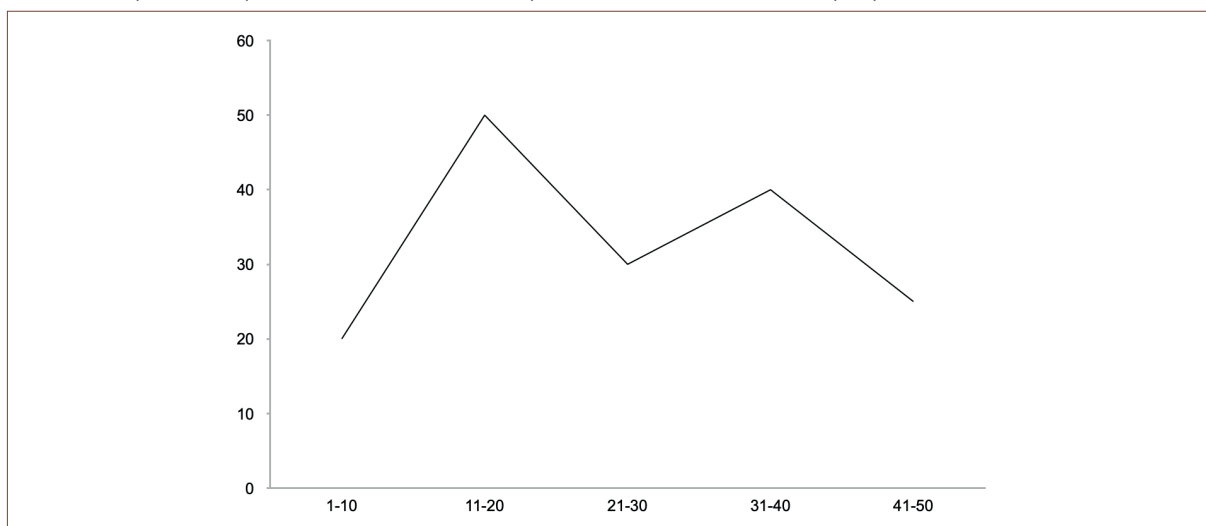


(Elaboración propia, datos simulados)

Por tanto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente el valor numérico y la categoría con la que se encuentra asociado y que permitan expresar cómo esos valores cuantitativos se encuentran repartidos o diseminados a lo largo del rango del conjunto al cual pertenecen, a través de las categorías que conforman esa variable de intervalo.

Figura 23

Número de pacientes por franja de edad en una población de enfermos de epilepsia.

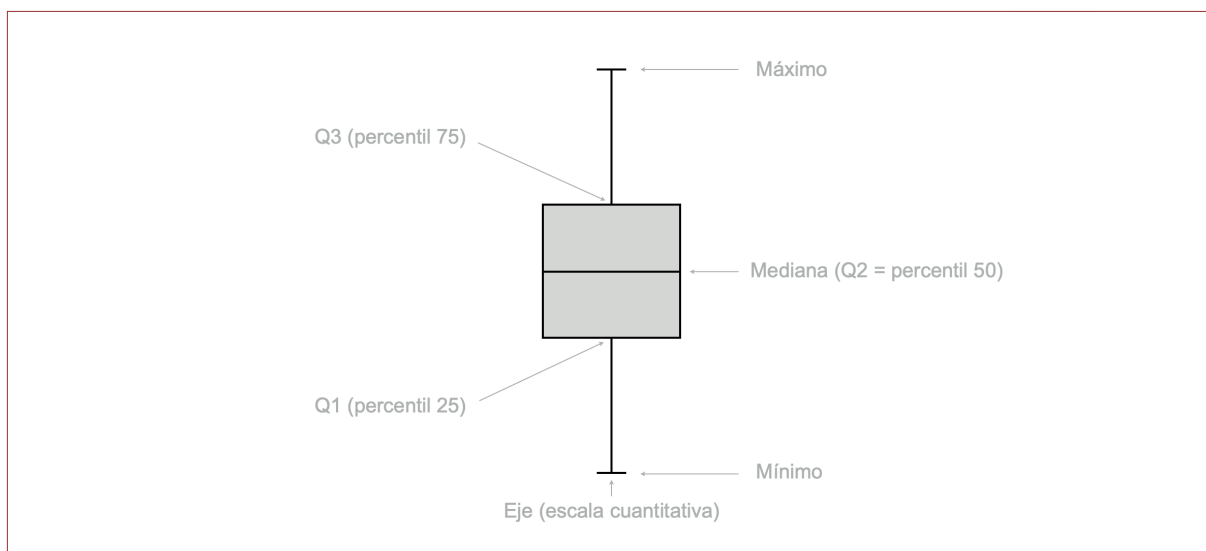


(Elaboración propia, datos simulados)

Otra de las restricciones que debemos tener en cuenta es si queremos representar la distribución de un único conjunto de datos o, por el contrario, queremos visualizar simultáneamente, en una misma propuesta, la distribución de dos o más conjuntos de datos.

Figura 24

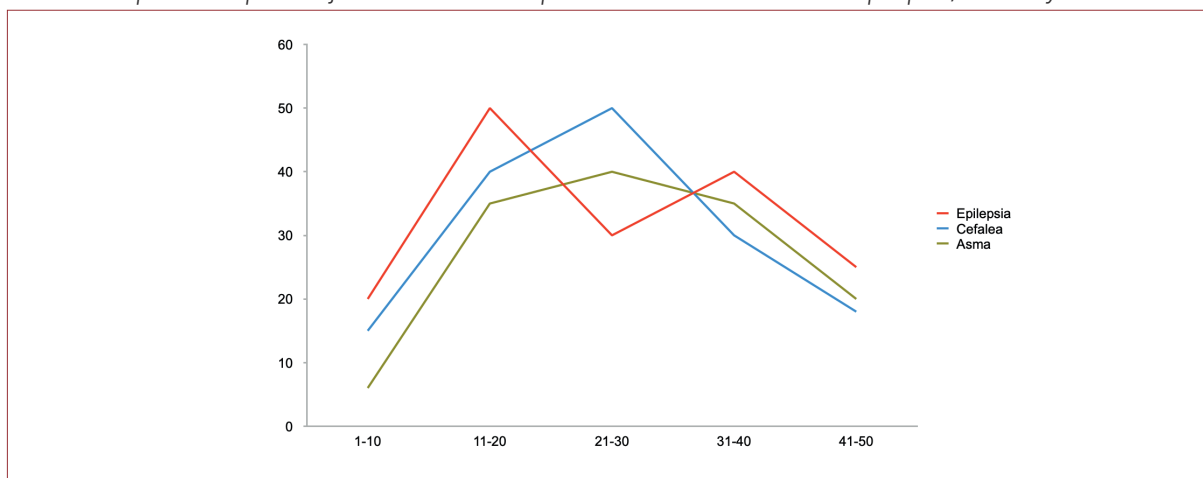
Diagrama de caja



(Elaboración propia, datos simulados)

Figura 25

Número de pacientes por franja de edad en unas poblaciones de enfermos de epilepsia, cefalea y asma.

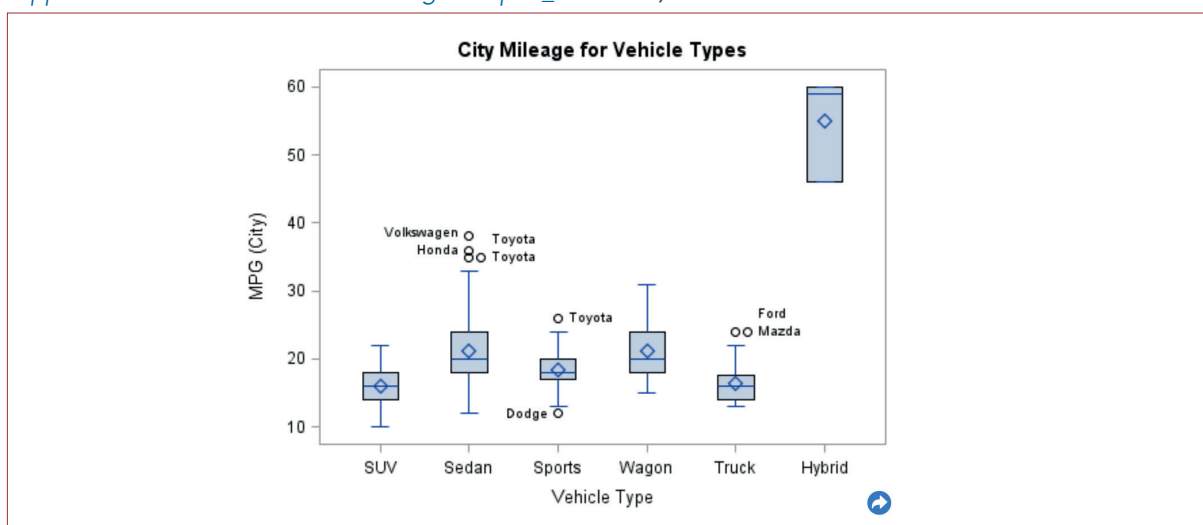


(Elaboración propia, datos simulados)

Centrémonos en el primero de los casos. Cuando nuestro objetivo es representar la distribución de un único conjunto de valores, las gráficas de distribución simple que mejor codifican esta relación son el histograma (figura 22), el polígono de frecuencia (figura 23) y el diagrama de caja (figura 24).

Figura 26

Distribución de vehículos según el modelo y el número de millas que recorren por galón de combustibles (https://support.sas.com/rnd/datavisualization/gtl/boxplot_sect2.htm)



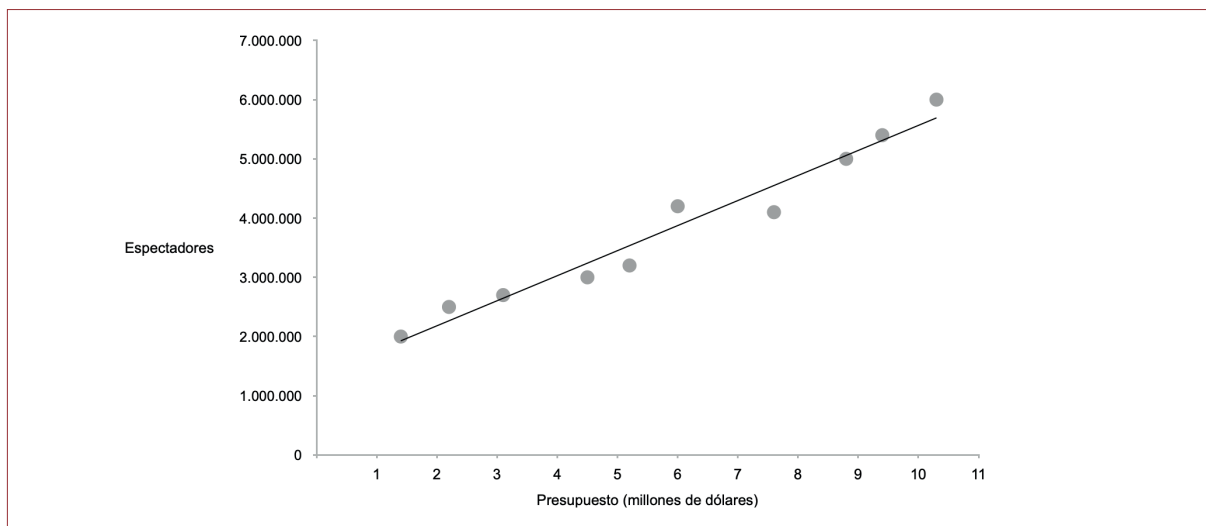
Abordemos ahora el segundo de los casos. Cuando nuestro objetivo es representar simultáneamente la distribución de dos o más conjuntos de valores, las gráficas de distribución múltiple que mejor codifican esta relación son la articulación de diferentes polígonos de frecuencia en una misma gráfica (figura 25) y la combinación de distintos diagramas de caja en la misma propuesta visual (figura 26).

3.7. Visualización de la relación correlación

Abordemos ahora la visualización del último de los contenidos semánticos. Como ya vimos anteriormente, una relación de correlación asigna o conecta dos valores cuantitativos (cada uno perteneciente a una variable cuantitativa distinta) a cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la identificación de algún patrón (fuerte o débil) en la variación entre esos dos tipos de valores numéricos relacionados con esos ítems.

Figura 27

Presupuesto y espectadores de una de las películas estrenadas la última semana.

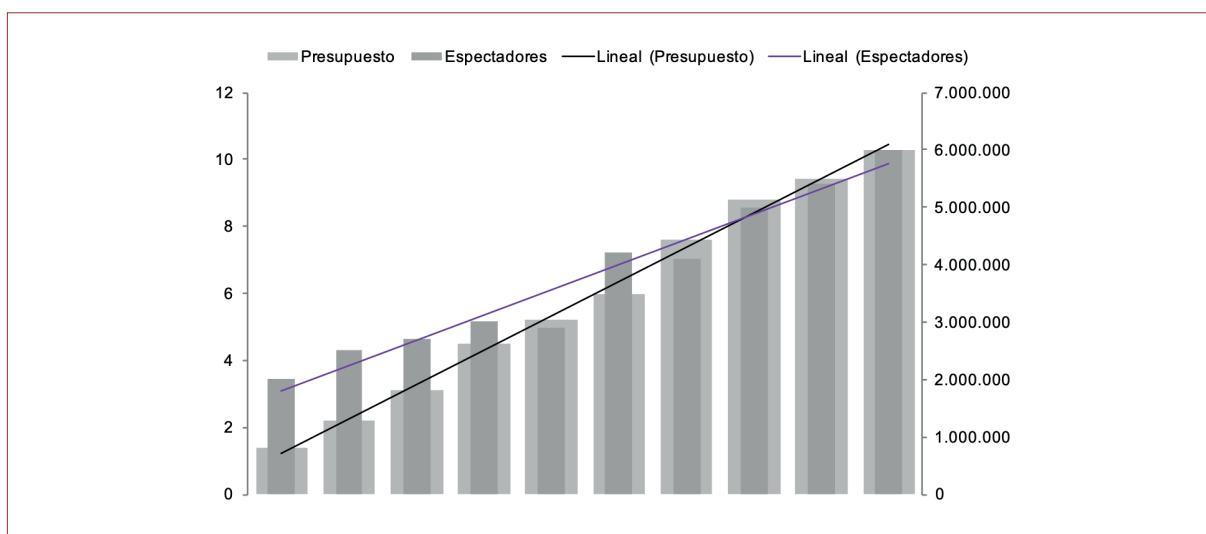


(Elaboración propia, datos simulados)

Por tanto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente los dos valores numéricos y la categoría con la que se encuentran asociados y que permitan expresar si existe o no una situación en los que los valores de una de las dos variables cuantitativas se modifican o cambian de manera sistemática con respecto a los valores de la otra.

Figura 28

Presupuesto y espectadores de las películas estrenadas la última semana.

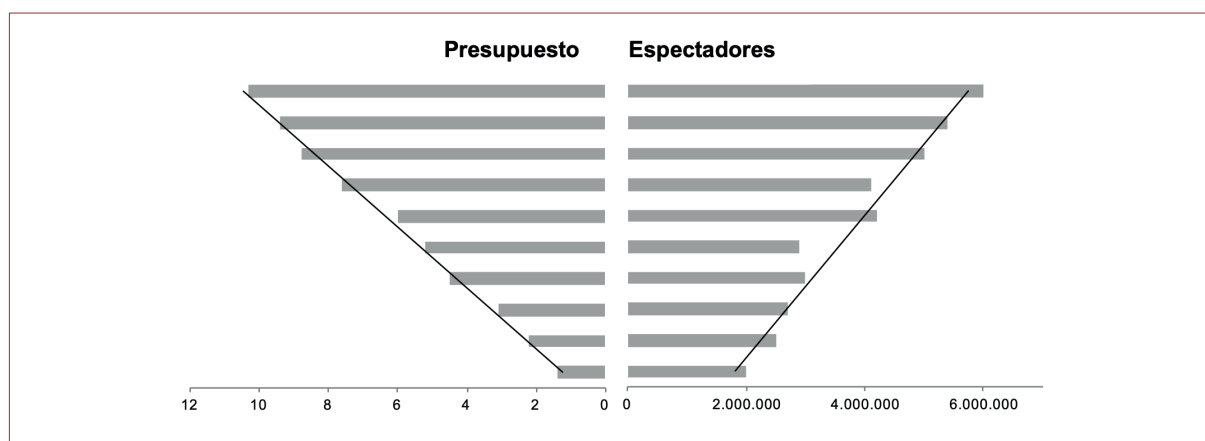


(Elaboración propia, datos simulados)

La propuesta visual que habitualmente se utiliza en la mayoría de los contextos comunicativos para representar este tipo de contenido semántico es el gráfico de dispersión (scatter plot, en inglés), que puede ser entendido como una variante particular de la gráfica de puntos (figura 27).

Figure 29

Presupuesto y espectadores de las películas estrenadas la última semana.



(Elaboración propia, datos simulados)

Algunos autores defienden que, principalmente en el contexto empresarial, el gráfico de dispersión presenta dificultades de interpretación entre los usuarios que no están familiarizados con este tipo de propuestas visuales (Few, 2012). Como alternativas más intuitivas, cuando estamos tratando de visualizar conjuntos de datos no muy extensos, es posible utilizar una gráfica de columnas de correlación (figura 28) o una gráfica de barras emparejadas (figura 29).

4. Nota

Este trabajo es fruto de la revisión, actualización y edición de una parte de un trabajo anterior (Pérez-Montoro, 2022).

5. Financiación

Este trabajo forma parte del proyecto "Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)". Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

6. Referencias

- Cherven, K. (2015). *Mastering Gephi Network Visualization*. Birmingham: Ed. Packt Publishing Ltd.
- Few, S. (2012). *Show me the numbers*. Oakland: Analytics Press.
- Hearn, D. (2011). *Computer graphics with OpenGL*. Boston: Pearson.
- Hugues, J. F., & Van Dam, A. (2013). *Computer Graphics: Principles and Practice*. Boston: Addison-Wesley.

- Jones, B. (2014). *Communicating data with Tableau: Designing, developing, and delivering data visualizations*. London: Ed. Safari Books Online.
- Khan, A. (2016). *Jumpstart Tableau: A step-by-step guide to better data visualization*. London: Ed. Apress.
- Pérez-Montoro, M. (2022). *Comunicación visual de la información. Qué y cómo podemos narrar con datos*. Río de Janeiro: IBICT-UNESCO.
- Shirley, P. (2009). *Fundamentals of computer graphics*. Natick, MA: AK Peters.
- Shneiderman, B. (1996). "The eyes have it: A task by data type taxonomy for information visualizations". *Proceedings of the IEEE Symposium on Visual Languages*, IEEE Computer Society Press, pp. 336-343.
- Zelazny, G. (2001). *Say it with charts*. New York: MacGraw-Hill.

¿Cómo movilizar las audiencias en medios de comunicación? Modelos de Activación de la Relevancia y el Engagement de Audiencias en Medios de Comunicación (MAREA)

Llúcia Castells-Fos

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0003-1124-7941>

Lluís Codina

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0001-7020-1631>

Carles Pont-Sorribes

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0003-1135-4245>

Castells-Fos, L., Codina, L., & Pont-Sorribes, C. (2025). ¿Cómo movilizar las audiencias en medios de comunicación? Modelos de Activación de la Relevancia y el Engagement de Audiencias en Medios de Comunicación (MAREA). En J. Guallar, M. Vállez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 110-125). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicecom.08.esp>

Resumen

LA MAREA ofrece una metodología práctica para que los medios de comunicación evalúen y refuercen su capacidad de relevancia y engagement, elementos cruciales para su viabilidad económica a largo plazo. Este marco metodológico operacionaliza la sostenibilidad mediática a través de tres dimensiones fundamentales: reputación, visibilidad y fidelización de audiencias. MAREA se basa en un conjunto de recomendaciones de expertos — 337 acciones y 226 métricas — sintetizadas en siete parámetros de acción: Escala, Orientación al Usuario, Desarrollo Organizacional, Abandono, Valor Diferencial, Transparencia y Reconocimiento. Estos parámetros, junto con métricas claras, permiten a los medios de comunicación —especialmente a los de tamaño reducido y recursos escasos— identificar áreas de mejora y elaborar planes de acción concretos. El capítulo detalla cada parámetro, ilustra su aplicación con un caso hipotético y subraya la importancia de reinterpretar las métricas más allá de la mera visibilidad, integrando la reputación y la fidelización de la audiencia, lo que facilita una gestión proactiva y una mayor resiliencia en un entorno mediático cambiante.

Palabras clave

Sostenibilidad de los medios de comunicación; Relevancia; Engagement; Audiencia; MAREA; Marco metodológico.

I. Introducción

La MAREA es una herramienta para analizar la relevancia de un medio y su capacidad de *engagement*. El objetivo es, en primer lugar, disponer de un procedimiento para evaluar la situación de un medio de comunicación en relación con un conjunto de acciones que han mostrado su capacidad para potenciar la relevancia y el *engagement*. En segundo lugar, sirve para identificar las estrategias que resultarían prioritarias afrontar si los responsables del medio necesitan aplicar mejoras.

El propósito de la MAREA es ampliar lo más posible el conjunto de medios sobre el que se puede aplicar. Para ello, propone cuestiones que pueden responderse con información al alcance de la mayoría de los medios sin realizar inversiones de dinero y sin una gran exigencia de dedicación temporal. Los medios con un cierto tamaño organizativo suelen disponer de equipos de análisis que cumplen todos o parte de los objetivos que se propone esta investigación. Los medios con recursos limitados, en cambio, no suelen disponer de esta capacidad; es aquí donde la MAREA tiene su campo natural de aplicación.

Su punto de partida fueron los análisis exhaustivos de las buenas prácticas y recomendaciones propuestas por expertos y *think tanks* de la industria de los medios (Castells-Fos et al., 2025). A lo largo de estas investigaciones, se observó, sin embargo, que la relevancia y el *engagement* necesitan de mayor concreción para poder ser activados en el día a día de los profesionales de los medios (Castells-Fos et al., 2022), tal y como se propone la MAREA.

Con el fin de responder a esta necesidad, propusimos desagregar el concepto de relevancia en dos dimensiones: (1) reputación y (2) visibilidad. Y, respecto del *engagement*, decidimos enfocarnos en una de sus dimensiones, (3) la fidelización de audiencias, dado que es el aspecto que mayor relación presenta con la sostenibilidad económica de los medios (Castells-Fos et al., 2023b). A estas tres dimensiones fundamentales las denominaremos, a partir de ahora, Dimensiones de Sostenibilidad Consideradas o DSC.

Las tres dimensiones propuestas han mostrado tener un impacto positivo en la generación de ingresos de los medios de comunicación. Por tanto, se considera que si los medios definen estrategias, líneas de actuación y objetivos expresamente dirigidos a incrementar su reputación, su visibilidad y la fidelización de sus audiencias, estarán contribuyendo de forma decisiva a la sostenibilidad de sus organizaciones.

La base de datos de recomendaciones que compone la MAREA se clasifica en dos categorías: acciones y métricas. Concretamente, 337 acciones y 226 métricas. La naturaleza de cada tipo de recomendación requiere un tratamiento particular y juega un rol muy diferente dentro del modelo de activación.

Las recomendaciones de la categoría acción nos permiten conocer el nivel de desarrollo de las DSC —reputación, visibilidad y fidelización de audiencias—, además de establecer prioridades para el diseño de un plan de activación de dichas dimensiones. Por otro lado, las métricas ayudarán a los medios a establecer un punto de partida, a partir del cual evaluar la efectividad de los planes de mejora que se podrían activar como consecuencia de la aplicación de la MAREA.

Este capítulo es una versión reducida del informe completo de la MAREA publicado de forma independiente y en acceso abierto gracias a la Cátedra Ideograma de la Universitat Pompeu Fabra (Castells-Fos et al., 2025). El informe presenta la metodología con todo detalle y ofrece explicaciones completas de cada una de las secciones abordadas en este capítulo.

2. El modelo de análisis de la MAREA

2.1. Acciones

Las 337 recomendaciones categorizadas como acciones se sintetizaron en indicadores y estos se agruparon en parámetros. El trabajo de extracción de datos y su posterior análisis arrojaron siete parámetros que contienen un total de 57 indicadores. El n.º de menciones recoge las veces que una recomendación se repite en la base de datos analizada, estableciendo un mayor consenso entre los expertos.

Los parámetros de la tabla 1 se presentan ordenados por número de menciones. Si bien el de “escala” recoge más de un tercio de las recomendaciones, el consenso de los expertos se evidencia más elevado en torno a los parámetros de “orientación a usuario” y “valor diferencial”, con un 10,7 y un 9,3 respectivamente, frente al 5,8 de “escala”. Esto indicaría, por un lado, una lista de recomendaciones más acotada y enfocada en el caso de «valor diferencial», y, por otro, la existencia de parámetros donde las posibilidades al alcance de los medios son más diversas, como evidencia el caso de “escala”.

Tabla 1

Listado de parámetros de acción.

Parámetros de acción	N. menciones	N. indicadores	Promedio menciones / indicador
Escala	116	20	5.8
Orientación a usuario	64	6	10.7
Desarrollo organizacional	46	6	7.7
Abandono	32	8	4.0
Valor diferencial	28	3	9.3
Transparencia	26	4	6.5
Reconocimiento	25	10	2.5
Total de menciones	337	57	5.9

Fuente: elaboración propia.

Parámetro “escala”

El parámetro denominado “escala” recoge los indicadores que ayudan a maximizar la presencia del medio en el entorno digital. En la tabla 2 se exponen los 20 indicadores del parámetro con una breve definición y el número de veces (o menciones) que se repiten en la base de la evidencia, totalizando 116 menciones, y un promedio de 5,8 menciones por indicador.

Tabla 2

Indicadores del parámetro “escala”.

Indicadores de escala	N. menciones	Descripción
Formatos	35	Generar diversidad de formatos de distribución.
Redes	26	Conjunto de acciones que tienen el objetivo de maximizar la presencia del medio en las redes sociales. Una parte muy importante de las audiencias accede a las noticias a través de las redes, y no a través del portal del medio.
SEO	17	La implementación del SEO necesita un doble nivel: el de la redacción, con redactores formados en SEO; y el estratégico, con una visión global.
Estrategia de precios	6	Definir una estrategia de precios para nuevos suscriptores / miembros / usuarios de pago.
Modo experimentación	5	El modo experimentación no requiere media labs. Cualquier medio puede experimentar con nuevos servicios, contenidos, diseños de página, formatos multimedia o incluso ofertas de suscripción. El objetivo es emprender un camino de mejora continua y de conocimiento de la audiencia.
Sindicación	3	Llegar a acuerdos con otros medios para re-publicar contenidos puede formar parte de una estrategia de ampliación de audiencias.
Adquisición	3	Realizar campañas para incrementar el tráfico. Además de las acciones regulares en las redes u otras plataformas, los medios pueden diseñar campañas de promoción de la marca o de los valores del medio, entre otros objetivos.
Llamadas a la acción	3	Resaltar las llamadas a la acción para comunicar de forma regular que el medio necesita el apoyo financiero de la audiencia. Este indicador se refiere a la ubicación óptima de la llamada a la acción en el portal del medio. Esta posición óptima debe consistir en lograr una adecuada visibilidad evitando los banners invasivos.
Branded content	2	Desarrollar una estrategia de content marketing o branded content. Esto supone crear contenidos financiados por un anunciante que le ayuden a posicionar su marca.

Indicadores de escala	N. menciones	Descripción
Referidos	2	Tener un programa de referidos para suscriptores. Estos programas pueden poner en marcha procesos de bola de nieve para incrementar suscriptores y ampliar la visibilidad de sus noticias. A modo de ejemplo, y entre otras características, el programa de referidos puede consistir en facilitar al suscriptor el envío y acceso a un número limitado de noticias a familiares o amigos.
Agrupación de audiencias	2	Desarrollar acuerdos con otros medios para generar audiencias más atractivas que vender a los anunciantes.
Influencers	2	Colaborar con influencers y/o firmas con reconocimiento amplio entre la población. El fenómeno de los influencers es una realidad que los medios pueden aprovechar de forma legítima para su beneficio.
First-party data	2	Recopilar datos «de propiedad» o first-party data.
Perfiles de autor	2	Fomentar la presencia activa del equipo de redacción en redes sociales. Los medios deben definir políticas en relación a la presencia de sus periodistas en las redes. Las políticas pueden incluir la recomendación de participar activamente en redes. Las formas y procedimientos deben consensuarse.
Multilingüismo	1	Publicar en más de una lengua en determinados contextos nacionales o internacionales. Ampliar la cobertura lingüística puede tener un impacto directo en la ampliación de la base de audiencias. Esto es efectivo en cualquier caso, pero de forma muy específica en comunidades con más de un idioma oficial.
Accesibilidad	1	La accesibilidad en el contexto digital se refiere a la construcción de una web sin barreras para determinadas discapacidades o situaciones de discapacidad. Una de las más significativas es la de subtitular los materiales audiovisuales. De forma adicional, favorece la visibilidad de este tipo de contenido al permitir su indexación por los buscadores.
Automatización	1	Automatizar la producción de noticias basadas en compilaciones o agregaciones de datos mediante APIs (application programming interface) u otras herramientas simples. La automatización de algunas categorías de noticias, altamente basadas en datos, puede ampliar la capacidad de cobertura del medio. Esto puede mejorar su viabilidad al incrementar su capacidad de producción. Debe haber transparencia en tales casos.
Volumen	1	Definir una estrategia para generar un mayor volumen de contenidos. El volumen no es intrínsecamente bueno, pero en el caso de bajos umbrales de producción, incrementarlo tiene un efecto positivo en la visibilidad del medio. Es decir, los medios deben, al menos, valorar el nivel de producción que realizan teniendo en cuenta el objetivo anterior.
Prueba gratuita	1	En el caso de medios con muro de pago, consiste en ofrecer un periodo de prueba con acceso gratuito para ganar posibles suscriptores.
Distribución	1	Tener acuerdos de distribución de contenidos con empresas o sociedades (no necesariamente medios de comunicación). Estos acuerdos pueden ayudar a los medios a ampliar sus audiencias.
Menciones totales	116	

Fuente: elaboración propia.

Parámetro “orientación a usuario”

El parámetro denominado «orientación a usuario» aglutina los indicadores que tienen como objetivo focalizar en el usuario las actividades del medio de comunicación. En la tabla 4 se recogen los seis indicadores del parámetro con el número de veces que se repiten en la base de la evidencia, totalizando 64 menciones y un promedio de 10,7 menciones por indicador.

Tabla 3
Indicadores del parámetro “orientación a usuario”

Orientación a usuario	N. menciones	Descripción
Participación	24	Fomentar la participación de la audiencia a través de cualquier tipo de canal, tanto digital como analógico.
Conocimiento de la audiencia	17	Solicitar retroalimentación de la audiencia a través de cualquier canal con el fin de monitorizar el consumo de los usuarios para analizar sus necesidades e intereses. Conocer los intereses y los valores de la audiencia es un aspecto clave para los medios.
Funnel para la mejora continua	9	Definir un <i>funnel</i> o embudo de conversión (de usuario gratis a usuario de pago) que sirva como base para elaborar un plan de mejora continua de la experiencia web del usuario. Esto implica fundamentalmente analizar comportamientos navegacionales.
Personalización	8	Personalizar la comunicación con el usuario en base a la monitorización de su comportamiento.
Lenguaje cercano	3	Utilizar un lenguaje cercano, asequible y adecuado para comunicar con la audiencia. Por ejemplo, algunos medios practican el periodismo explicativo. En tal caso, es necesario reflexionar sobre el registro del lenguaje utilizado. Otros medios pueden buscar una audiencia joven en temas relacionados con estilos de vida, etc.
Política de comentarios	3	Definir una política para responder a los comentarios. Algunos medios han eliminado la posibilidad de enviar comentarios, pero otros los mantienen. En este caso, es necesario establecer políticas claras sobre los mismos, para evitar que generen el desgaste del medio como marca.
Menciones totales	64	

Fuente: elaboración propia.

Parámetro “desarrollo organizacional”

Cuando los expertos abordan temas de sostenibilidad, siempre señalan la importancia de los aspectos organizacionales. En concreto, este parámetro agrupa a los indicadores vinculados con los recursos humanos y tecnológicos necesarios en una organización periodística. La tabla 6 recoge los seis indicadores del parámetro, con sus 46 menciones, y un promedio de 7,7 menciones por indicador.

Tabla 4
Indicadores del parámetro “desarrollo organizacional”.

Desarrollo organizacional	Nº menciones	Descripción
Herramientas	24	Consiste en dotarse de los recursos tecnológicos necesarios para facilitar la tarea periodística en conjunción con el resto de la organización y con enfoque a dar servicio a las audiencias.
Equipo	12	Consiste en disponer de personas dedicadas a las siguientes tareas, definidas sin ánimo de exhaustividad: diseño web, tecnología, gestión, desarrollo y análisis de audiencia, programa de membresía, venta de espacios publicitarios, dinamización de la comunidad, servicios al cliente, y comunicación corporativa; tareas que se consideran clave y complementarias al equipo de redacción.
Formación	7	Consiste en llevar a cabo formaciones y/o reuniones para mejorar el conocimiento operativo que garantice que el equipo obtiene el máximo partido de las herramientas a su disposición. Es importante también compartir buenas prácticas y los objetivos de los diferentes equipos del medio.
Reportes	1	Realizar reportes estandarizados para los anunciantes.
Retención de talento	1	Desarrollar estrategias de retención del talento en la redacción.
Interconexión	1	Conseguir la interconexión de los diferentes sistemas de gestión (en especial los que contienen datos de los usuarios, suscriptores, etc.), aporta sinergias a los procesos y evita información redundante.
Menciones totales	46	

Fuente: elaboración propia.

Parámetro “abandono”

Existe un amplio consenso entre los expertos sobre la relativa facilidad para captar nuevos suscriptores con motivo de campañas especiales y ofertas de suscripción muy reducidas, a veces por debajo del 50 % del coste tarifado normal. En cambio, hay un auténtico problema con las bajas o los ratios de cancelación (*churn*). Por tanto, no es sorprendente que en los informes de los think tanks y entre los expertos haya recomendaciones para mantener lo más bajo posible el abandono de suscriptores, que puede producirse por una gran variedad de causas, como se podrá ver en el desarrollo de los indicadores.

En concreto, bajo el parámetro denominado “abandono” se agrupan los indicadores que tratan de incrementar la tasa de retención de los usuarios de pago —o de reducir su tasa de abandono—. La tabla 8 ofrece un listado de los ocho indicadores que componen el parámetro, con un total de 32 menciones dentro de la base de la evidencia y un promedio de 4,0 menciones por indicador.

Tabla 5
Indicadores del parámetro abandono.

Abandono	Nº menciones	Descripción
Incentivos	10	Ofrecer contenidos exclusivos para suscriptores. Algunos medios no tienen muros de pago absolutos, es decir, las audiencias que no son suscriptores pueden acceder al menos a parte de los contenidos. Otros dejan el acceso abierto a todos, pero añaden prestaciones para suscriptores. En estos contextos, puede ser muy necesario hacer más evidentes las ventajas de ser suscriptor.
Medios de pago caducados	6	Disponer de una herramienta para gestionar los medios de pago caducados, como puede ser el caso muy frecuente de la caducidad de las tarjetas bancarias con las que se realizó el pago de la suscripción. Este indicador es un ejemplo de la diversidad de niveles de este parámetro. Los expertos han detectado que una fracción importante de las bajas entre los suscriptores se deben a cuestiones como esta.
Onboarding	6	Definir un proceso de onboarding para nuevos suscriptores. El onboarding es como se denomina el proceso por el cual las organizaciones incorporan a nuevos empleados. En los medios de comunicación se refiere a la manera en que un suscriptor ingresa a un medio, por primera vez o con un nuevo perfil.
Riesgo de cancelación	4	Definir el nivel de uso por debajo del cual existe un riesgo de cancelación. La experiencia indica que, en general, es posible identificar los suscriptores de un medio que están en riesgo de cancelación mediante la monitorización de sus hábitos. Por ejemplo, suscriptores que hace días que no entran en el portal del medio, o suscriptores que abandonan el portal sin consultar ninguna noticia, etc. Otro indicador puede ser la tasa de apertura de las newsletters. Los medios deben utilizar los datos de uso de sus suscriptores para detectar estas situaciones y realizar acciones, como campañas de comunicación dirigidas a los usuarios con menor nivel de consumo, para revertir la tendencia.
Suscripciones largas	2	Ofrecer descuentos para suscripciones largas. Para un medio de comunicación es especialmente importante conseguir suscriptores de larga duración (un año o más), ya que denotan por parte del usuario un mayor grado de convencimiento en el proceso de suscripción. El medio puede diseñar políticas de descuentos para convertir, por ejemplo, suscripciones mensuales en anuales.
Motivos de cancelación	2	Analizar los motivos de cancelación. Una vez se ha llevado a cabo una cancelación, siempre es posible intentar aprovecharla para saber más sobre los usuarios y sus motivos de abandono. Algunas empresas han desarrollado medios por los cuales se solicita a los suscriptores que cancelan su suscripción cumplimentar formularios simples, pero muy importantes para conocer los motivos.
Recuperaciones	1	Realizar campañas de recuperación de clientes que hayan efectuado cancelaciones. Una vez se ha dado de baja un suscriptor, podemos programar campañas ligeras (deben evitar ser persistentes o agresivas) para intentar motivarle a reiniciar su suscripción.
Fatiga del suscriptor	1	Crear packs de contenido para evitar la fatiga del suscriptor. Los expertos han detectado el fenómeno que señalan con este nombre como un motivo importante de abandono.
Menciones totales	32	

Fuente: elaboración propia.

Parámetro “valor diferencial”

Bajo el parámetro denominado «valor diferencial», se agrupan los indicadores en torno a la necesidad de posicionar el medio de comunicación frente a la competencia. Lo componen tres indicadores que se mencionan 28 veces en la base de datos, lo que supone un promedio de 9,3 menciones por indicador. La siguiente tabla presenta un resumen.

Tabla 6
Indicadores del parámetro “valor diferencial”.

Valor diferencial	N. menciones	Descripción
Posicionamiento	15	Definir una estrategia de posicionamiento que sea diferencial frente a la competencia y única en la percepción del usuario.
Géneros periodísticos	8	Una posible enumeración de los géneros periodísticos puede ser la siguiente: Informativos (noticia, reportaje, entrevista...); Opinión (editorial, columna, crítica); Interpretativos (crónica, perfil, reportaje interpretativo). Este indicador valora la opción estratégica que puede adoptar un medio respecto a su posible priorización como parte de su valor diferencial.
Temáticas	5	Especializarse en alguna temática. Las empresas de medios con muchos recursos humanos y tecnológicos pueden ser generalistas. Sin embargo, empresas medianas y pequeñas, así como, en general, los nuevos emprendimientos, pueden optar por enfocarse en ciertos temas o nichos temáticos.
Menciones totales	28	

Fuente: elaboración propia.

Parámetro “transparencia”

Fenómenos crecientes como la desconfianza en los medios y otros asociados, como la evitación de noticias, han llevado a los expertos a insistir en la transparencia. Elementos éticos muy importantes también inciden en lo mismo. En concreto, este parámetro aglutina a los indicadores que promueven el ejercicio de la transparencia para generar confianza en la audiencia. Como muestra la siguiente tabla, lo componen cuatro indicadores que suman un total de 26 menciones, lo que promedia 6,5 por indicador.

Tabla 7
Indicadores del parámetro “transparencia”.

Transparencia	N. menciones	Descripción
Código ético	15	Disponer de código ético o documento similar que establezca estándares periodísticos éticos y de calidad periodística.
Rendición de cuentas	5	Ser transparentes en la gestión y financiación de la organización. Los expertos insisten en que la transparencia es esencial para conseguir la fidelización de la audiencia. Puede ejercerse mediante la publicación de información corporativa.
Verificación	4	Diseñar procesos consistentes, bien definidos y rigurosos de verificación de datos por parte de la redacción.
Información vs. opinión y otros	2	Diferenciar claramente secciones patrocinadas como el branded content o la publicidad, así como los contenidos de opinión respecto de contenidos informativos.
Menciones totales	26	

Fuente: elaboración propia.

Parámetro "reconocimiento"

El parámetro agrupa un conjunto de indicadores que, puestos en práctica y de acuerdo con los expertos consultados, pueden suponer una mejora en el reconocimiento por parte del público. Según la tabla siguiente, los 10 indicadores que componen el parámetro reúnen un total de 25 menciones, lo que supone un promedio de 2,5 por indicador.

Tabla 8
Indicadores del parámetro "reconocimiento"

Reconocimiento	Nº menciones	Descripción
Mediatización	6	Definir una estrategia mediática para el equipo de redacción y dirección que se aplique en sus apariciones públicas.
Colaboradores	4	Captar firmas de prestigio que colaboren con el medio.
Divulgación	3	Publicar con objetivo de divulgación. Se trataría de publicar contenidos de divulgación o piezas de periodismo explicativo en momentos clave, o como parte de una política general.
Proyectos conjuntos	2	Realizar proyectos periodísticos con otros medios. En el periodismo de investigación, el desarrollo de proyectos conjuntos entre medios se ha revelado muy eficaz.
Innovación	2	Desarrollar procesos de innovación que tengan como objetivo mejorar el producto periodístico. Esta cultura de innovación y mejora continua favorece la imagen del medio entre sus stakeholders.
Equilibrio calidad vs. cantidad	2	Definir una estrategia que equilibre el volumen de publicaciones con la calidad de las producciones. En otro indicador se señala la conveniencia de aumentar el volumen de producción en algunos contextos; ahora se trata de equilibrar el volumen con la calidad.
Exclusivas	2	Definir una política de exclusivas en algún ámbito del medio en el que tenga fortalezas, que ayude a caracterizarlo delante de las audiencias y los stakeholders en general.
Ecuanimidad	2	Alejarse del partidismo en la cobertura mediática. En temas esenciales para los valores del medio, así como en aspectos relacionados con derechos humanos y similares, no es factible mantener posiciones de imparcialidad. Pero hay muchas otras situaciones en las que los medios pueden exhibir esa capacidad y mostrarse ajenos a luchas o debates con un carácter muy partidista.
Patrocinadores	1	Escoger patrocinadores que estén relacionados con el posicionamiento de la marca del medio. Los medios de comunicación deben establecer criterios sobre la esponsorización o el mecenazgo que aceptan, para que no condicione su línea editorial o la percepción de condicionamiento por parte de las audiencias. También, porque el mecenazgo adecuado puede contribuir a construir la reputación del medio.
Premios	1	Presentar candidatura a premios periodísticos.
Menciones totales	25	

Fuente: elaboración propia.

Ejemplo de "aplicación"

A continuación, se presenta el resultado de la matriz aplicada a un ejemplo hipotético, donde se compara el resultado conseguido con el idóneo, que supondría cumplir con todas las recomendaciones de la MAREA. El caso hipotético que nos sirve de ejemplo se identifica como «caso», y el resultado de cumplir con todos los requerimientos como «idóneo». Cada medio puede decidir evaluar todos o parte de los indicadores propuestos, según su perfil concreto. Por simplicidad, se simula un caso en el que los responsables del medio dan respuesta a todos. El parámetro que se toma para el ejemplo es Desarrollo Organizacional.

La tabla 9 muestra el resultado de la MAREA aplicada a un caso hipotético. El resultado máximo que conseguiría un medio idóneo es 39,5 y el obtenido en el caso es 29,9. El gráfico 1 muestra este nivel de consecución del 76 %.

Tabla 9

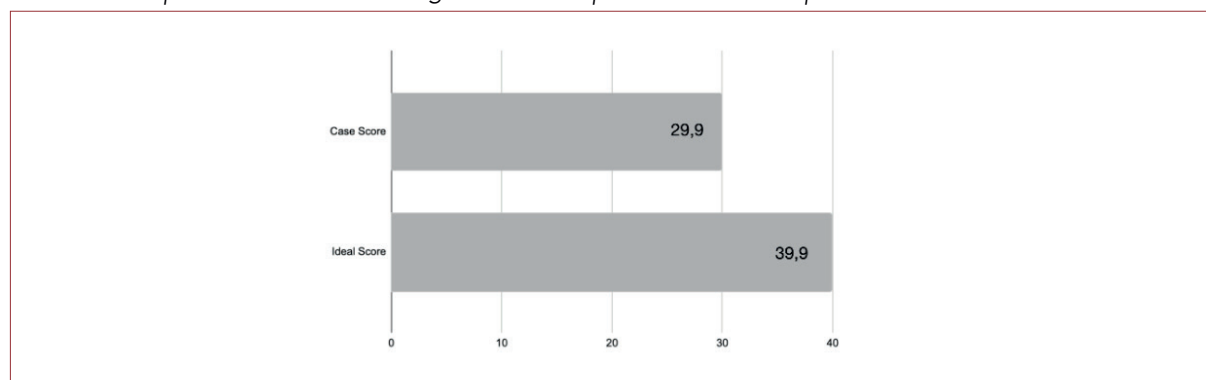
Aplicación del parámetro “desarrollo organizacional” sobre un caso hipotético.

Desarrollo organizacional	Puntuación		FdF			Resultados	
	Idóneo	Caso	FdF1	FdF2	FdF3	Idóneo	Caso
Herramientas	1	1	4.0	3	2	24.0	24.0
Equipo	1	0	1.6	3	2	9.4	0.0
Formación	1	1	0.9	3	2	5.5	5.5
Reportes	1	1	0.1	1	1	0.1	0.1
Retención de talento	1	0	0.1	1	2	0.3	0.0
Interconexión	1	1	0.1	1	2	0.3	0.3
Total						39.5	29.9

Fuente: elaboración propia.

Gráfico 1

Resultado del parámetro “desarrollo organizacional” aplicado a un caso hipotético .



Fuente: elaboración propia.

2.1. Métricas

La principal aportación de la MAREA consiste en su voluntad de activar en la práctica profesional diaria, de forma ordenada y priorizada, una batería de acciones que contribuyan a construir medios más sostenibles. En este sentido, las métricas juegan un papel secundario. La MAREA quiere dar respuesta a la ineludible y responsable necesidad de medir utilizando las métricas disponibles por imperfectas que sean (Carlson, 2018).

Es por esta razón que, tal y como se describe en el capítulo 2 del informe completo (Castells-Fos et al., 2025), Objetivos generales y procedimientos, se decide utilizar los indicadores métricos simples y concretos sobre los que existe un mínimo consenso entre los expertos consultados. Estos suponen el 53% de las recomendaciones métricas recogidas en la base de datos.

Tabla 10

Listado de indicadores métricos simples concretos

Indicadores métricos simples concretos	Nº menciones	Descripción
Usuarios únicos	15	Personas diferentes que visitan una web en un periodo de tiempo.
Usuarios de pago	13	Personas, empresas o instituciones que han contribuido a financiar el medio en un período de tiempo. Se recomienda una visión acumulada anual.
Usuarios recurrentes	13	Personas que visitan la web más de una vez en un periodo determinado de tiempo.
Ingresos	10	Totalidad de los ingresos generados con independencia de su origen. Se recomienda una visión acumulada anual.
Seguidores	8	Número de seguidores o miembros de los perfiles de redes sociales del medio en un período de tiempo determinado.
Interacciones	8	Totalización de las posibles interacciones simples: comparticiones, likes y comentarios en redes sociales en un tiempo determinado.
Newsletter	7	Número total de usuarios registrados en las diferentes newsletter del medio en un período de tiempo determinado.
Páginas vistas	6	Número total de páginas visitadas en un período de tiempo determinado.
Duración	5	Duración media de una visita.
Experiencia	4	Número de años desde el inicio de la operación.
Ratio apertura	4	Se refiere al porcentaje de apertura de los emails enviados. Número de veces que se abre un email sobre el número total de enviados en un período de tiempo. El medio puede decidir totalizar la métrica para el total de emails que envía, o analizar la newsletter que considere más importante.
Rebote	3	Es el porcentaje de sesiones de la página en las que no se ha producido una interacción significativa.
Tráfico referido	3	Número de visitas que recibe la web desde dominios diferentes al propio del medio.
Clics en newsletter	3	El número de clics en la newsletter en un período de tiempo determinado.
Premios recibidos	2	Número de reconocimientos públicos recibidos por su labor periodística.
Usuarios de pago sobre audiencia	2	Porcentaje resultante de dividir el número de usuarios de pago (en forma de suscripción, membresía o donaciones) entre el número de usuarios únicos en un tiempo determinado.
Bajas	2	Porcentaje de suscriptores que se han dado de baja en un tiempo determinado.
Menciones totales	121	

Fuente: elaboración propia.

3. Conclusiones

El modelo MAREA logra articular un marco de trabajo para la sostenibilidad de los medios al activar tres dimensiones críticas: reputación, visibilidad y fidelización de audiencias. El modelo resulta fácil de aplicar y útil para definir un plan de mejora, ya que identifica las acciones a priorizar.

Sus siete parámetros de acción cubren un amplio espectro que va desde temas estratégicos, como el valor diferencial, hasta cuestiones tácticas, como el abandono. A continuación, se presentan las principales conclusiones extraídas de cada uno.

3.1. Parámetros de acción

Escala: Este es el parámetro con más menciones, con un enfoque técnico y una fuerte conexión con la visibilidad. Sin embargo, su efectividad mejora si se apoya en la reputación del medio, utilizando indicadores como el SEO, la sindicación o la adquisición de usuarios. La «escala» acumula el mayor número de indicadores de MAREA, posiblemente porque los expertos ofrecieron recomendaciones muy variadas, con poco consenso en algunos casos (de 20 indicadores, seis tienen solo una mención).

Orientación al usuario: Con una alta tasa de consenso (10.7 menciones por indicador), este parámetro se basa en el conocimiento de la audiencia para acercarse a ella a través de la personalización o la participación. Este proceso genera una gran cantidad de información que debe utilizarse para la mejora continua.

Desarrollo organizacional: Un parámetro eminentemente estratégico que busca sentar las bases para el crecimiento sostenible y a largo plazo de la organización. Su objetivo es dotar al medio del equipo y las herramientas necesarias para potenciar su capacidad de acción y generar dinámicas de trabajo eficientes.

Abandono: Es un parámetro táctico e indispensable para mantener los ingresos por suscripciones. Su importancia radica en un entorno de alta competencia, abundancia de productos sustitutivos y la doble amenaza del “techo” de suscriptores y la fatiga de la audiencia.

Valor diferencial y Transparencia: Estos dos parámetros son estratégicos y definitorios para los medios. El primero es fundamental para la calidad del contenido, mientras que el segundo lo es para la credibilidad del proceso de su generación.

Reconocimiento: Este es el segundo parámetro en número de indicadores, lo que se traduce en un consenso relativamente bajo (10 indicadores con 2,5 menciones/indicador, frente a un promedio de 5,9 en la base de datos total). Esto refleja el esfuerzo de los expertos por generar ideas para accionar la reputación, un concepto que a lo largo de esta investigación ha aparecido también como “impacto”, “influencia” o “confianza”.

3.2. Conclusiones sobre las métricas

MAREA propone una lista de métricas accesibles para todos los medios con un esfuerzo mínimo, lo que ayuda a la toma de conciencia antes de definir un plan de mejora. Las métricas de tamaño son las que presentan mayor consenso entre los expertos. Tanto la industria como la academia parecen haberse resignado a trabajar con las herramientas disponibles, siempre con la intención de reinterpretar los resultados en clave de reputación y fidelización, y no solo de visibilidad.

Esta reinterpretación lleva a una mutualización de las métricas: de los 24 indicadores métricos que se facilitan, la mayoría miden dos o tres dimensiones. Dieciséis se proponen para la reputación, 12 para la visibilidad y 17 para la fidelización.

Para finalizar, recordamos que la MAREA está basada, entre otras investigaciones, en opiniones de los expertos recogidas a lo largo de una batería de investigaciones realizadas entre 2022 y 2024 sobre la evolución de un sector que cambia rápidamente. No obstante, la mayor parte de los indicadores facilitados deberían resistir bastante bien cualquier cambio al menos a medio plazo.

Conscientes en todo caso de las posibilidades reales de cambio constante, la principal aportación de la MAREA pretende ser a la vez metodológica y conceptual, aportando la base principal de un modelo de análisis que puede ser enriquecido y modificado de forma libre por parte de los profesionales de los medios de comunicación y los estudiosos de los medios. Esto les permitirá contar con más herramientas para enfrentar los cambios a los que están sujetos los indicadores que impactan en la sostenibilidad de los medios de comunicación.

4. Financiación

Este trabajo forma parte del proyecto "Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)". Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

5. Referencias citadas y consultadas

Este apartado recoge las referencias que han contribuido a la construcción de la MAREA.

ADEPA. (2022). *Industria periodística 2021. Anuario 2021*. <https://adepa.org.ar/wp-content/uploads/2021/12/anuario-adepa-2021.pdf>

Amoedo-Casais, A., Vara-Miguel, A., Negredo-Bruna, S., Moreno-Moreno, E., & Kaufmann-Argueta, J. (2021). *Digital News Report España*. <https://dadun.unav.edu/handle/10171/62862>

Carlson, M. (2018). Confronting measurable journalism. *Digital Journalism*, 6(4), 406-417. <https://doi.org/10.1080/21670811.2018.1445003>

Castells-Fos, L., Codina, L., & Pont-Sorribes, C. (2025). *Modelos de activación de la relevancia y el engagement de audiencias en medios de comunicación*. UPF. Cátedra Ideograma. <https://doi.org/10.31009/com.2025.01>

Castells-Fos, L., Lopezosa, C., Pont-Sorribes, C., & Codina, L. (2023). Reputación, visibilidad y fidelización de audiencias de medios: una propuesta operativa desde la academia. *Estudios Sobre el Mensaje Periodístico*, 29(2), 419-430. <https://doi.org/10.5209/esmp.85655>

Castells-Fos, L., Pont-Sorribes, C., & Codina, L. (2023). Decoding news media relevance and engagement through reputation, visibility and audience loyalty: A scoping review. *Journalism Practice*, 19(5), 1154-1173. <https://doi.org/10.1080/17512786.2023.2239201>

Castells-Fos, L., Pont-Sorribes, C., & Codina, L. (2022). La sostenibilidad de los medios a través de los conceptos de engagement y relevancia: scoping review. *Doxa Comunicación*, 35, 19-38. <https://doi.org/10.31921/doxacom.n35a1627>

Clear, J. (s.f.). *All models are wrong, some are useful*. James Clear. <https://jamesclear.com/all-models-are-wrong>

Day, B., & Wills, F. (2020). 2020 Advertiser assessment of local media digital advertising.

Del Pozo, M. (2022). Netflix se lanza a la publicidad. *Expansión*. <https://www.expansion.com/blogs/peon-de-dama/2022/10/23/netflix-se-lanza-a-la-publicidad.html>

Escolar, I. (2021, 3 julio). Las cuentas de 2020: los socios y socias salvan elDiario.es. *ElDiario.es*. https://www.eldiario.es/escolar/cuentas-socios-socias-salvan-eldiario_132_8102443.html

- Escolar, I. (2020). Un periódico que se debe a su comunidad: las cuentas de eldiario.es en 2019. *ElDiario.es*. https://www.eldiario.es/escolar/cuentas_132_6071575.html
- Espada, A. (2021). *Impact of the COVID-19 Pandemic on Media sustainability in Latin America*. https://unesdoc.unesco.org/ark:/48223/pf0000377631_eng
- Evolok. (2021). *The paywall revolution goes global*. <https://pressgazette.co.uk/subscription-news-media-must-re-think-model-for-creating-new-revenues/>
- Ferrer-Conill, R., & Tandoc, E. C. (2018). The audience-oriented editor: Making sense of the audience in the newsroom. *Digital Journalism*, 6(4), 436-453. <https://doi.org/10.1080/21670811.2018.1440972>
- Forman, C. I. (2021, October 12). Solutions to America's Local Journalism Crisis: Consolidated Literature Review. <https://shorensteincenter.org/solutions-americas-local-journalism-crisis-consolidated-literature-review>
- Graves, A. L., & Kelly, J. (2010). *Confusion online: Faulty metrics and the future of digital journalism*. Tow Center for Digital Journalism. <https://search.isuelab.org/resources/15315/15315.pdf>
- Gundlach, H., & Hofmann, U. (2021). Information search, behavioural economics, and relevance decisions in the online media industry: how strongly do the algorithms of intermediaries influence the relevance evaluation of information? *Journal of Media Business Studies*, 18(3), 179-198. <https://doi.org/10.1080/16522354.2020.1854602>
- Guzik, S. (2022). *The global survey on journalism's futures*. Future Today Institute. https://www.linkedin.com/posts/samguzik_the-global-survey-on-journalisms-futures-activity-6912124871074299904-lmke
- Iastrebnier, M., Warner, J., & Best, J. (2021). Punto de Inflexión.
- Kovach, B., & Rosenstiel, T. (2001). *The elements of journalism* (3rd ed.). Three Rivers Press.
- Krebs, I., Bachmann, P., Siegert, G., Schwab, R., & Willi, R. (2021). Non-journalistic competitors of news media brands on Google and YouTube: From solid competition to a liquid media market. *Journal of Media Business Studies*, 18(1), 27-44. <https://doi.org/10.1080/16522354.2020.1832746>
- Li, A. (2021). How customer service can build trust and engagement with audiences. <https://www.americanpressinstitute.org/publications/reports/strategy-studies/how-customer-service-can-build-trust-and-engagement-with-audiences>
- López-García, G. (2015). *Periodismo digital: redes audiencias y modelos de negocio*. Comunicación Social.
- López-Linares, C. E. (2021). *Innovación y periodismo local en América Latina*. Fundación Gabo. <https://fundaciongabo.org/es/recursos/publicaciones/fundacion-gabo-y-google-news-initiative-lanzan-libro-digital-innovacion-y>
- Lopezosa, C., Codina, L., & Freixa, P. (2022). *ATLAS.ti para entrevistas semiestructuradas : guía de uso para un análisis cualitativo eficaz*. DigiDoc Research Group | Pompeu Fabra University, 30. <http://repositori.upf.edu/handle/10230/52848>
- Matsa, K. E., & Naseer, S. (n.d.). News platform fact sheet. <https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/>

- McKisson, I. & Pallack, B. (2021). What types of messages, pricing, and perks motivate #ThisTucson super users to become paying members.
- Nelson, J. L., & Tandoc, E. C. (2019). Doing “well” or doing “good”: What audience analytics reveal about journalism’s competing goals. *Journalism Studies*, 20(13), 1960-1976. <https://doi.org/10.1080/1461670X.2018.1547122>
- Newman N. (2019). *Journalism, media, and technology trends and predictions 2019*. Reuters Institute. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-01/Newman_Predictions_2019_FINAL_2.pdf
- Newman, N. (2021). *Digital news report: Journalism, media, and technology trends and predictions 2021*. Reuters Institute. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-01/Newman_Predictions_2021_FINAL.pdf
- Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., & Nielsen, R. K. (2022). *Digital news report 2022*. Reuters Institute. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf
- Newman, N., Fletcher, R., Eddy, K., Robertson, C. T., & Nielsen, R. K. (2023). *Digital news report 2023*. Reuters Institute. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf
- Oliver, L. (n.d.). *Private gatekeepers: Encrypted messaging Apps and news audiences*. https://www.cima.ned.org/wp-content/uploads/2022/02/CIMA_EMA-Report_web_150ppi.pdf
- Piechota, G. (2021). *The benefits and risks of media data democratisation*.
- Pope, K. (2017). Is there a business model for real journalism? *Nation*, 304(9).
- Powers, E. (2018). Selecting metrics, reflecting norms: How journalists in local newsrooms define, measure, and discuss impact. *Digital Journalism*, 6(4), 454-471. <https://doi.org/10.1080/21670811.2018.1445002>
- Puente-Rivera, I., Campos-Freire, F & López-García, X (Coords.). (2018). *Periodismo con futuro*. Media XXI.
- PWC. (2021). *Entertainment and media outlook 2021-2025*. PricewaterhouseCoopers. <https://www.pwc.es/es/entretenimiento-medios/entertainment-media-outlook-espana-2021-2025.html>
- Radcliffe, D., & Wallace, R. (2021, October 7). Life at local newspapers in a turbulent era. *Columbia Journalism Review*. https://www.cjr.org/tow_center_reports/life-at-local-newspapers-in-a-turbulent-era-findings-from-a-survey-of-more-than-300-newsroom-employees-in-the-united-states.php
- Roseman, E., McLellan, M., & Holcomb, J. (2021). The state of nonprofit news.
- Schulte, A. (2021). *A quemar las naves. Guía para activar el cambio cultural en medios de comunicación*. WAN-IFRA. <https://wan-ifra.org/wp-content/uploads/2021/04/WAN-IFRA-Cultural-Change-Espanol-New.pdf>
- Sonderman, J., & Vargo, G. (2021). What news publishers do to retain subscribers. <https://www.americanpressinstitute.org/publications/reports/survey-research/what-news-publishers-do-to-retain-subscribers/single-page/#:~:text=Almost all publishers said they,welcome emails to new subscribers>

- Stocking, G., & Khuzam, M. (2021). *Digital news fact sheet*. Pet Research Center. <https://www.pewresearch.org/journalism/fact-sheet/digital-news/>
- Stroud, N. J. (2017). Helping newsrooms work toward their democratic and business objectives. In P. J. Boczkowski, & C. W. Anderson (Eds.), *Remaking the news: Essays on the future of journalism scholarship in the digital age*. MIT Press.
- The Economist. (2023, December 14). Can you have a healthy democracy without a common set of facts? *The Economist*. <https://www.economist.com/leaders/2023/12/14/can-you-have-a-healthy-democracy-without-a-common-set-of-facts>
- The Economist. (2023, December 14). How American journalism lets down readers and voters. *The Economist*. <https://www.economist.com/united-states/2023/12/14/how-american-journalism-lets-down-readers-and-voters>
- Vallés, M. S. (2003). *Entrevistas cualitativas*. Centro de Investigaciones Sociológicas. <https://doi.org/978-84-7476-342-3>
- Villafañe, J., Ortiz-de-Guinea-Ayala, Y., & Martín-Sáez, J.-L. (2020). Reputación de los medios de comunicación españoles. *Profesional de la Informacion*, 29(4). <https://doi.org/10.3145/epi.2020.jul.07>

Factores para aumentar la visibilidad en repositorios digitales: metadatos, interoperabilidad, identificadores persistentes y optimización SEO-GEO

Danilo Reyes-Lillo

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0002-0141-8324>

Cristòfol Rovira

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0002-6463-3216>

Alejandro Morales-Vargas

Universidad de Chile, Chile

<https://orcid.org/0000-0002-5681-8683>

Reyes-Lillo, D., Rovira, C., & Morales-Vargas, A. (2025). Factores para aumentar la visibilidad en repositorios digitales: metadatos, interoperabilidad, identificadores persistentes y optimización SEO-GEO. En J. Guallar, M. Vállez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 126-141). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicom.09.esp>

Resumen

En el contexto de la ciencia abierta y la comunicación académica, mejorar la visibilidad de los repositorios digitales es esencial para maximizar el alcance, el impacto y la capacidad de descubrimiento del contenido que albergan. Este capítulo de libro explora cinco estrategias clave para mejorar la visibilidad de los repositorios: optimización de la calidad de los metadatos, habilitación de protocolos de interoperabilidad, adopción de identificadores persistentes, implementación de acciones de optimización para motores de búsqueda (SEO) y adopción de la optimización para motores generativos (GEO). A través de un análisis detallado y recomendaciones prácticas, este trabajo destaca cómo los metadatos estandarizados, los vocabularios controlados y los identificadores persistentes, como DOI, Handle y ARK, contribuyen a mejorar la visibilidad. Este capítulo también enfatiza la importancia de alinear los repositorios con las tecnologías web en evolución y los motores impulsados por inteligencia artificial para garantizar que el contenido siga siendo accesible, rastreable e integrado en la experiencia de búsqueda de los usuarios.

Palabras clave

Repositorios digitales; Visibilidad; Calidad de metadatos; Interoperabilidad; Identificadores persistentes; Optimización para Motores de Búsqueda; SEO; Optimización para Motores Generativos; GEO.

I. Introducción

En la era del acceso abierto y la ciencia abierta, los repositorios digitales juegan un papel fundamental en la difusión del conocimiento. Incrementar su visibilidad no solo amplía el alcance de los contenidos depositados, sino que también fortalece el impacto institucional y contribuye a la democratización del acceso a la información. Diversos estudios han demostrado que los trabajos disponibles en acceso abierto –especialmente a través de repositorios institucionales– tienden a recibir más citas y resultan más accesibles que aquellos restringidos por barreras de pago (Piwowar et al., 2018; Swan, 2010).

Además, una mayor visibilidad en los motores de búsqueda académicos como Google Scholar o bases de datos como OpenAIRE y CORE permite que los repositorios cumplan con los requisitos de mandatos de acceso abierto establecidos por financiadores nacionales e internacionales (UNESCO, 2021). Estrategias como la correcta implementación de metadatos normalizados, el uso de identificadores persistentes –como DOIs y ORCID– y la interoperabilidad con otros sistemas a través de protocolos como OAI-PMH (Open Archive Initiative-Protocol for Metadata Harvesting) son claves para lograr una mayor exposición (OpenAIRE, 2020).

Por lo tanto, trabajar activamente en la optimización de la visibilidad de los repositorios no es solo una cuestión técnica (Reyes-Lillo et al., 2025), sino una estrategia institucional para asegurar que la producción intelectual cumpla su objetivo final: ser consultada, utilizada y citada por comunidades académicas y la sociedad en general.

A continuación, analizamos cinco técnicas de optimización que pueden ser utilizadas para mejorar la visibilidad del contenido de un repositorio digital:

1. Optimización de calidad de metadatos.
2. Habilitación de los protocolos de interoperabilidad.
3. Adopción de identificadores persistentes.
4. Optimización SEO del repositorio.
5. *Generative Engine Optimization* en repositorios.

2. Optimización de calidad de metadatos

La normalización de metadatos es crucial para garantizar la interoperabilidad, la visibilidad, el acceso y la reutilización de los contenidos depositados. Una estrategia de metadatos debe considerar tanto los aspectos técnicos como los organizativos y normativos del repositorio.

Si bien el modelo de metadatos puede estar estrechamente ligado a la plataforma en la que se sostiene el repositorio, es fundamental que este esquema se adapte a las necesidades locales sin romper la compatibilidad con los estándares internacionales.

La tabla 1 muestra una comparativa entre algunos de los softwares de repositorios digitales y su esquema de metadatos de base:

Tabla 1
Softwares de repositorios digitales y su esquema de metadatos de base

Software	Esquema de metadatos de base	Información adicional
DSpace	Qualified Dublin Core https://github.com/DSpace/DSpace/blob/main/dspace/config/registries/dublin-core-types.xml	Usa un perfil propio de Dublin Core calificado; permite extensiones como METS/MODS. Desde DSpace 7, hay soporte para esquemas múltiples.
Fedora	No posee esquema fijo https://wiki.lyrasis.org/display/FEDORA6x/Data+Modeling	Usa modelos RDF/Linked Data; el esquema depende de la implementación (MODS, DC, PREMIS, etc.).
EPrints	Esquema de metadatos configurable y extensible https://wiki.eprints.org/w/Metadata#Metadata_Field_Types	Esquema basado en campos definidos por el administrador del repositorio. Puede interoperar con otros esquemas mediante exportaciones e importaciones.
InvenioRDM	JSON es una estructura alineada con DataCite https://inveniordm.docs.cern.ch/reference/metadata	El esquema se ajusta a DataCite's Metadata Schema v4.x con menores añadidos y modificaciones.
TIND	MARC21 https://www.tind.io	Utiliza como base MARC21 ya que está basado en INVENIO, desarrollado por el CERN. Sin embargo, tiene diversos módulos y se puede adaptar a otros formatos.
Digital Commons	No especifica esquema, pero puede ser mapeado a Dublin Core dimple o <i>qualified</i> https://digitalcommons.elsevier.com/en_US/organization-content-planning/metadata-options-in-digital-commons	Declara un esquema de metadatos flexible.
Dataverse	Utiliza diferentes metadatos compatibles con estándares https://guides.dataverse.org/en/latest/user/appendix.html	Garantiza la interoperabilidad y preservación de los metadatos mediante el uso de esquemas compatibles con estándares (como DDI, DataCite, Dublin Core, ISA-Tab y VOResource), lo que permite una fácil asignación y exportación en formatos estructurados.

Para iniciar una estrategia de optimización de metadatos, se recomienda iniciar con un diagnóstico inicial que incorpore al menos la evaluación de los siguientes elementos:

- Auditoría de metadatos existentes: el objetivo es revisar una muestra representativa de registros para detectar inconsistencias, errores de formato, campos vacíos o mal utilizados.
- Identificación de esquemas de metadatos utilizados: es esencial identificar el esquema de metadatos de base, pero también los diversos mapeos y posibilidades de exportación a otros esquemas.
- Revisión de vocabularios y autoridades: revisar y validar la implementación de vocabularios o listas controladas, como tesauros, autoridades e identificadores, así como listas de tipo de ítems o control de formato en los campos de lenguas y fechas.

Una vez realizado el diagnóstico, se recomiendan cuatro procesos fundamentales para optimizar los metadatos del repositorio, que se describen a continuación.

2.1. Limpieza y depuración de datos

Esta etapa resulta fundamental para optimizar la visibilidad de los contenidos almacenados en un repositorio. Los metadatos de alta calidad “should allow digital users to intuitively conduct the tasks such as identifying, describing, managing and searching data” (Ma et al., 2009, p. 1).

En ese sentido, la *data cleaning* (limpieza de datos) es un proceso de detección y corrección de errores, inconsistencias y campos incompletos orientado a mejorar la interoperabilidad y la experiencia del usuario (van Hooland y Verborgh, 2015; Westbrook et al., 2012). De esta forma, se fortalece la precisión y se mejora la recuperación de información a través de las herramientas de búsqueda de un sistema.

Para mejorar la calidad de los metadatos en repositorios, existen diversas herramientas, pero destaca ampliamente el uso de OpenRefine (<https://openrefine.org>). Esta es una herramienta de código abierto y gran potencia para limpiar, transformar y reconciliar datos tabulares desordenados. OpenRefine es especialmente útil para estandarizar y enriquecer metadatos en bibliotecas, archivos y conjuntos de datos de investigación.

Para utilizar OpenRefine en el proceso de *data cleaning* es fundamental: 1) exportar los metadatos en un formato compatible con OpenRefine, como CSV, TSV, Excel o JSON; y 2) que cada fila representa un recurso –por ejemplo, un documento o imagen–, y cada columna un elemento del esquema de metadatos. Si usáramos Dublin Core (DC), el ejemplo sería *dc:title*, *dc:creator*, *dc:subject*, entre otros.

De esta manera, es posible hacer una limpieza en lotes de registros que tengan inconsistencias, errores de tipeo o errores de formato. A continuación, se expone a modo de ejemplo una tabla de registros DC que potencialmente pueden ser mejorados con herramientas como OpenRefine (figura 1).

Figura 1

Ejemplo de registros DC con potencial de optimización a través de data cleaning.

dc:title	dc:creator	dc:subject	dc:date	dc:language
History of Science	Laura Gonzalez	Science; history	2020-06-10	eng
history of science	Laura gonzález	science; history	10/06/2020	en
Artificial Intelligence	María Gómez	Technology, AI; Machine Learning	2023	english
Big Data		Massive data; analytics	2021-03-14	en
Machine Learning	Carlos Ruiz	machine learning, artificial intelligence	2024-11-25	en
Cybersecurity	Ana Lopez	information security	2021/05/03	English

Cabe señalar que este proceso se puede efectuar sin importar el esquema de metadatos de base que tenga el sistema. Para sistematizar los metadatos en un formato aceptado por OpenRefine, es posible utilizar herramientas anexas como MarcEdit (<https://marcedit.reeset.net>), si el esquema de base es MARC21.

Una vez importados los registros, OpenRefine puede corregir inconsistencias como espacios innecesarios, mayúsculas desiguales, duplicados y formatos incorrectos –especialmente en fechas–, así como normalizar términos mediante vocabularios controlados. Se pueden dividir celdas, agrupar valores similares y aplicar transformaciones con expresiones GREL¹. Además, OpenRefine permite detectar valores faltantes, validar formatos y reconciliar datos con fuentes externas. Finalmente, los datos limpios se exportan en el formato deseado, listos para su reutilización o reintegración en el repositorio.

2.2. Implementación de vocabularios controlados

Para una mejora de metadatos con el objetivo de la optimización de visibilidad e interoperabilidad con otros sistemas, los vocabularios controlados juegan un rol relevante (Chipangila et al., 2024). Por una parte, mejoran la experiencia de búsqueda en el propio repositorio, controlando el vocabulario en campos clave como autores, idioma y palabras clave. Por otro lado, optimizan la integración con sistemas externos, a través del uso de identificadores, vocabularios o listas controladas, y formatos preestablecidos en campos clave para la interoperabilidad.

Para incorporar vocabularios controlados en repositorios digitales, el primer paso es identificar los campos de metadatos que se podrían beneficiar de la normalización. Por ejemplo, autores, fechas, idioma, palabras clave o temas, y tipo de recurso.

Posteriormente, es necesario seleccionar los vocabularios adecuados para controlar determinados campos. A continuación, listamos algunas opciones según categoría:

- Para controlar materias o palabras clave: algunas alternativas son las Library of Congress Subject Headings (LCSH), el tesauro de la UNESCO o el tesauro de Naciones Unidas (UNBIS). Luego de decidir el vocabulario que se quiere integrar, es necesario obtener el tesauro en un formato SKOS (Simple Knowledge Organization System), RDF (Resource Description Framework) o similar, e integrarlo a la plataforma. Dicha integración va a depender estrictamente de la plataforma y debe asegurar que los campos

1 GREL es un lenguaje de programación simple diseñado para facilitar la organización, transformación y consulta de datos en OpenRefine.

de materias o descriptores estén mapeados a los términos del tesauro. Por ejemplo, el campo dc.subject en Dublin Core o 650 en MARC21. Finalmente, es necesario verificar que los términos se autocompleten correctamente y que estos estén enlazados a sus URLs.

- Para controlar autoridades: es posible utilizar vocabularios controlados como VIAF (Virtual International Authority File) o identificadores persistentes como ORCID o ROR (Research Organization Registry). El procedimiento de integración de estos elementos va a depender también estrictamente de la plataforma. De igual manera, se debe asegurar que los campos del esquema de metadatos relacionados con autoridades estén adecuadamente mapeados al vocabulario o lista que se quiera integrar.
- Para controlar tipo de ítem: existen diversos vocabularios para controlar el tipo de ítem y fortalecer la interoperabilidad, por ejemplo, con sistemas como gestores de referencias. Entre los más destacados, podemos mencionar DCIM Type Vocabulary (Dublin Core), COAR Resource Type Vocabulary, MODS Resource Types, Schema.org / CreativeWork Types, OpenAIRE Guidelines Types, entre otros. Luego de seleccionar el vocabulario, se debe mapear los valores locales existentes al vocabulario elegido, ajustar formularios de ingreso de datos para que utilicen listas controladas y asegurar que los tipos sean consistentes en los registros. Además, se adapta la exportación de metadatos –por ejemplo, en Dublin Core o XML– para cumplir con estándares de interoperabilidad y facilitar la cosecha (harvest) por agregadores. El proceso puede incluir validación de datos, ajustes en la interfaz de usuario y pruebas para garantizar que los valores se reflejan correctamente tanto en la interfaz interna de administración (backend) como en la visualización pública del repositorio.
- Para controlar campos clave: la normalización de campos clave como el idioma o la fecha de publicación es fundamental para estandarizar estos metadatos. Una adecuada normalización favorece la experiencia de búsqueda en el propio repositorio y garantiza un adecuado intercambio de datos con servicios externos. En estos campos es recomendable adoptar la norma ISO-8601 para normalizar fechas y la norma ISO 639 para normalizar idiomas. Para lograr esta incorporación, se debe usar el formato YYYY-MM-DD para fechas y códigos de dos o tres letras para idiomas. Esto se obtiene mediante listas controladas, validación automática en formularios de ingreso y ajustes en las plantillas de metadatos del sistema –por ejemplo, en Dublin Core, MARC21 o MODS (Metadata Object Description Schema)–. Así se garantiza consistencia, interoperabilidad y compatibilidad con cosechadores externos.

2.3. Auditorías periódicas y ajuste según cambios de normativas

Para realizar auditorías periódicas de calidad de metadatos, se recomienda establecer una rutina regular, por ejemplo, trimestral o semestral, en la que se revisen muestras representativas de registros. Durante esta revisión, se debe verificar la coherencia, completitud, la utilización correcta de vocabularios controlados, formatos normalizados –como ISO 8601 e ISO 639-1– y la ausencia de errores tipográficos o duplicados. El uso de herramientas de validación automática y generación de informes de calidad ayuda a detectar problemas de forma eficiente.

Frente a cambios en normativas o estándares, es fundamental mantener un monitoreo activo de actualizaciones en esquemas como Dublin Core, COAR o OpenAIRE. Ante una modificación, se debe revisar el mapeo de metadatos en el repositorio, ajustar formularios, vocabu-

larios o plantillas de exportación, y capacitar al personal sobre los cambios. Documentar cada ajuste garantiza la trazabilidad y facilita las auditorías futuras.

3. Habilidad de los protocolos de interoperabilidad

Habilitar los protocolos de interoperabilidad en un repositorio digital es fundamental para aumentar su visibilidad y alcance. Estos protocolos permiten que servicios externos, como agregadores, motores de búsqueda académicos y portales nacionales o regionales, puedan cosechar automáticamente los metadatos y contenidos del repositorio, sin necesidad de intervención manual (Eells et al., 2024).

Gracias a esta interoperabilidad, los registros del repositorio pueden integrarse en plataformas como OpenAIRE, BASE, CORE, WorldCat, o incluso repositorios nacionales o regionales, lo que mejora su descubribilidad por parte de investigadores, estudiantes y el público en general. Además, facilita la integración con otros sistemas institucionales y garantiza que los contenidos cumplan con estándares abiertos y políticas de acceso abierto.

Existen diversas maneras de promover esta interoperabilidad, a través de diversos protocolos que enlistamos a continuación:

- Protocolo OAI-PMH: Es un protocolo basado en HTTP y XML que permite a otros sistemas cosechar metadatos de un repositorio de forma automática. Es ampliamente usado para que agregadores como OpenAIRE o BASE recojan información sobre los recursos disponibles.
- REST-API: Es una interfaz que permite a otros sistemas consultar, crear o modificar recursos en un repositorio mediante peticiones HTTP, como GET, POST, PUT, o DELETE. Es muy flexible y usada para integraciones con sistemas externos o aplicaciones personalizadas.
- SWORD: es un protocolo que permite depositar contenidos, como artículos, datasets o tesis, de manera remota en un repositorio. Facilita integraciones entre plataformas de publicación, sistemas institucionales y repositorios digitales.

La habilitación de los diversos protocolos para interoperar con otros sistemas va a depender estrictamente del software y la versión que se esté utilizando. Se deben utilizar las herramientas disponibles en cada repositorio para activar la API o SWORD con el objetivo de permitir la carga o actualización desde clientes externos.

Asimismo, cada repositorio permite la configuración de su protocolo OAI-PMH, donde se define la URL de base del servicio y se exponen los metadatos en un formato adecuado. En este ámbito, los metadatos son fundamentales y deben ser expuestos correctamente en formatos legibles para el protocolo como Dublin Core, MARC o DataCite.

La tabla 2 explica cómo activar el protocolo OAI-PMH según cada software:

Tabla 2

Activación de protocolo OAI según cada software.

Software	¿Cómo activar OAI-PMH?
DSpace https://wiki.lyrasis.org/display/DSDOC8x/OAI	Si bien OAI-PMH viene activado por defecto, es necesario verificar que en "local.cfg" o "dspace.cfg" estén "oai.enabled=true" y "oai.path=oai".
EPrints https://wiki.eprints.org/w/OAI	El soporte OAI viene activado por defecto. Es necesario configurar la URL de base ("oai.base_url"), "archive_id", "sets" y cualquier <i>stylesheet</i> XSL en "cfg/cfg.d". Es necesario mirar particularmente el archivo <i>oai.pl</i> .
Fedora https://github.com/saw-leipzig/foaipmh	Fedora 6 no incluye OAI-PMH nativo. Requiere implementar un endpoint externo (ej. Django + foaipmh) conectado con su REST API para operar.
InvenioRDM https://inveniordm.docs.cern.ch/reference/oai_pmh/	OAI-PMH activado por defecto en "/oai2d". Desde la interfaz de administración puedes definir sets y formatos ("oai_dc", "oai_datacite").
TIND	Desde el panel de administración, ir a "OAI Repository Admin", activar el proveedor y definir sets.
Digital Commons https://digitalcommons.elsevier.com/integration-preservation/digital-commons-and-oai-pmh	Soporte OAI habilitado por defecto. Se configuran los campos que se exponen a través del administrador de metadatos.
Dataverse https://guides.dataverse.org/en/latest/admin/harvestserver.html	Desde "Harvesting Server" en el Dashboard, se habilita el servicio OAI-PMH y se definen sets. El endpoint suele ser "/oai".

Para garantizar un correcto funcionamiento del protocolo OAI-PMH existen herramientas conocidas como validadores, que permiten evaluar la operatividad del protocolo. Una de las herramientas más conocidas (figura2) es el OAI-PMH validator (<https://validator.oaipmh.com>). Este validador permite verificar que los conjuntos (sets) estén bien definidos y que los registros exporten los campos correctamente: fechas, identificadores, tipos de ítem, idioma, etc. Asimismo, permite revisar las respuestas XML para detectar errores o formatos mal estructurados. Para analizar el funcionamiento del protocolo de un repositorio, solo se necesita ingresar la URL de base y el sistema realiza el análisis.

Figura 2

OAI-PMH Validator.

Para sostener un correcto funcionamiento del protocolo es necesario realizar actividades de mantenimiento y monitorear periódicamente el estado del protocolo. Asimismo, es importante adoptar los mapeos de metadatos si cambian las recomendaciones de COAR, OpenAIRE u otros agregadores. Asimismo, para el seguimiento, es relevante documentar los *end-points*² y mantener coordinación con otros sistemas que dependan del repositorio.

4. Adopción de identificadores persistentes

Un identificador persistente es una referencia digital única, duradera y resoluble a un objeto concreto, por ejemplo, un artículo, *dataset*, software, persona u organización. Estos identificadores están diseñados para mantener su validez y accesibilidad a lo largo del tiempo, incluso si cambia su ubicación física o su servidor (Meadows et al., 2019).

Normalmente, un identificador persistente tiene tres componentes esenciales: 1) unicidad global, lo que implica que incluye una sintaxis controlada y espacio de nombres gobernados por autoridades claramente definidas; 2) persistencia, lo que garantiza enlaces y funciones de resolución estables, así como esquemas y objetos referidos persistentes, y 3) resolubilidad, lo que implica que son accesibles tanto para humanos como para máquinas, proporcionando información sobre cómo encontrar, acceder o usar el objeto referido (de Castro et al., 2023).

Los identificadores persistentes (tabla 3) son relevantes para optimizar la visibilidad de las publicaciones y la citabilidad, ya que facilitan que buscadores, repositorios académicos y herramientas de análisis encuentren y enlacen automáticamente los documentos, sin depender de URLs fluctuantes. Asimismo, entregan estabilidad a largo plazo, promoviendo la accesibilidad de recursos, dado que combaten el "*link rot*"³ y el "*content drift*"⁴: incluso si el objeto se mueve, su identificador persistente seguirá resolviendo correctamente. Finalmente, si los identificadores persistentes están asociados a metadatos estructurados y respaldados por infraestructuras robustas, mejoran la confianza y reputación del contenido.

Entre los identificadores persistentes más reconocidos, encontramos el Digital Object Identifier (DOI). Es el esquema más empleado para artículos, libros, *datasets* y software. Este sistema combina un identificador permanente con metadatos obligatorios y resolución garantizada. Al resolver un DOI, este conduce a una *landing page* con metadatos, promoviendo la visibilidad y el seguimiento de citaciones. Los DOIs son asignados por agencias registradores como CrossRef y DataCite y, en líneas generales, resulta más costoso que otros identificadores.

Por otro lado, también existe el sistema Handle, que es un identificador no-comercial usado desde 1995. Su principal objetivo es proveer servicios de identificación persistente y resolución, operados de manera centralizada por la Corporation for National Research Initiatives (CNRI). Un identificador Handle consiste en un prefijo que identifica la autoridad, en conjunto con un sufijo, que se refiere al objeto al cual hace referencia. Handle es la base técnica de DOI y resulta más económico en su implementación que DOI. Algunos sistemas como DSpace

2 Un *endpoint* es una dirección o URL específica que permite a un sistema externo interactuar con el repositorio y acceder a sus servicios o datos.

3 *Link rot* ocurre cuando un hipervínculo ya no conduce al contenido previsto porque la página ha sido movida, eliminada o el dominio ya no está activo.

4 *Content drift* ocurre cuando el contenido en una URL determinada cambia con el tiempo, por lo que ya no refleja lo que originalmente se citó o se pretendía, aunque el enlace aún funcione.

integran el sistema Handle por defecto, pero necesita ser adquirido y configurado para que resuelva correctamente los recursos a través del identificador (<https://wiki.lyrasis.org/display/DSDOC8x/Handle.Net+Registry+Support>).

Otra alternativa es el identificador persistente ARK (Archival Resource Key), un sistema diseñado para proporcionar enlaces duraderos y confiables a objetos digitales, particularmente útil en bibliotecas, archivos y museos. A diferencia de otros identificadores como el DOI o Handle, ARK es más barato, descentralizado y altamente flexible (<https://arks.org>), lo que permite a las instituciones generar y gestionar sus propios identificadores sin depender de una entidad registradora central. Su formato típico es ark:/NAAN/identificador, donde el NAAN identifica a la organización emisora. Una característica distintiva de ARK es su capacidad de ofrecer acceso no solo al objeto digital, sino también a sus metadatos y a una declaración de compromiso de persistencia, lo que refuerza la transparencia y la confianza en la conservación a largo plazo. Este sistema se ha adoptado ampliamente en instituciones como la Biblioteca del Congreso de EE. UU. y la California Digital Library, favoreciendo la visibilidad y trazabilidad de recursos culturales y académicos.

Tabla 3
Comparación entre identificadores persistentes.

	DOI	Handle	ARK
Gestión	Centralizada (DataCite, Crossref)	Distribuida (CNRI)	Descentralizada (institucional)
Estructura	10.1234/abc123	20.5000/xyz456	ark:/12345/x6789
Resolución	Sí, mediante https://doi.org/	Sí, mediante https://hdl.handle.net/	Sí, mediante https://n2t.net/ o local
Acceso a metadatos	Sí (<i>landing page</i> obligatoria)	Sí (dependiendo de uso)	Sí (mediante inflexión, es decir, un carácter modificador de búsqueda)
Persistencia garantizada	Alta (por contrato)	Alta (depende de repositorio)	Variable (según política institucional)
Coste	Requiere membresía de pago y puede incluir coste adicional por cada asignación de DOIs	Cuota anual de 50 USD	No requiere pago, ni membresía con costo asociado.
Usos típicos	Artículos, <i>datasets</i> , software	Objetos en repositorios (DSpace, Fedora)	Archivos, bibliotecas, museos digitales

Para implementar un identificador persistente, lo primero es seleccionar cuál se va a implementar. Cabe señalar que los identificadores no son excluyentes entre sí y se pueden combinar. Por ejemplo, añadir un DOI y un ARK a un recurso.

A continuación, es necesario registrarse con un proveedor de identificadores persistentes. Para obtener DOIs, se puede hacer a través de una de las agencias registradoras, como Crossref o DataCite, aunque ambos requieren una membresía de pago y puede conllevar un costo asociado por acuñar cada DOI.

Para implementar el sistema Handle, se puede adquirir con la CNRI, pagando alrededor de 50 USD y vinculándolo con un sistema de repositorios compatible (<https://www.handle.net/payment.html>). Una vez configurado, cada objeto depositado recibe un identificador con un prefijo autorizado, asignado por CNRI, y un sufijo único. Estos identificadores se resuelven a través de <https://hdl.handle.net/>, lo que garantiza su accesibilidad a largo plazo, incluso si el recurso cambia de ubicación física.

Para implementar ARK, es necesario solicitar un Name Assigning Authority Number (NAAN) en [arks.org](https://arks.org/about/getting-started-implementing-arks) (<https://arks.org/about/getting-started-implementing-arks>). Un NAAN es un número único que identifica a la institución emisora de ARKs dentro del sistema. Funciona como un prefijo oficial que garantiza que cada organización que crea identificadores ARK tenga un espacio de nombres propio y exclusivo.

Posteriormente, es necesario configurar el repositorio para emitir y mantener identificadores persistentes. Esto va a depender de cada software que se utilice para la gestión del repositorio institucional, y se puede realizar mediante las agencias registradoras, el uso de *plugins*, la integración en el propio sistema o el manejo de archivos de configuración clave en el sistema.

Finalmente, es necesario integrar al flujo de trabajo de la ingesta de nuevos recursos la asignación de identificadores persistentes en el campo "*identifier*", junto con asegurar que el identificador se resuelva adecuadamente en una *landing page* del objeto con sus metadatos y acceso al recurso.

De esta manera, los identificadores persistentes como DOI, Handle y ARK son herramientas fundamentales para fortalecer la visibilidad de los documentos en un repositorio digital. Al proporcionar enlaces estables, únicos y duraderos, permiten que los recursos sean fácilmente localizables, accesibles y citables, incluso cuando su ubicación técnica cambia con el tiempo. Además, al estar integrados en infraestructuras globales de resolución y asociados a metadatos estructurados, estos identificadores facilitan el descubrimiento por motores de búsqueda, recolectores, sistemas de citación académica y redes de datos abiertos. En síntesis, estos identificadores garantizan que los documentos no solo sean preservados, sino también difundidos y reconocidos en los entornos digitales actuales.

5. Optimización SEO del repositorio

La optimización para motores de búsqueda o SEO (Search Engine Optimization) en repositorios digitales es una estrategia clave para aumentar la visibilidad, el acceso y el impacto de los contenidos académicos, científicos y culturales que albergan. A pesar de contar con metadatos estructurados y estándares de preservación, muchos repositorios no logran posicionarse adecuadamente en buscadores como Google o Bing, lo que limita que los usuarios descubran sus recursos de forma orgánica.

Implementar buenas prácticas SEO, como el uso correcto de etiquetas HTML, la exposición de metadatos Dublin Core en formato schema.org, la generación de URLs amigables, el uso de *sitemaps* y la indexación automática de nuevos contenidos, permite que los documentos sean mejor comprendidos, clasificados y priorizados por los motores de búsqueda.

Además, la combinación de SEO con identificadores persistentes (como DOI o Handle) refuerza la estabilidad y la trazabilidad de los recursos en la web.

En un entorno digital donde la atención es limitada y la competencia por visibilidad es alta, optimizar el SEO de un repositorio no es solo una mejora técnica, sino una acción estratégica para garantizar que los recursos depositados cumplan su misión de ser encontrado, utilizado y citado.

Para optimizar el SEO en un repositorio, se recomienda tener en consideración los siguientes elementos:

5.1. Uso correcto de etiquetado semántico HTML

Utilizar etiquetas semánticas como <title>, <meta name="description">, <meta name="citation_doi">, <meta name="citation_author">, <h1>, <h2>, <article>, <section>, entre otras, para ayudar a los motores de búsqueda a comprender la estructura del contenido. Todo esto incluso ayuda a que herramientas como Altmetric puedan rastrear de mejor manera las métricas de un recurso en particular (<https://help.altmetric.com/support/solutions/articles/6000240582-required-metadata-for-content-tracking>) (Reyes-Lillo y Pastor-Ramon, 2024).

Hay que asegurar que cada página de recurso, como un documento, tenga un <title> único y descriptivo. Adicionalmente, es recomendable incluir metadatos enriquecidos usando schema.org o Dublin Core embebido en etiquetas <meta> o en JSON-LD.

A continuación, es posible distinguir un uso adecuado del etiquetado semántico en Dublin Core (figura 3) o en JSON-LD (figura 4):

Figura 3

Ejemplo de Dublin Core embebido en <meta>.

```
<head>
  <meta name="DC.title" content="How to improve your metadata" />
  <meta name="DC.creator" content="Laura Martinez" />
  <meta name="DC.date" content="2023-11-15" />
  <meta name="DC.identifier" content="http://doi.org/10.1234/test-doi.2025.001" />
  <meta name="DC.language" content="es" />
  <meta name="DC.type" content="Book" />
</head>
```

Figura 4

Ejemplo de JSON-LD.

```
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@type": "Book",
  "name": "How to improve your metadata",
  "author": {
    "@type": "Person",
    "name": "Laura Martinez"
  },
  "datePublished": "2023-11-15",
  "identifier": {
    "@type": "PropertyValue",
    "propertyID": "Handle",
    "value": "http://doi.org/10.1234/test-doi.2025.001"
  },
  "inLanguage": "es",
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "publisher": {
    "@type": "Organization",
    "name": "Universidad Nacional"
  }
}
</script>
```

5.2. Crear y mantener un *sitemap* XML

La creación y mantención de un *sitemap* XML en un repositorio digital es una práctica esencial para mejorar la indexación y visibilidad de los contenidos por parte de los motores de búsqueda como Google o Bing. Un *sitemap* actúa como un mapa estructurado que lista todas las páginas relevantes del repositorio, permitiendo a los buscadores descubrir rápidamente nuevos documentos, actualizaciones o recursos depositados.

Para que funcione correctamente, es fundamental que el *sitemap* incluya únicamente URLs públicas, permanentes y accesibles (como las que contienen identificadores persistentes tipo DOI o Handle), y que se actualice automáticamente cada vez que se agrega o modifica contenido. Además, debe estar correctamente referenciado en el archivo robots.txt y ser enviado a herramientas como Google Search Console para maximizar su eficacia.

Una buena implementación del *sitemap* no solo agiliza la indexación, sino que también mejora el posicionamiento SEO de los recursos, aumentando su alcance y uso dentro del ecosistema académico y científico.

5.3. Configuración adecuada del archivo robots.txt

El archivo robots.txt desempeña un papel clave en la optimización SEO de repositorios digitales, ya que controla cómo los motores de búsqueda acceden, rastrean e indexan sus contenidos. Este archivo, ubicado en la raíz del sitio web, indica a los robots, como Googlebot o Bingbot, qué partes del repositorio pueden ser exploradas y cuáles deben ser excluidas.

Una configuración adecuada permite que los buscadores accedan a las páginas de los objetos digitales, como las *landing pages* con identificadores persistentes. Por otra parte, se puede bloquear rutas irrelevantes o sensibles, como áreas administrativas o filtros de navegación que podrían generar contenido duplicado. Por esta razón, es indispensable que el archivo no bloquee rutas relevantes vinculadas a identificadores persistentes, como por ejemplo, /handle/ o /ark:/.

Un ejemplo simple (figura 5) es el siguiente:

Figura 5

Ejemplo de elementos a considerar en el archivo robots.txt.

```
User-agent: *  
Allow: /handle/  
Disallow: /admin/  
Disallow: /private/  
Sitemap: https://repository.org/sitemap.xml
```

Si no se configura correctamente, el robots.txt puede impedir accidentalmente la indexación de recursos importantes, afectando negativamente su visibilidad y descubrimiento en los resultados de búsqueda. Además, debe incluir una referencia al archivo sitemap.xml, facilitando a los buscadores el rastreo estructurado del contenido.

5.4. Aplicar SSR (Server-Side Rendering) o renderizado híbrido

Aplicar SSR (Server-Side Rendering) o un enfoque de renderizado híbrido en un repositorio digital es una estrategia técnica que contribuye a mejorar la visibilidad y rendimiento del sitio, especialmente en un contexto web cada vez más orientado a la inteligencia artificial y la indexación automatizada.

A diferencia del CSR (Client-Side Rendering) o renderizado puramente en el cliente, donde el contenido se genera dinámicamente en el navegador, el SSR permite que las páginas se generen en el servidor antes de ser enviadas al usuario o al robot de búsqueda. Esto tiene múltiples beneficios: en primer lugar, mejora el SEO, ya que los motores de búsqueda pueden acceder inmediatamente al contenido estructurado sin depender de JavaScript para renderizarlo. En segundo lugar, acelera los tiempos de carga inicial, lo que mejora la experiencia del usuario y favorece la navegación desde dispositivos móviles o redes lentas.

Además, al implementar un renderizado híbrido (SSR combinado con CSR), se logra un equilibrio óptimo entre rendimiento, interactividad y visibilidad, permitiendo que el repositorio sea eficaz tanto para humanos como para *bots* de indexación y motores generativos.

Existen herramientas como Next.js, Nuxt o Rendertron, que permiten adaptar sitios a un SEO amigable y orientado a estos renderizados.

6. Generative Engine Optimization en repositorios: un factor a considerar

Generative Engine Optimization (GEO) es un concepto emergente que se refiere a la optimización de contenidos digitales para motores de búsqueda generativos, como ChatGPT, Google Search Generative Experience (SGE) o Perplexity, que usan inteligencia artificial (IA) para responder con texto generado directamente, en lugar de solo mostrar enlaces como lo hace el SEO tradicional (Daniels, 2025).

GEO busca adaptar la forma en que los contenidos digitales están estructurados y etiquetados para que modelos de lenguajes (LLMs) los interpreten correctamente, los referencien en sus respuestas y sean integrados en el contenido generado por inteligencia artificial (Aggarwal et al., 2024).

De esta forma, así como el SEO optimiza los contenidos para que sean más visibles en Google, GEO optimiza los contenidos para ser comprendidos, citados y usados por los motores de generación de respuestas sobre la base de la IA.

En este ámbito, los repositorios también pueden tomar acciones específicas para optimizar su contenido basadas en estrategias GEO. Por ejemplo, estructurar los metadatos utilizando schema.org, JSON-LD, Open Graph, entre otras tecnologías es altamente recomendable para contribuir a la incorporación de su contenido en respuestas generadas por herramientas de IA.

Por otro lado, es necesario proporcionar contenido claro y accesible, evitando el texto oculto o técnico difícil de procesar para modelos LLMs. Asimismo, es aconsejable indicar claramente licencias de propiedad intelectual flexibles como Creative Commons para facilitar la reutilización de contenido.

Finalmente indicar que también son factores clave para mejorar el contenido para ser procesado por LLMs: el uso de *landing pages* descriptivas, la incorporación de identificadores persistentes para garantizar trazabilidad, proporcionar ficheros RDF o JSON y facilitar el acceso a la API.

En síntesis, GEO representa un nuevo paradigma de visibilidad para los repositorios digitales, donde ya no basta con aparecer en Google, sino que el contenido debe ser estructurado, accesible y comprensible por modelos de lenguaje. Implementar estrategias GEO no solo aumenta el alcance de los recursos, sino que prepara al repositorio para integrarse en el ecosistema de búsqueda y descubrimiento basado en inteligencia artificial que está definiendo el futuro del acceso al conocimiento.

7. Conclusión

La optimización de visibilidad en repositorios digitales requiere un enfoque integral que combine aspectos técnicos, normativos y estratégicos. La calidad de los metadatos garantiza que los recursos sean comprensibles y reutilizables tanto por humanos como por máquinas; la interoperabilidad permite su integración fluida en redes globales de información; y los identificadores persistentes aseguran su trazabilidad y acceso duradero. Al mismo tiempo, una sólida optimización SEO mejora el posicionamiento en buscadores, mientras que incorporar una perspectiva GEO amplía el alcance de los contenidos hacia agentes de inteligencia artificial generativa. En conjunto, estos elementos fortalecen la visibilidad, el impacto y la circulación del conocimiento alojado en los repositorios, alineándolos con los principios de la ciencia abierta y el acceso equitativo a la información.

8. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

9. Referencias

- Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., & Deshpande, A. (2024). *GEO: Generative Engine Optimization* (arXiv:2311.09735). arXiv. <https://doi.org/10.48550/arXiv.2311.09735>
- Chipangila, B., Liswaniso, E., Mawila, A., Mwanza, P., Nawila, D., M'sendo, R., Nyirenda, M., & Phiri, L. (2024). Controlled vocabularies in digital libraries: Challenges and solutions for increased discoverability of digital objects. *International Journal on Digital Libraries*, 25(2), 139–155. <https://doi.org/10.1007/s00799-023-00374-1>
- Daniels, C. (2025). *From SEO to GEO: How agencies are navigating LLM-driven search*. Campaign Asia. <https://www.campaignasia.com/article/from-seo-to-geo-how-agencies-are-navigating-llm-driven-search/501593>
- De Castro, P., Herb, U., Rothfritz, L., & Schöpfel, J. (2023). *Building the plane as we fly it: The promise of Persistent Identifiers*. Zenodo. <https://doi.org/10.5281/zenodo.7258286>

- Eells, L., Kelly, J., & Farrell, S. (2024). Repository (R)evolution: Metadata, interoperability, and sustainability. *Journal of Librarianship and Scholarly Communication*, 12(1). <https://doi.org/10.31274/jlsc.16890>
- Ma, S., Lu, C., Lin, X., & Galloway, M. (2009). Evaluating the metadata quality of the IPL. *Proceedings of the American Society for Information Science and Technology*, 46(1), 1–17. <https://doi.org/10.1002/meet.2009.1450460249>
- Meadows, A., Haak, L. L., & Brown, J. (2019). Persistent identifiers: The building blocks of the research information infrastructure. *Insights*, 32(1). <https://doi.org/10.1629/uksg.457>
- OpenAIRE. (2020). OpenAIRE Guidelines—OpenAIRE Guidelines documentation. Retrieved from: <https://guidelines.openaire.eu/en/latest>
- Piowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6, e4375. <https://doi.org/10.7717/peerj.4375>
- Reyes-Lillo, D., Morales-Vargas, A. & Rovira, C. (2025). Visibility, discoverability, findability, Search Engine Optimization (SEO) and Academic SEO in digital repositories: A scoping review. *BiD: textos universitaris de biblioteconomia i documentació*, (54). <https://doi.org/10.1344/BID2025.54.06>
- Reyes-Lillo, D., & Pastor-Ramon, E. (2024). Use of handle in institutional repositories and its relationship with alternative metrics: A case study in Spanish-speaking America. *Hipertext. Net*, 29. <https://doi.org/10.31009/hipertext.net.2024.i29.17>
- Swan, A. (2010, Febrero). *The open access citation advantage: Studies and results to date*. <https://eprints.soton.ac.uk/268516/>
- UNESCO. (2021). *UNESCO recommendation on open science*. <https://unesdoc.unesco.org/ark:/48223/pf0000379949>
- Van Hooland, S., & Verborgh, R. (2015). *Linked data for libraries, archives and museums: How to clean, link and publish your metadata*. Facet. <https://doi.org/10.29085/9781783300389>
- Westbrook, R. N., Johnson, D., Carter, K., & Lockwood, A. (2012). Metadata clean sweep: A digital library audit project. *D-Lib magazine*, 18(5/6). <https://doi.org/10.1045/may2012-westbrook>

IA en salas de prensa iberoamericanas: casos de uso y mejores prácticas

Alexis Apablaza-Campos

Universidad UNIACC, Chile

<https://orcid.org/0000-0003-4407-1145>

Jaime Andrés Wilches Tinjacá

Institución Universitario Politécnico Grancolombiano, Colombia

<https://orcid.org/0000-0002-4425-9394>

Apablaza-Campos, A., & Wilches Tinjacá, J. A. (2025). IA en salas de prensa iberoamericanas: casos de uso y mejores prácticas. En J. Guallar, M. Vállez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 142-156). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicom.10.esp>

Resumen

Este capítulo examina la creciente integración de la inteligencia artificial (IA) en las rutinas periodísticas iberoamericanas, acelerada por la irrupción de modelos generativos. El objetivo es explorar los principales usos de la IA en la generación de contenidos y proponer buenas prácticas. Se identifican siete áreas clave de aplicación: 1) presentadores virtuales, 2) redacciones automatizadas, 3) uso de IA generativa, 4) conversión de audios, 5) herramientas de SEO asistido, 6) resúmenes de artículos y 7) chatbots personalizados. A través de numerosos ejemplos regionales, se demuestra que la IA opera como una herramienta de apoyo que agiliza procesos y potencia la creatividad, sin sustituir la labor humana. Las conclusiones resaltan la importancia de la ética, la transparencia, la capacitación continua, la adaptación a contextos locales y las colaboraciones estratégicas para una adopción equilibrada y responsable de la IA en el periodismo.

Palabras clave

Inteligencia artificial; Generación de contenidos; Inteligencia artificial generativa; Medios de comunicación; Iberoamérica.

I. Introducción

En los últimos años la inteligencia artificial (IA) ha dejado de ser una especie de ‘tecnología del futuro’ para convertirse en un elemento central en las rutinas periodísticas. La irrupción de modelos generativos como ChatGPT, el 30 de noviembre de 2022, marcó un punto de inflexión que aceleró la adopción de herramientas de IA en la industria de los medios (Apablaza-Campos y Codina, 2023).

Estudios recientes muestran que los editores de noticias utilizan la IA para automatizar contenidos, mejorar la distribución, crear productos con supervisión humana y ofrecer recomendaciones a sus audiencias (Segarra Saavedra et al., 2019; Apablaza-Campos, 2024; Newman y Cherubini, 2025). Esta rápida evolución plantea oportunidades, pero también desafíos estratégicos y éticos que los gestores de empresas periodísticas y los periodistas deben considerar.

El objetivo de este capítulo es sintetizar las principales formas en que los medios de comunicación de Iberoamérica están incorporando la IA en la generación de contenidos y, a partir de esos usos, proponer buenas prácticas y recomendaciones para responsables de medios, periodistas y personas interesadas en el futuro del periodismo.

Se trata de un texto breve y práctico, centrado en conclusiones y sugerencias más que en extensas discusiones teóricas. La estructura se basa en siete grandes áreas de aplicación identificadas –principalmente– en la segunda edición del libro *Inteligencia artificial para la generación de contenidos: experiencias editoriales en medios de comunicación de Iberoamérica* (Apablaza-Campos y Wilches Tinjacá, 2025); cada sección describe la categoría, presenta ejemplos y extrae lecciones útiles para la industria mediática regional.

2. Presentadores virtuales

Los presentadores virtuales o avatares generados por IA son modelos sintéticos que “presentan” noticias en vídeo o audio. Esta tecnología, basada en sistemas de síntesis de voz y animación facial, permite crear conductores digitales capaces de leer un guion, mantener un estilo coherente y ser adaptados a distintas lenguas. Para los medios, los presentadores virtuales pueden reducir costes en la producción de contenidos audiovisuales, ampliar la oferta lingüística y cubrir eventos de forma continua.

2.1. Ejemplos destacados por país

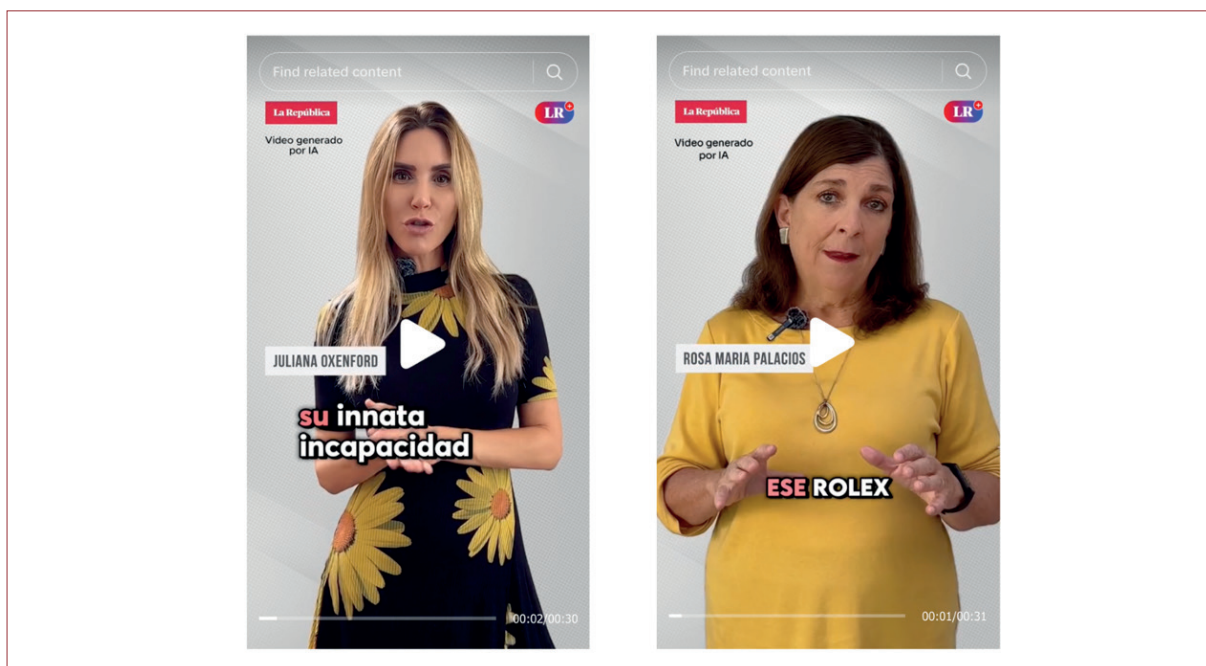
- **México:** NAT, de *Grupo Fórmula*, se convirtió en la primera presentadora IA de noticias del país (Kusunoki, 2024). Otra iniciativa es CLARA, un acrónimo de *Conductora Lógica de Asistencia y Respuesta Automática*, creada por medios universitarios como *Canal 44* y *Radio Universidad de Guadalajara* (Del Campo y Ramírez Santos, 2024).
- **Perú:** el canal universitario *Letras TV* lanzó Illariy, un avatar que presenta noticieros en quechua (Kusunoki, 2024). El diario *La República* utiliza avatares que reproducen las voces de sus columnistas para convertir artículos en videocolumnas publicadas en TikTok (Lozano Chávez, 2024).
- **Otros países:** en España, RTVE emplea a *Hiperia*, un avatar de contenidos musicales (Salaverría, 2024); mientras que medios televisivos como *Megacadena* en Paraguay o *Radio Televisión Dominicana* experimentan con presentadores para participar en su informativo principal e incluso narrar la lucha por la independencia (Lora & Álvarez Álvarez, 2024; Ferreira Candia et al., 2025). En Ecuador, *Teleamazonas* permitió a su audiencia escoger el nombre de un presentador IA para su sección de tecnología (Espinosa, 2024).

2.2. Buenas prácticas y recomendaciones

- **Valorar la diversidad lingüística y cultural:** los ejemplos de *Illariy* y *La República* muestran que los presentadores virtuales pueden acercar las noticias a comunidades quechuahablantes y a segmentos jóvenes de la audiencia. Adaptar los avatares a dialectos o lenguajes locales refuerza la identidad cultural y puede mejorar la relevancia del contenido.
- **Transparencia con la audiencia:** es recomendable que los medios identifiquen claramente cuándo están usando un presentador sintético y expliquen su propósito. La transparencia evita percepciones de engaño y ayuda a gestionar expectativas sobre la confiabilidad de la información.
- **Complementar, no sustituir:** aunque los avatares reducen costes, es importante que convivan con periodistas humanos. La supervisión humana garantiza la calidad, aporta criterio editorial y previene errores o sesgos automatizados (Ventura Pocino, 2021).

Figura 1

Videocolumnas en La República (Perú) generadas con IA.



(Lozano Chávez, 2024).

3. Redacciones con IA automatizada

Las redacciones con IA automatizada utilizan algoritmos para generar textos de manera autónoma a partir de datos estructurados. Son comunes en deportes, finanzas o resultados electorales, donde hay formatos repetitivos y abundancia de datos (Segarra-Saavedra et al., 2019). Estas herramientas permiten producir grandes volúmenes de notas con rapidez, liberando a periodistas para tareas de mayor valor añadido (Apablaza-Campos, 2024).

3.1. Ejemplos destacados según aplicación

- **Crónicas deportivas:** la empresa DataFactory provee servicios de redacción automática a medios de Chile (*El Rancagüino*), Perú (*Radio RPP*), Venezuela (*Meridiano*) y Colombia (*Noticias RCN*, *Gol Caracol*, *Revista Semana* y *Caracol Radio*), generando notas de partidos de fútbol de diferentes ligas en tiempo real (Apablaza-Campos y Wilches Tinjacá, 2025).
- **Cobertura electoral:** en Brasil, *TV Globo* y *G1* utilizaron IA para narrar los resultados de las elecciones presidenciales de 2022 (Barbosa y Costa Pinto, 2025); en España, *RTVE* aplicó sistemas automatizados para informar sobre votaciones locales en pueblos con una población menor a mil habitantes (Apablaza-Campos, 2025).
- **Actualidad y tecnología:** medios como *ABC Color*, en Paraguay; *El Comercio* y *Gadgeros*, en Ecuador, utilizan la automatización para noticias de actualidad y artículos sobre innovaciones tecnológicas (Ferreira Candia et al., 2025; Espinosa, 2024).

3.2. Buenas prácticas y recomendaciones

- **Priorizar la verificación de datos:** los algoritmos dependen de fuentes numéricas; por tanto, los medios deben asegurarse de que las fuentes de datos sean fiables y estén libres de errores. Un sistema de revisión humana o validación automatizada evitará la publicación de información incorrecta.
- **Personalización de la voz editorial:** aunque la redacción sea automática, los medios pueden ajustar el estilo, tono y estructuras de los textos para que mantengan la identidad de la marca. Esto refuerza la coherencia y genera confianza entre los lectores.
- **Uso estratégico:** la automatización debe servir para liberar recursos humanos y permitir que los periodistas se concentren en reportajes, análisis y piezas de investigación. Integrar la IA en las rutinas diarias puede aumentar la productividad sin sacrificar la calidad (Segarra-Saavedra et al., 2019; Murcia Verdú et al., 2022; Apablaza-Campos, 2024).

Figura 2

El Rancagüino (Chile) ofrece contenidos deportivos mediante IA automatizada en web y redes sociales con la debida advertencia a sus lectores.



(Erlandsen, 2024).

4. Experiencias con IA generativa

Las herramientas de IA generativa –como ChatGPT, Gemini u otras similares– permiten redactar borradores, sugerir ideas de guion y crear contenido multimedia. Son especialmente útiles para acelerar la investigación, redactar guiones y explorar enfoques novedosos. Utilizadas con criterio, estas herramientas pueden potenciar la creatividad periodística.

4.1. Ejemplos destacados por país

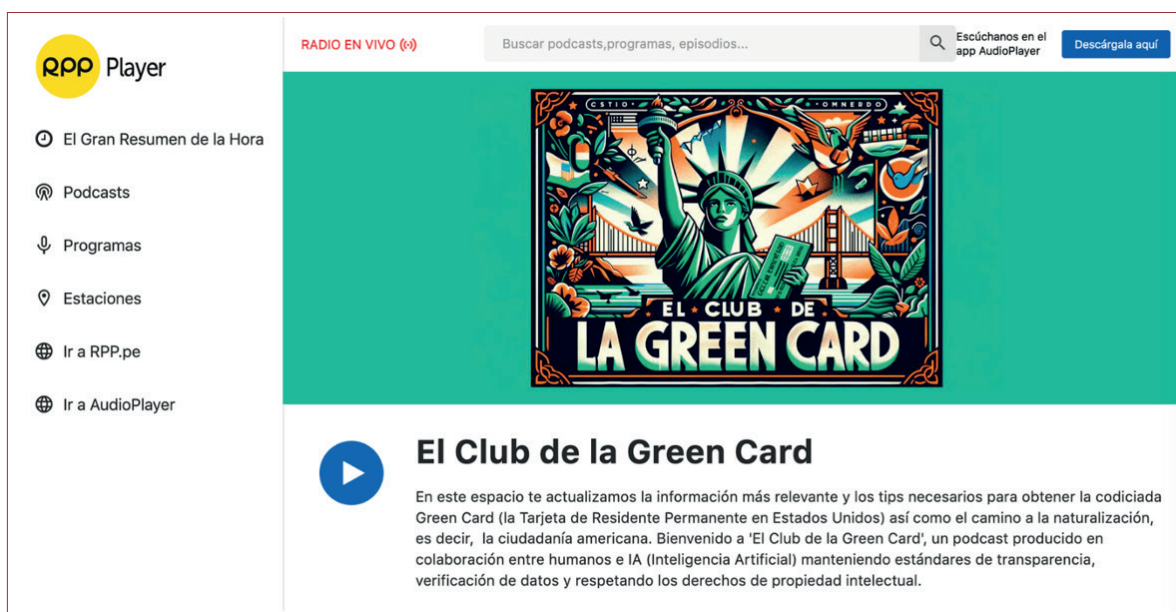
- **Bolivia:** los periódicos *Opinión* y *El Deber* utilizan ChatGPT y Gemini para búsqueda de información e investigación como parte del proceso de reporte (Banegas Flores, 2024).
- **Guatemala:** el verificador de datos *Ojoconmipisto* recurre a la IA para generar ideas de guiones para su podcast (Alpírez, 2024).
- **República Dominicana:** *Diario Libre* emplea estas herramientas para optimizar redacciones humanas (Lora y Álvarez Álvarez, 2024).
- **Perú:** *RPP Radio* genera podcasts asistidos por ChatGPT (Lozano Chávez, 2024).
- **Uruguay:** *Canal 4* produce textos e imágenes que se proyectan en pantalla (Roba, 2024).
- **Venezuela:** *2001online.com* emplea diversas herramientas IA para generación de vídeos que comparte en sus perfiles de redes sociales (De Los Reyes y Sarmiento Altuve, 2025).

4.2. Buenas prácticas y recomendaciones

- **Uso como asistente creativo:** la IA generativa no debe sustituir la labor de investigación y la creatividad humana, sino actuar como un asistente que propone ideas, corrige redacciones o traduce borradores. Los periodistas deben validar la información generada y aportar su criterio profesional.
- **Control del sesgo y la precisión:** los sistemas generativos pueden producir respuestas inexactas o sesgadas. Es aconsejable contrastar los resultados con fuentes confiables y mantener un protocolo editorial claro para determinar qué insumos de la IA se integran en el producto final.
- **Formación y alfabetización:** los equipos deben capacitarse en el uso de estas tecnologías, entender sus limitaciones y aprender a formular preguntas efectivas para obtener resultados útiles (Gutiérrez Caneda et al., 2023).

Figura 3

El podcast El Club de la Green Card, de RPP Radio (Perú), se elabora con apoyo de ChatGPT.



(Lozano Chávez, 2024).

5. Conversión de audios y síntesis de voz

Las herramientas de conversión de audio facilitan la transcripción de entrevistas o la generación de audio a partir de textos. Automatizan tareas tediosas, permiten la accesibilidad de contenidos y abren nuevas formas de consumo (podcasts automatizados o artículos que se leen solos). También incluyen sistemas de traducción y limpieza de audio.

5.1. Ejemplos destacados según aplicación

- **Transcripción de entrevistas:** medios como *El Deber* (Bolivia), *ABC Color* (Paraguay), *Canal 4* y *El Observador* (Uruguay) utilizan la IA para transcribir audios de entrevistas (Apablaza-Campos y Wilches Tinjacá, 2025).
- **Texttospeech:** en Colombia, el sistema se utiliza de forma masiva en medios como *El Espectador*, *El Tiempo*, *Blu Radio*, *Pulzo*, *El Heraldo de Barranquilla* y *El Colombiano* de Medellín (Lozada Rodríguez, 2025).
- **Otras aplicaciones:** *UOL*, en Brasil, transcribe vídeos de YouTube; *Infobae México* utiliza IA para limpiar y transcribir audios; *Quispe Chequea*, proyecto de *Ojo Público* en Perú, traduce audios a lenguas originarias como quechua y aimara; y medios españoles, como *Antena 3* y *Radio Bilbao*, experimentan con clones de voz (Apablaza-Campos y Wilches Tinjacá, 2025).

Figura 4

El Espectador (Colombia) incluye la herramienta de text-to-speech en sus contenidos.



(Lozada Rodríguez, 2024).

5.2. Buenas prácticas y recomendaciones

- **Confidencialidad y protección de datos:** al transcribir entrevistas, se deben aplicar protocolos de seguridad que protejan la identidad de las fuentes y eviten filtraciones. Elegir herramientas que cumplan estándares de privacidad es clave.

- **Corrección y edición manual:** las transcripciones generadas por IA pueden contener errores; por ello conviene revisarlas y editarlas para garantizar la fidelidad de las citas.
- **Adaptación a lenguas locales:** proyectos como *Quispe Chequea* demuestran la importancia de traducir y sintetizar audios en idiomas indígenas. Esta práctica mejora la inclusión y puede inspirar iniciativas similares en otras regiones (Lozano Chávez, 2024).

6. Técnicas de SEO asistido

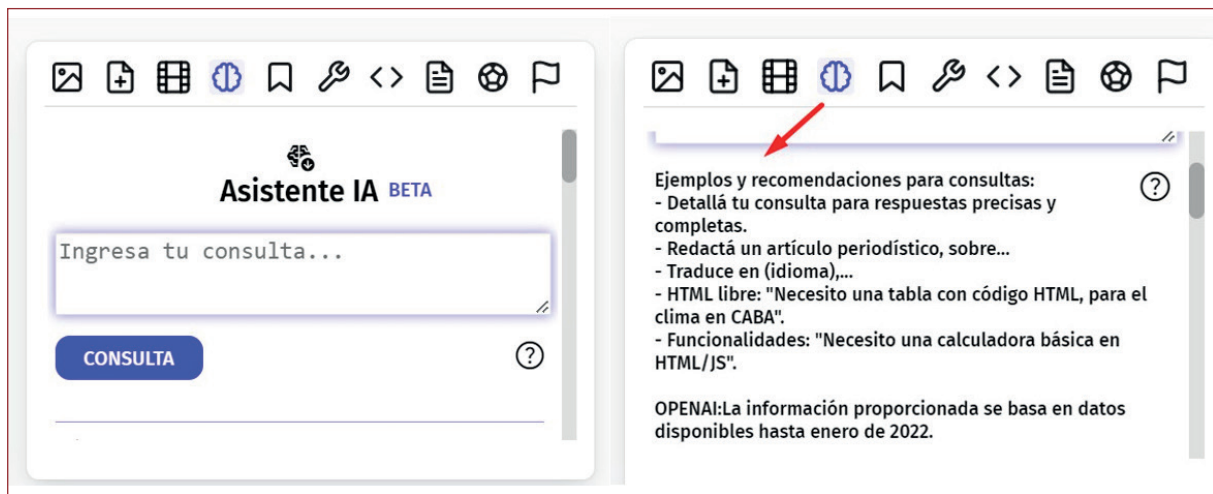
El SEO asistido agrupa a las herramientas integradas en los sistemas de gestión de contenidos (CMS) que dan consejos de redacción y optimización para mejorar la visibilidad de los artículos en motores de búsqueda. Ayudan a ajustar títulos, descripciones y etiquetas, y se apoyan en la IA para sugerir palabras clave y estructuras.

6.1. Ejemplos destacados según aplicación

- **Integración en CMS:** los medios *Guatemala.com*, *Infobae México*, *ABC Color* (Paraguay) y *Meridiano* (Venezuela) son pioneros en implementar estas herramientas que facilitan la redacción ciberperiodística (Apablaza-Campos y Wilches Tinjacá, 2025).
- **SEO asistido e IA generativa:** en Argentina, medios como *Olé* y *Clarín* utilizan asistentes basados en IA que ofrecen consejos de redacción y recomendaciones para mejorar el posicionamiento (Macías, 2024).

Figura 5

Asistente IA, con recomendaciones SEO y opciones de consulta a herramientas como ChatGPT, que está disponible para redactores de *Olé* y *Clarín* (Argentina).



(Macías, 2024).

6.2. Buenas prácticas y recomendaciones

- **Integrar el SEO en la rutina editorial:** emplear estas herramientas desde las fases iniciales de la redacción agiliza el trabajo y evita retrabajos posteriores.

- **Evitar fórmulas automatizadas:** aunque el SEO asistido sugiere estructuras y palabras clave, es recomendable adaptar las recomendaciones al contexto y a la voz del medio para no sacrificar la calidad narrativa ni caer en contenidos genéricos.
- **Formación continua:** mantener a los equipos actualizados en las prácticas de SEO y entender cómo interactúan con los algoritmos de búsqueda ayuda a maximizar el alcance y la relevancia de los contenidos (Apablaza-Campos et al., 2025).

7. Resúmenes de artículos

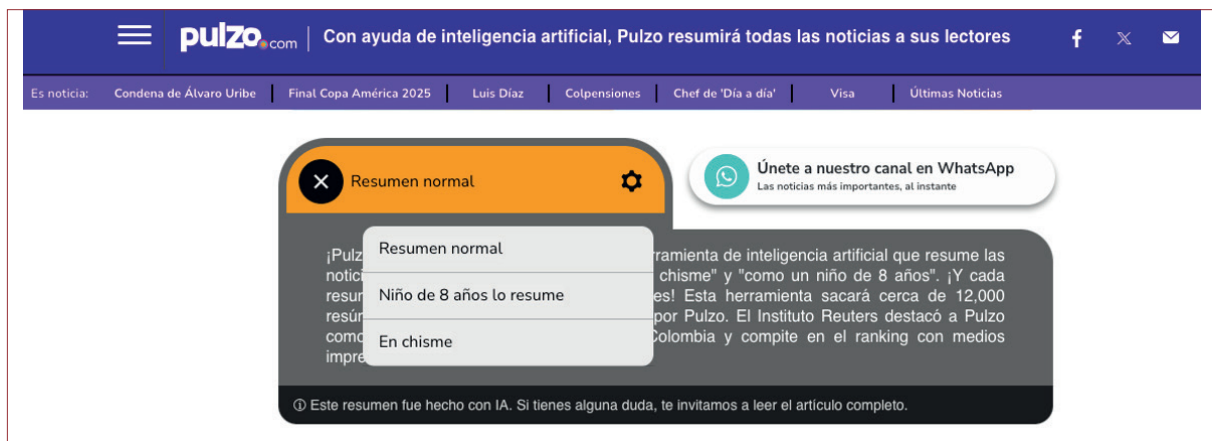
Las herramientas de resúmenes automáticos generan versiones abreviadas de artículos extensos. Algunas ofrecen distintos niveles de compresión o adaptan el lenguaje a públicos específicos, permitiendo a los lectores elegir cómo consumir la información. Son útiles para captar la atención de audiencias con poco tiempo y acercar el contenido a niños o personas con menor nivel de alfabetización mediática.

7.1. Ejemplos destacados según aplicación

- **Modos de síntesis:** mientras *BioBio Chile* sintetiza en un párrafo sus contenidos extensos (Erlandsen, 2024); Pulzo (Colombia) aplica herramientas que sintetizan sus artículos en tres modos: convencional, chisme y para niños de 8 años (Lozada Rodríguez, 2025).
- **Proyecto UALTER:** iniciativa del *Grupo Clarín* (Argentina) que ofrece varias versiones de un mismo texto: una versión convencional, otra abreviada con consejos básicos y opciones adaptadas para lectores jóvenes o con lenguaje coloquial (Macías, 2024).

Figura 6

Pulzo (Colombia) ofrece tres modos de resumen de sus artículos.



(Lozada Rodríguez, 2025).

7.2. Buenas prácticas y recomendaciones

- **Diversificar formatos:** ofrecer diferentes niveles de resumen (por ejemplo, lectura rápida, versión explicada para niños o modo "chisme") amplía el alcance y permite adaptar el contenido a distintos segmentos de audiencia.

- **Mantener la exactitud:** las versiones sintetizadas deben preservar los hechos y el contexto clave. La supervisión humana es necesaria para asegurar que la esencia de la noticia no se distorsione.
- **Presentar opciones claras:** es recomendable señalar explícitamente que se trata de un resumen y ofrecer un enlace al artículo completo, para que los lectores puedan profundizar si lo desean.

8. Chatbots personalizados

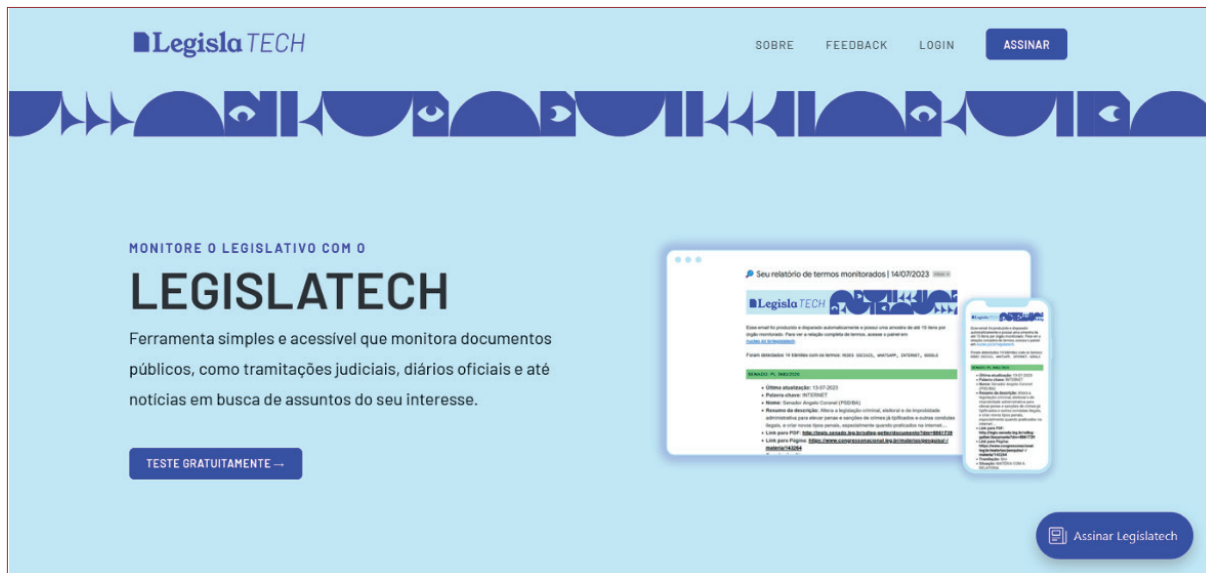
Las herramientas conversacionales basadas en IA permiten a los usuarios interactuar con el contenido de un medio mediante preguntas naturales. Pueden operar como sistemas de verificación de datos, asistentes de búsqueda o guías temáticas. Estos chatbots aportan nuevas vías de distribución y acercan la relación del público con los medios.

8.1. Ejemplos destacados según país

- **Brasil:** *FátimaGPT*, de *Aos Fatos*, y *Legislatech*, de *Núcleo Jornalismo* analizan documentos públicos y permiten realizar consultas personalizadas (Barbosa y Costa Pinto, 2025).
- **Perú:** *Quispe Chequea*, de *Ojo Público*, –ya mencionado en el punto 4– utiliza GPTs personalizados¹ para generar textos de verificación periodística (Lozano Chávez, 2024).

Figura 7

Legislatech es un chatbot, de Núcleo Jornalismo (Brasil), integrado en ChatGPT.



(Barbosa & Costa Pinto, 2025).

8.2. Buenas prácticas y recomendaciones

- **Definir el propósito del chatbot:** algunos chatbots se enfocan en brindar acceso a bases de datos públicas, otros sirven como verificador de información y otros como asistentes de consumo de noticias. Antes de implementarlo, conviene definir la necesidad específica que se quiere resolver.

1 Modelos desarrollados dentro de la interfaz de ChatGPT.

- **Actualización y entrenamiento continuos:** los chatbots deben incorporar datos recientes y ajustar sus respuestas según las consultas de los usuarios. Es recomendable evaluar regularmente su precisión y corregir sesgos o errores.
- **Privacidad y seguridad:** dado que los usuarios pueden compartir datos sensibles, es esencial cumplir la normativa de protección de datos y comunicar cómo se gestionan las conversaciones (Ventura Pocino, 2021).

9. Conclusiones generales y desafíos futuros

La revisión de casos de uso desarrollada demuestra que la inteligencia artificial ya forma parte del ecosistema mediático iberoamericano (Infobae, 2024). Si bien la producción científica respecto a la relación entre periodismo e inteligencia artificial en la región aún es incipiente, con España concentrando el 93% de las publicaciones indexadas en bases de datos como Scopus o Web of Science (Apablaza Campos et al., 2024), los ejemplos recopilados evidencian una adopción progresiva en América Latina, la cual también ha sido adecuadamente reseñada en otros trabajos (Salas et al., 2023; Soto-Sanfiel et al., 2022; Navarro Zamora, 2023; Pinto y Barbosa, 2024).

La IA aparece en roles muy diversos: producción de contenidos, generación de voces, traducción, optimización de visibilidad y servicio al usuario (Newman y Cherubini, 2025). Esta multiplicidad de aplicaciones revela que no existe una única manera correcta de incorporar la IA; cada medio debe evaluar sus recursos, objetivos y público (Apablaza-Campos y Wilches Tinjacá, 2025).

Desde una perspectiva estratégica, conviene resaltar algunas ideas transversales:

- **La IA como herramienta de apoyo:** en todas las áreas analizadas la IA complementa la labor humana. Las redacciones automatizadas y los presentadores virtuales agilizan procesos, pero la supervisión editorial y la creatividad siguen siendo indispensables. La IA debe entenderse como un recurso para ampliar capacidades y no como un sustituto del periodista (Apablaza-Campos y Codina, 2023; Apablaza-Campos, 2024).
- **Ética y transparencia:** la confianza del público depende de que los medios sean transparentes en el uso de IA, informen sobre posibles limitaciones y mantengan un compromiso con la veracidad. Es crucial establecer políticas claras sobre la autoría de los contenidos y la protección de datos personales (Ventura Pocino, 2021).
- **Capacitación continua:** la rápida evolución de estas tecnologías exige programas de formación para periodistas y gestores (Gonçalves & Melo, 2022). La alfabetización en IA contribuirá a un uso más responsable y a la identificación de nuevas oportunidades.
- **Adaptación a contextos locales:** el éxito de proyectos como *Illariy* o *Quispe Chequea* se explica por su atención a lenguas y realidades locales. Los medios deben considerar las necesidades de sus comunidades y aprovechar la IA para fortalecer su vínculo con las audiencias (Kusunoki, 2024; Lozano Chávez, 2024).
- **Colaboraciones estratégicas:** muchas experiencias exitosas (por ejemplo, DataFactory² o acuerdos con plataformas de IA) surgen de alianzas entre medios y empresas tecnológicas. Establecer colaboraciones puede facilitar el acceso a capacidades avanzadas y reducir costos (Apablaza-Campos y Wilches Tinjacá, 2025).

2 Empresa que provee contenidos generados con IA a medios de comunicación de Iberoamérica. Más información <https://www.datafactory.la/>

En el futuro inmediato, los medios de comunicación de Iberoamérica enfrentan el reto de incorporar herramientas cada vez más sofisticadas sin precarizar la profesión periodística ni generar temores de sustitución humana. El desarrollo de políticas públicas, códigos deontológicos y marcos normativos será clave para promover una adopción equilibrada de la IA (Ventura Pocino, 2024). La combinación de innovación tecnológica y compromiso social permitirá que el periodismo aproveche las oportunidades de la IA y contribuya al mismo tiempo a combatir la desinformación.

10. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

11. Referencias

- Alpírez, A. C. (2024). Guatemala: IA: más tiempo para más periodismo. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 97–99). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.
- Apablaza-Campos, A. (2024). Redacción humana versus redacción con inteligencia artificial automatizada: el caso de diario El Rancagüino y las notas algorítmicas. *Hipertext.net*, (28), 143-152. <https://doi.org/10.31009/hipertext.net.2024.i28.11>
- Apablaza-Campos, A. (2025). Intel·ligència artificial en mitjans de comunicació: l'experiència de RTVE a les eleccions al Parlament de Catalunya del 2024. *Comunicació. Revista de Recerca i d'Anàlisi*, 42(1), 9–29. <https://doi.org/10.2436/20.3008.01.250>
- Apablaza-Campos, A., & Codina, L. (2023). *ChatGPT en medios digitales: experiencias periodísticas con inteligencia artificial generativa*. Barcelona: Departament de Comunicació, Universitat Pompeu Fabra. Communication Reports 07. <https://doi.org/10.31009/cr.2023.07>
- Apablaza-Campos, A.; Codina, L.; & Lopezosa, C. (2025). SEO de contenidos e inteligencia artificial: experiencias en medios digitales. En E. Ortega Fernández, M. J. Ufarte Ruiz, & La Rosa Barrolleta, L. (Eds.), *La integración de la inteligencia artificial en la comunicación digital* (pp. 335-350). Editorial Tirant Humanidades. <http://eprints.rclis.org/47034/>
- Apablaza-Campos, A., & Wilches Tinjacá, J.A. (Eds.) (2025). *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (1ª ed. Impresa & 2ª ed. Digital). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica. <https://doi.org/10.15765/librosic.v5i60>
- Apablaza-Campos, A., Wilches Tinjacá, J. A. & Salaverría, R. (2024). Generative artificial intelligence for journalistic content in IberoAmerica: perceptions, challenges and regional projections. *BiD*, 52. <https://doi.org/10.1344/BID2024.52.06>

Banegas Flores, N. C. (2024). Bolivia: entusiasmo y escepticismo ante la IA. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 72–74). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Barbosa, S. & Costa Pinto, M. (2025). Brasil: entre ejemplos significativos y dificultades tecnológicas y financieras (trad. de J. Schena). En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 81–85). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

De Los Reyes, L. & Sarmiento Altuve, Y. J. (2025). Venezuela: primeros (e incipientes) pasos periodísticos. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 120–123). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Del Campo, A.M. & Ramírez Santos, F. (2024). México: mucha cobertura, poca integración y nada de desarrollo. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 100–103). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Erlandsen, P. (2024). Chile: pasos tímidos y en secreto. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 83–85). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Espinosa, C. (2024). Ecuador: desde la automatización hacia la transformación de contenidos. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 90–93). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Ferreira-Candia, R., Daporta, N., Jara, R., & Sales, C. (2025). Paraguay: experiencias a paso lento. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 107–109). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Gonçalves, A. & Melo, P. (2022). Artificial Intelligence and Journalism: an Approach to the Portuguese Context. *Fonseca*, 25, 23–34. <https://doi.org/10.14201/fjc.29682>

Gutiérrez-Caneda, B., Vázquez-Herrero, J., & López-García, X. (2023). AI application in journalism: ChatGPT and the uses and risks of an emergent technology. *Profesional de la información*, 32(5). <https://doi.org/10.3145/epi.2023.sep.14>

Infobae (24 de junio de 2024). Apablaza Campos (docente de la UNIACC): “las rutinas periodísticas han tenido un proceso de evolución con la IA”. <https://www.infobae.com/america/agencias/2024/06/24/apablaza-campos-docente-de-la-uniacc-las-rutinas-periodisticas-han-tenido-un-proceso-de-evolucion-con-la-ia>

Kusunoki, J. (2024). América Latina: generar valor, un desafío regional para la IA. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 66-68). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Lora, E., & Álvarez Álvarez, A. (2024). República Dominicana: ¿probar “lo prohibido” en la redacción? En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 110-113). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Lozada Rodríguez, A. D. (2025). Colombia: manifiestos y (auto)contenido: un ‘Macondo’ generativo. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 89-92). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Lozano Chávez, F. (2024). Perú: primeros pasos para seguir avanzando. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 107-109). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Macías, D. (2024). Argentina: la necesidad de defender nuestra identidad. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 69-71). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Murcia Verdú, F., Ramos Antón, R., & Calvo Rubio, L. (2022). Análisis comparado de la calidad de crónicas deportivas elaboradas por inteligencia artificial y periodistas. *Revista Latina de Comunicación Social*, 80, 91-111. <https://doi.org/10.4185/RLCS-2022-1553>

Navarro Zamora, L. (2023). Inteligencia Artificial en el Ciberperiodismo de América Latina: Estudio exploratorio de las prácticas de los cibermedios y del ciberperiodista. *Question/ Cuestión*, 3(75), e806. <https://doi.org/10.24215/16696581e806>

Newman, N., & Cherubini, F. (9 de enero de 2025). *Journalism, media, and technology trends and predictions 2025*. Reuters Institute for the Study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/journalism-media-and-technology-trends-and-predictions-2025>

Pinto, M. C., & Barbosa, S. O. (2024). Artificial Intelligence (AI) in Brazilian Digital Journalism: Historical Context and Innovative Processes. *Journalism and Media*, 5(1), 325-341. <https://doi.org/10.3390/journalmedia5010022>

Roba, N. (2024). Uruguay: IA, aliada y sospechada. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 114-116). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Salas, M., Siles, I., & Carazo, C. (2023). Imaginarios de periodistas costarricenses sobre la inteligencia artificial y otras tecnologías digitales. *Textual & Visual Media*, 17(1), 61-75. <https://doi.org/10.56418/txt.17.1.2023.4>

Salaverría, R. (2024). España: mucho interés, poca innovación. En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 97–99). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

Segarra-Saavedra, J., Cristófol, F., & Martínez-Sala, A. (2019). Inteligencia artificial aplicada a la documentación informativa y redacción periodística deportiva. El caso de BeSoccer. *Doxa Comunicación*, (29), 275–286. <https://doi.org/10.31921/doxacom.n29a14>

Soto-Sanfiel, M., Ibiti, A., Machado, M., Marín-Ochoa, B., Mendoza, M., Rosell, C., & Angulo-Brunet, A. (2022). In search of the Global South: assessing attitudes of Latin American journalists to artificial intelligence in journalism. *Journalism studies*, 23(10), 1197–1224. <https://doi.org/10.1080/1461670X.2022.2075786>

Ventura Pocino, P. (2021). *Algoritmos en las redacciones: Retos y recomendaciones para dotar a la inteligencia artificial de los valores éticos del periodismo*. Barcelona: Consell de la Informació de Catalunya

Ventura Pocino, P. (2024). Lo más importante es dar sentido de uso a la IA poniéndola al servicio de la misión periodística En A. Apablaza-Campos, & J. A. Wilches Tinjacá (Eds.), *Inteligencia artificial para la generación de contenidos en Iberoamérica: experiencias editoriales en medios de comunicación* (pp. 29–35). DataFactory, Institución Universitaria Politécnico Grancolombiano & Iniciación Científica.

IA y bancos de imágenes: una metodología de investigación

Pere Freixa

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0002-9199-1270>

Mar Redondo-Arolas

Universitat de Barcelona, España

<https://orcid.org/0000-0002-0000-8593>

Lluís Codina

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0001-7020-1631>

Carlos Lopezosa

Universitat de Barcelona, España

<https://orcid.org/0000-0001-8619-2194>

Freixa, P., Redondo-Arolas, M., Codina, L., & Lopezosa, C. (2025). IA y bancos de imágenes: una metodología de investigación. En J. Guallar, M. Váñez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 157-170). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvicom.11.esp>

Resumen

Este capítulo presenta un marco metodológico para analizar el sesgo de género y la presencia de estereotipos socioculturales en bancos de imágenes profesionales, con especial atención a los resultados visuales generados por plataformas fotográficas y generadas con IA. El estudio se basa en la hipótesis de que las indicaciones neutrales (aquellas que carecen de referencias explícitas al género, la edad o la etnia) deberían, en ausencia de sesgos culturales o técnicos, generar una representación visual equilibrada entre las diferentes categorías sociales. Cualquier desviación significativa de dicha proporcionalidad puede indicar la existencia de sesgos implícitos o clichés visuales recurrentes. Para explorar esto, se analizan imágenes recuperadas de cuatro plataformas profesionales: dos basadas en fotografía convencional y dos basadas en la generación de imágenes con IA. Se ha desarrollado un sistema de indicadores codificados para clasificar las representaciones en términos de género, edad, etnia, diversidad funcional, pautas de belleza y acciones representadas. La metodología excluye imágenes de grupo y variantes casi idénticas para garantizar la diversidad y el rigor analítico. Los hallazgos revelan que las plataformas basadas en IA se alinean con mayor consistencia con las indicaciones del usuario (60,36 %) en comparación con las bases de datos fotográficas tradicionales (44,84 %). Sin embargo, ambos tipos de plataformas presentan patrones estereotipados, lo que sugiere la persistencia de tropos visuales y clichés. La metodología propuesta resulta eficaz para detectar estos sesgos y ofrece un marco analítico transferible. El capítulo busca contribuir a esfuerzos más amplios hacia culturas visuales más inclusivas, fomentando la investigación interdisciplinaria sobre la generación y representación algorítmica de imágenes en medios digitales.

Palabras clave

Sesgo de género; Estereotipos; Bancos de imágenes; Inteligencia artificial; Representación visual; Prompts de imágenes; Interpretación algorítmica; Análisis iconográfico; Representación mediática.

I. Introducción

Los bancos de imágenes son uno de los principales recursos usados por los medios de comunicación para complementar con imágenes los contenidos que publican (Codina, 2011; Kamin, 2023). Según varios autores (Gynnild, 2017; Mortensen y Gade, 2023), la reducción de las plantillas de fotógrafos profesionales en los medios de comunicación tradicionales ha propiciado el aumento del uso de imágenes de *stock* para ilustrar informaciones periodísticas, en contra de la imagen comisionada (Mortensen et al., 2023; Ferry, 2023; Mortesen et al., 2024; Hugues, 2024). En su estudio pionero, Tsang (1984) ya detectó cómo a finales de los años setenta y principios de los ochenta del siglo pasado la producción propia de fotografías representaba alrededor de una cuarta parte de las imágenes publicadas. El resto provenía de agencias, bancos de imágenes y fotógrafos *freelance*. No es sorprendente que los bancos de imágenes, junto con las agencias de información, se hayan convertido en el recurso visual con mayor capacidad de influencia y de creación de estándares visuales (Machin, 2004; Machin y Polzer, 2015; Frosh, 2015; 2020).

Las empresas dedicadas a la distribución de fotografía de stock con la incorporación de la inteligencia artificial (IA) han multiplicado su capacidad para generar toda clase de ilustraciones y recursos visuales, tanto fotorrealistas como ficcionales (Perdices-Castillo y Perianes-Rodríguez, 2011; Codina y Lopezosa, 2020; Vrabčič-Dežman, 2024). La IA ha propiciado, además, la irrupción de nuevos actores en el mercado de la ilustración informativa que utilizan las herramientas de generación de imágenes para elaborar ingentes cantidades de contenido visual. A través de interfaces de búsqueda, los medios pueden acceder, localizar y seleccionar aquellas imágenes que se adecúan mejor a sus necesidades en un mercado enormemente competitivo (Allard, 2023; Bright, 2023; Sojit-Pejcha y Crabapple, 2023; Freixa y Redondo-Arolas, 2023; 2024).

La relevancia cada vez mayor de las imágenes de archivo en los medios suscita reacciones dispares. Los detractores del uso tanto de bancos de imágenes como de contenido generado con IA en el periodismo alegan, entre otras cuestiones, que estos recursos reproducen clichés y estereotipos, y facilitan, en consecuencia, una mayor perpetuación de sesgos, afectando a la credibilidad de los medios (Mortensen et al., 2023; Aiello et al., 2023) así como que estas imágenes carecen de calidad editorial o testimonial, al estar desligadas de acontecimientos concretos (Mortensen et al., 2024). Como describe Frosh (2003) "La característica más llamativa de las imágenes de archivo es su estrecha relación con la clasificación. Inscritas en la ideología publicitaria, estas imágenes participan en una representación de la realidad sesgada por la categorización selectiva" (p. 91).

Pero, aunque se critique a los proveedores de imágenes de archivo de reproducir códigos y formulas tipificadas y estereotipadas, al margen de si las fotografías provienen de una u otra fuente, Kwak y An (2016) demostraron mediante sistemas de aprendizaje profundo y grandes volúmenes de datos, cómo los medios utilizan las imágenes publicadas en prensa para reproducir determinados mensajes y, a menudo, consolidar clichés y sesgos, independientemente del origen de esas imágenes.

La preocupación por la reproducción de sesgos (de género, de raza, culturales, ideológicos) en las imágenes de prensa ha preocupado y ha sido tratado tanto de forma genérica (Miller, 1975; Luebke, 1989; Rodgers y Thorson, 2000; Rodgers et al., 2007; Thurlow et al., 2020; Freixa et al., 2025) como en comunicación política (Waldman y Devitt, 1998; Goodnow, 2010; Rönnback et al., 2025).

En estos trabajos se han utilizado metodologías de análisis cualitativas y cuantitativas combinadas, basadas en el recuento estadístico y el análisis de contenido tanto iconográfico como iconológico y simbólico. A menudo las observaciones se han complementado con el análisis de los textos que acompañan las imágenes. El trabajo pionero de Miller, por ejemplo, recuenta el número de veces que aparecen imágenes de hombres y mujeres en *Los Angeles Times* y *The Washington Post* durante un año, las secciones de los periódicos en que se publican y los roles representados (Miller, 1975). Por el contrario, Goodnow (2010) focaliza su estudio en un grupo reducido de imágenes a las que somete a un análisis semiótico detallado, profundizando en los códigos de lectura e interpretación inherentes a las imágenes periodísticas. Como se ha visto, Kwak y An (2016) en su investigación han trabajado de forma cuantitativa con una muestra mucho mayor, formada por dos millones de imágenes. Thurlow et al. (2020) y Freixa et al. (2025) por su parte, han utilizado análisis semiótico y muestras de 600 imágenes para valorar la presencia de estereotipos en las plataformas de imagen de stock.

En este capítulo se presenta una propuesta metodológica que permite la realización de investigaciones visuales de imágenes, determinando claramente los parámetros de observación, así como la dimensión de la muestra. El sistema se ha mostrado eficaz aplicado al estudio de sesgo de género y estereotipos tanto en imágenes de stock de origen fotográfico como creadas por inteligencia artificial (Freixa et al., 2025).

2. Los bancos de imágenes y las plataformas de búsqueda

Los bancos de imágenes utilizan las plataformas digitales para ofrecer sus fondos a los clientes. A través de interfaces de búsqueda, los usuarios pueden acceder, localizar y seleccionar aquellas imágenes que se adecúan mejor a sus necesidades. Se trata de un mercado enormemente competitivo. Las búsquedas, realizadas bien por palabras clave o por medio de *prompts* (peticiones textuales), ofrecen como respuesta galerías visuales en forma de conjuntos de imágenes de pequeño formato, que los usuarios pueden seleccionar y descargar. En función de la complejidad de la búsqueda realizada, los sistemas pueden ofrecer desde unas pocas a cantidades ingentes de imágenes.

Para desarrollar este modelo metodológico, partimos de la conjetura de que estas preselecciones que ofrecen los programas se adecúan a las peticiones de los usuarios, pero que, no obstante, pueden ofrecer sesgos y evidenciar la presencia de estereotipos, entre ellos los de género o raciales, al preponderar en los algoritmos determinados constructos visuales en detrimento de otros.

Se ha partido del supuesto de que los sistemas aplican algoritmos entrenados para ofrecer a los usuarios los mejores resultados posibles, es decir, aquellos que se adecúan mejor a sus demandas. Con algunas diferencias, prácticamente todas las interfaces de búsqueda ofrecen funcionalidades que permiten contextualizar y parametrizar las búsquedas. Algunas de ellas son de carácter técnico, como el tamaño y formato de las imágenes; otras comerciales, como el *copyright*. Algunas permiten justamente indicar preferencias en relación con la edad, la etnia o el género de las personas representadas. Todas ellas añaden la posibilidad de añadir palabras clave que ayuden a obtener resultados más precisos.

Pero situémonos en el escenario más simple y, más común: ¿qué sucede cuando se realiza una búsqueda especialmente neutra, sin ninguna indicación que permita determinar un grupo específico de personas a representar, y tampoco se utilizan las funciones que ofrecen las aplicaciones para limitarla o parametrizarla? ¿Cómo puede actuar el sistema para asegurar el éxito de la respuesta que va a dar al cliente? Se parte de la hipótesis de que, cuanto menor sea la información dada a la plataforma, mayor será el grado de compromiso que el sistema deberá asumir para satisfacer a la demanda. Cuando se realiza una demanda de estas características, dada su ambigüedad, las plataformas ofrecen un gran número de respuestas, con muchos matices y variantes, facilitando la posibilidad de que, alguna de ellas, se adecúe bien a las necesidades del usuario.

Con esta metodología buscamos provocar esta situación. Consideramos que este escenario es el entorno de investigación más adecuado para observar la presencia de clichés, sesgos y estereotipos en los bancos de imágenes. Cuando se lanza un *prompt* neutro a la plataforma sin ninguna otra indicación que permita su parametrización, el sistema debe ofrecer respuestas numerosas y variadas, con el fin de adaptarse lo mejor posible a cualquier tipo de usuario. La observación, recuento y caracterización de las variables ofrecidas por el sistema ante una

demanda expresamente neutral puede indicar la presencia de constructos visuales dominantes, formas de respuesta fuertemente estereotipadas y, en consecuencia, traslucir códigos y sistemas de codificación sesgados.

3. Selección de plataformas, definición del *prompt* y tamaño de la muestra

El mercado de imágenes de stock es enormemente competitivo, en el que conviven numerosas plataformas que ofrecen sus servicios a particulares y medios, tanto gratuitos como de pago, tanto de fotografía convencional como de imagen generada por IA. A partir del criterio de selección por mayor reputación profesional (Piironen, 2022; 2023; May, 2024), para esta investigación se han elegido cuatro plataformas profesionales, dos de imagen tradicional fotográfica y dos de imagen generada con IA. En la elección se ha tenido en cuenta que las plataformas elegidas ofrezcan como resultado a las consultas formuladas un mínimo de 50 imágenes en la galería visual que ofrecen de respuesta, sin necesidad de reformular la búsqueda ni modificar los *prompts*.

Tabla 1

Selección de plataformas profesionales de imagen de stock

Bancos de imágenes		Bancos de imágenes creadas con IA	
Shutterstock	https://www.shutterstock.com/photos	Lexica (Stable diffusion)	https://lexica.art/
Getty Images	https://www.gettyimages.com/	Adobe Stock (opción IA)	https://stock.adobe.com/es/

Fuente: elaboración propia.

Para la formulación de las solicitudes se ha optado por emplear la posibilidad que ofrecen estas plataformas de usar *prompts* en lenguaje natural, en lugar de recurrir a ecuaciones de búsqueda basadas en operadores booleanos (Battelle, 2006; Abadal y Codina, 2008; Codina, 2018). Tal como se ha mencionado anteriormente, partimos de la hipótesis de que el uso de *prompts* neutros activa los sistemas algorítmicos de interpretación incorporados en estas aplicaciones, con el fin de ofrecer la respuesta que el sistema considere más adecuada en función de la petición recibida.

Este enfoque se plantea como una vía idónea para la detección de posibles sesgos y estereotipos: los resultados devueltos por el sistema tenderán a alinearse con aquello que se considera común, normativo o estandarizado dentro de su base de datos o corpus de entrenamiento.

La decisión de utilizar *prompts* en lenguaje natural en lugar de ecuaciones de búsqueda responde a tres motivaciones principales. En primer lugar, por su potencial para activar la lógica de interpretación algorítmica que deseamos observar. En segundo lugar, porque nos interesa explorar y poner a prueba esta modalidad de interacción con bancos de imágenes, dado su carácter emergente e innovador. Y, en tercer lugar, porque queremos analizar qué sucede cuando los usuarios —aun disponiendo en algunos casos de la posibilidad de realizar búsquedas mediante ecuaciones— optan por interactuar a través de *prompts*.

Para la construcción del marco experimental de la investigación, se llevó a cabo una prueba preliminar (pre-test) utilizando dos plataformas: Lexica y Shutterstock. El objetivo era comprobar la eficacia de los *prompts* seleccionados para el estudio. En esta fase inicial se formularon

seis solicitudes que combinaban tres elementos: uno constante y dos variables. El elemento constante fue la expresión "*smiling person*", seleccionada por las connotaciones que implica, entre las que destacan las siguientes:

- El uso del singular debería facilitar la representación de una sola persona en la imagen.
- El término no presupone un género específico.
- El término no presupone una edad.
- El término no presupone una etnia.
- El término no presupone una actividad determinada.
- Se espera que la representación de las personas será frontal, a pesar de no haber formulado indicaciones de encuadre ni composición.

Como variables, se introdujeron dos elementos diferenciadores. En primer lugar, se incorporó un indicador espacial con el objetivo de contextualizar geográficamente las solicitudes. Para ello, se seleccionaron tres escenarios alternativos: *city*, *landscape* y *beach*.

La elección del término *landscape* se fundamentó en una prueba preliminar que consistió en comparar los resultados obtenidos al utilizar *landscape* frente a *countryside*. Al no observarse diferencias significativas en los resultados devueltos por las plataformas, se optó por *landscape* como término de uso más común.

En segundo lugar, se introdujeron variaciones relativas al acabado formal de las imágenes, con el fin de analizar cómo interpretan los sistemas de inteligencia artificial el concepto de representación fotográfica. Se probaron tres formulaciones: *photography of*, *picture of* e *image with photorealistic look*. A través de esta comparativa, se busca examinar no solo la forma en que los sistemas comprenden y representan el grado de realismo o hiperrealismo, sino también los elementos compositivos y de encuadre que emergen en las imágenes generadas o seleccionadas en respuesta a cada *prompt*.

Finalmente, para el pretest, los *prompts* lanzados fueron los siguientes:

- *Picture of a smiling person in a landscape.*
- *Photography of a smiling person in a landscape.*
- *Image of a smiling person in a landscape. Photorealistic look.*

Se realizaron pruebas similares para los otros dos escenarios elegidos, "*city*" y "*beach*".

En todos los casos, las combinaciones que emplearon el término "*photography*" ofrecieron unos resultados más hiperrealistas que el término "*picture*" y que "*image of (...) photorealistic look*". Esta segunda opción proporcionó resultados con alguna imagen caricaturizada, con aspectos claramente distorsionados, propios de la ilustración gráfica.

Se observó que la incorporación del término "*photography*" en el *prompt* —añadido, como se ha explicado, con la finalidad de delimitar formalmente la naturaleza de la imagen en contraposición a la ilustración— puede introducir variaciones en los resultados, particularmente cuando se analizan las "actividades" que desarrollan las personas representadas en las escenas.

Con el fin de valorar el impacto de esta variable, se realizaron pruebas preliminares utilizando dos *prompts* comparables: "*Photography of a smiling person in the city*" y "*Smiling person in the city*", aplicados a las cuatro plataformas analizadas.

En el caso de Shutterstock, la inclusión del término “*photography*” dio lugar, efectivamente, a un cambio en los resultados: varias de las imágenes mostraban a personas tomando fotografías, mientras que este tipo de escenas apenas aparecía cuando el término no se incluía. Por el contrario, en Getty Images, los resultados obtenidos fueron prácticamente idénticos en ambos casos, sin que el uso de “*photography*” introdujera variaciones relevantes.

En la plataforma Lexica, como se ha señalado previamente, el término “*photography*” actuó como filtro eficaz para eliminar ilustraciones no deseadas de los resultados. No obstante, su presencia o ausencia no modificó de manera significativa ninguno de los parámetros de análisis que se quería considerar. De forma similar, en Adobe Stock tampoco se detectaron variaciones sustanciales en los resultados visuales en función del uso del término, aunque, al igual que en Lexica, cuando no se incluye “*photography*”, es frecuente la aparición de ilustraciones entre las imágenes recuperadas.

En resumen, se constató que la presencia o ausencia del término “*photography*” no afectó en ningún caso a la representación de sesgos vinculados con el sexo, la raza, la edad o los estereotipos asociados a modelos de belleza. Solo en la plataforma Shutterstock se identificó una ligera desviación relacionada con las acciones representadas, atribuible al uso de *prompts* que incluyen explícitamente el término “*photography*”.

Es por ello por lo que, finalmente, se optó por lanzar a las cuatro plataformas estos tres *prompts*:

- *Photography of a smiling person in a landscape.*
- *Photography of a smiling person in the city.*
- *Photography of a smiling person on the beach.*

A partir de estas combinaciones de *prompts* se obtuvo la muestra de análisis, compuesta por un total de 600 imágenes. Esta muestra se generó mediante la recopilación de las primeras 50 imágenes devueltas por cada una de las cuatro plataformas, siempre que cumplieran con los criterios establecidos y ofrecieran una respuesta coherente a las tres búsquedas planteadas. La dimensión de esta muestra —600 imágenes— ha sido previamente empleada en estudios similares, como el de Thurlow et al. (2020), orientados a la identificación de sesgos y estereotipos en bancos de imágenes digitales.

4. Qué podemos observar y cómo

Se parte de la hipótesis de que el uso de *prompts* formulados sin ninguna referencia explícita al género, la edad o la etnia de las personas representadas debería generar, en ausencia de cualquier sesgo cultural o técnico, una distribución proporcional entre las distintas categorías observables. Cualquier desviación respecto a ese equilibrio teórico es interpretada, desde la literatura especializada, como un posible indicio de brechas de género y presencia de clichés o estereotipos (Cook y Cusack, 2010; Castillo-Mayén y Montes-Berges, 2014).

La sobrerrepresentación de determinados grupos, por ejemplo, en términos de género, fenotipo o rango etario, puede evidenciar la existencia de reiteraciones visuales: imágenes similares entre sí, con estructuras compositivas o elementos icónicos repetitivos, que reflejan codificaciones culturalmente aceptadas y que, por tanto, pueden ser consideradas estereotipos (Ángeles-Galiano, 2023).

En el ámbito de los medios de comunicación, estas fórmulas visuales codificadas, los tropos y estereotipos, han demostrado una notable eficacia como recurso comunicacional para ilustrar acontecimientos. Constituyen parte de las rutinas profesionales de producción visual y tienen una presencia destacada, especialmente, en la imagen fotográfica (Baeza, 2001; Freixa y Redondo-Arolas, 2022). Los medios no solo reproducen estos constructos visuales, sino que los actualizan constantemente a través de pequeñas variaciones, adaptándolos a las preferencias y expectativas de sus audiencias. De este modo, se convierten en corresponsables de su mantenimiento y renovación simbólica (Quin y McMahon, 1997; Novaes-Cirjanic, 2017).

La observación de repeticiones formales y temáticas, así como la identificación de elementos comunes en las imágenes analizadas, facilita la detección y descripción de los estereotipos más recurrentes presentes en el corpus visual. Para ello, se recurre a los métodos de análisis iconológico, y más específicamente, a la descripción iconográfica, con el objetivo de localizar aquellos atributos que puedan estar asociados a sesgos visuales derivados de clichés o estereotipos socioculturales (Drainville, 2018; Hariman y Lucaites, 2007; 2016; Panofsky, 1979).

Como instrumento metodológico para evaluar la presencia de sesgos de género y estereotipos, se propone la aplicación de una ficha sistemática de análisis, estructurada a partir de los parámetros e indicadores que se detallan en el apartado siguiente.

En cuanto al procedimiento, aquellas imágenes que incluían a más de una persona, que eran claramente fotografías de grupo o que no mostraban ninguna figura humana, fueron descartadas del análisis, con el fin de asegurar la consistencia y comparabilidad de los resultados. Asimismo, la información textual asociada a cada imagen fue empleada como apoyo para facilitar su clasificación.

4.1. Parámetros e indicadores. Sesgos y estereotipos

A continuación, se describen los indicadores que se propone observar, la definición de cada uno de ellos, el procedimiento de observación y los valores utilizados.

4.1.1. Indicador: Género

- **Definición:** Los *prompts* lanzados presuponen la obtención de imágenes en las que pueden aparecer una o varias personas. A priori no existe ninguna indicación que determine si esa persona debe ser de uno u otro género.
- **Procedimiento:** Si en la imagen aparece una sola persona, clasificar si se trata de hombres, mujeres u otros. Anotar el número de ítems que aparecen representando a cada categoría.
- **Valores:** Mujer / Hombre / no definido.

4.1.2. Indicador: Edad

- **Definición:** Los *prompts* lanzados presuponen la obtención de imágenes en las que pueden aparecer una o varias personas. A priori no existe ninguna indicación que determine la edad de esas personas.
- **Procedimiento:** Si en la imagen aparece una sola persona, clasificar la franja de edad representada. Anotar el número de ítems pertenecientes a cada categoría.
- **Valores:** Infantil / joven / adulto / tercera edad.

4.1.3. Indicador: Etnia

- **Definición:** Los *prompts* lanzados presuponen la obtención de imágenes en las que pueden aparecer una o varias personas. A priori no existe ninguna indicación que determine la etnia o etnias de esas personas.
- **Procedimiento:** Si en la imagen aparece una sola persona, clasificar las etnias representadas. Anotar el número de ítems de cada categoría.
- **Valores:** caucasiano / asiático / afroamericano / latino-mediterráneo / indoamericano / oriente medio.

4.1.4. Indicador: Diversidad funcional

- **Definición:** Los *prompts* lanzados presuponen la obtención de imágenes en las que pueden aparecer una o varias personas. A priori no existe ninguna indicación que determine si las personas representadas pueden mostrar o no diversidad funcional.
- **Procedimiento:** Si en la imagen aparece una sola persona, clasificar distintas categorías de disfuncionalidad. Anotar el número de ítems de cada categoría.
- **Valores:** disfuncionalidad motora / disfuncionalidad sensorial / disfuncionalidad intelectual / sin disfuncionalidad.

4.1.5. Indicador: Modelos de belleza

- **Definición:** Los *prompts* lanzados presuponen la obtención de imágenes en las que pueden aparecer una o varias personas. A priori no existe ninguna indicación que determine si las personas representadas responden o no a patrones de belleza y a códigos de representación.
- **Procedimiento:** Si en la imagen aparece una sola persona, clasificar distintos elementos visuales que puedan asociarse a patrones de belleza. La reiteración de elementos permitirá establecer un etiquetado común. Anotar el número de ítems de cada categoría.
- **Valores apriorísticos:** ninguno.
- **Posibles valores:** maquillaje / gafas de sol / gafas / sombreros, gorras, cascos / joyas / relojes / ropa verano / ropa invierno / mochilas / bolsos / otros.

4.1.6. Indicador: Inventario de acciones

- **Definición:** Los *prompts* lanzados presuponen la obtención de imágenes en las que pueden aparecer una o varias personas en lugares indefinidos. A priori no existe ninguna indicación que determine si las personas representadas deben o no realizar alguna acción. Las acciones pueden actuar como indicadores de códigos de conducta preestablecidos y estereotipados.
- **Procedimiento:** Si en la imagen aparece una sola persona, clasificar las acciones que realiza. La reiteración de elementos permitirá establecer un etiquetado común. Anotar el número de ítems de cada categoría.
- **Valores apriorísticos:** ninguno, aunque se presupone una mayor incidencia de representaciones de dispositivos fotográficos por la naturaleza del *prompt* utilizado.

- **Posibles valores:** realizar selfie / mirar mapa / leer libro / consultar smartphone, tablet / realizar gestualidad / trabajar / beber, comer / utilizar cámara o tecnología / jugar / conducir, estar en un coche / bailar / utilizar transporte público / realizar deporte / andar / correr / otros.

Una vez obtenidos los resultados, se propone realizar comparativas de cada parámetro e indicador entre las distintas plataformas, para confirmar o no la presencia de sesgos y de estereotipos.

5. Resultados esperados y limitaciones

El proceso de obtención de la muestra presenta particularidades derivadas del funcionamiento algorítmico de las plataformas empleadas. A pesar de la precisión que pueden indicar los *prompts*, los sistemas no siempre devuelven imágenes que se correspondan estrictamente con lo solicitado. En el caso de esta investigación, se constató que las galerías de resultados incluyen, con frecuencia, fotografías grupales, retratos de parejas o imágenes en las que no aparece ninguna figura humana.

Asimismo, se observó que los bancos de imágenes fotográficos tradicionales tienden a ofrecer múltiples versiones de una misma escena. Estas variantes suelen presentar mínimas diferencias, generalmente relacionadas con el encuadre o la composición. Por ejemplo, ante un *prompt* que solicita el retrato de una mujer en una calle urbana, el sistema puede devolver cinco o seis imágenes prácticamente idénticas, con leves ajustes de encuadre y composición.

Para la conformación de la muestra definitiva se eliminaron, por un lado, los denominados falsos positivos —esto es, las imágenes que no se ajustaban al contenido solicitado— y, por otro, las variaciones redundantes de una misma toma, con el fin de evitar distorsiones en el análisis y garantizar una mayor diversidad visual.

En este sentido, se comprobó que los bancos de imágenes generados mediante inteligencia artificial tienden a ofrecer respuestas más precisas frente a las demandas formuladas. A modo ilustrativo, en el marco de esta investigación, los resultados mostraron que los bancos de IA alcanzaron un 60,36 % de adecuación a los *prompts* planteados, frente a un 44,84 % en el caso de los bancos fotográficos convencionales. Esta diferencia puede atribuirse, en buena medida, a los distintos modelos de negocio que sustentan cada tipo de plataforma. Mientras que los bancos tradicionales están orientados a la explotación de un archivo preexistente —lo que los lleva a ofrecer imágenes ya disponibles, aunque no coincidan plenamente con la búsqueda—, los bancos basados en IA están diseñados para generar contenidos nuevos que se ajusten con mayor precisión a las demandas del usuario. En consecuencia, priorizan la eficacia de la respuesta por encima de la diversidad.

El sistema metodológico propuesto se ha mostrado eficaz para la observación de los indicadores definidos. Los resultados obtenidos (Freixa et al., 2025) revelan pequeñas pero significativas variaciones en la perpetuación de sesgos, comparando el comportamiento de las plataformas de generación de imágenes por IA con el de los bancos de fotografía convencional.

Con esta metodología se busca aportar una herramienta sistematizada que facilite la identificación de sesgos y estereotipos tanto en bancos de imágenes como en cualquier otro sistema visual que ofrezca respuestas gráficas a través de *prompts* en lenguaje natural. Proponemos,

de este modo, un *framework* de análisis que puede ser adoptado por otros equipos de investigación con el objetivo común de visibilizar estas dinámicas y contribuir a su corrección. Esta tarea exige una acción coordinada en múltiples dimensiones, entre ellas las que nuestra investigación pone en evidencia, con el propósito de avanzar hacia una sociedad más justa e inclusiva.

6. Nota final

Dada su posible utilidad para otros investigadores y su potencial de transferencia, este capítulo explica la metodología de la investigación sobre la presencia de sesgos de género y estereotipos en los bancos de imagen que se expone en el artículo:

Freixa, P., Redondo-Arolas, M., Codina, L., & Lopezosa, C. (2025). AI, Stock Photography, and Image Banks: Gender Biases and Stereotypes. *Hipertext.net*, (30), 197-214. <https://doi.org/10.31009/hipertext.net.2025.i30.05>

7. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

8. Referencias

Abadal, E., & Codina, L. (2008). La diversidad cultural en Google y los motores de búsqueda: una aproximación conceptual. *Profesional de la Información*, 17(2), 191–198. <https://doi.org/10.3145/epi.2008.mar.09>

Aiello, G., Thurlow, C., & Portmann, L. (2023). Desocializing social media: The visual and media ideologies of stock photography. *Social media+society*, 9(1). <https://doi.org/10.1177/20563051231156363>

Allard, D. (August 8, 2023). Navigating the Intersection of AI and Journalism: the Future of Storytelling? *Kultur*. <https://www.kultur.art/navigating-the-intersection-of-ai-and-journalism-the-future-of-storytelling/>

Ángeles-Galiano, T. J. (2023). Análisis del estereotipo como imagen literaria: una propuesta de estudio. *Tropelías. Revista de Teoría de la Literatura y Literatura Comparada*, (39), 145-156. https://doi.org/10.26754/ojs_tropelias/tropelias.2023397277

Baeza, P. (2001). *Por una función crítica de la fotografía de prensa*. Gustavo Gili.

Battelle, J. (2006). *Buscar: cómo google y sus rivales han revolucionado los mercados y transformado nuestra cultura*. Urano.

Bright, J. (20 de noviembre 2023). Photography & AI. *The independent photographer*. <https://independent-photo.com/news/photography-ai/>

Castillo-Mayén, R., & Montes-Berges, B. (2014). Análisis de los estereotipos de género actuales. *Anales de Psicología*, 30(3), 1044-1060. <https://www.redalyc.org/pdf/167/16731690027.pdf>

- Codina, L. (2011). Entender los bancos de imágenes. *Profesional de la Información*, 20(4), 417–423. <https://doi.org/10.3145/epi.2011.jul.08>
- Codina, L. (2018). Sistemas de búsqueda y obtención de información: componentes y evolución. *Anuario ThinkEPI*, 12, 77-82. <https://doi.org/10.3145/thinkepi.2018.06>
- Codina, L., & Lopezosa, C. (2020). *Documentación periodística: búsqueda de imágenes en Internet y bancos de fotografía y vídeo 2020*. Universitat Pompeu Fabra. <https://hdl.handle.net/10230/44063>
- Cook, R. J., & Cusack, S. (2010). *Estereotipos de género. Perspectivas legales transnacionales*. Profamilia. https://www.law.utoronto.ca/utfl_file/count/documents/reprohealth/estereotipos-de-genero.pdf
- Drainville, R. (2018). Iconography for the Age of Social Media. *Humanities*, 7(1), 12. <https://doi.org/10.3390/h7010012>
- Ferry, S. (2 de mayo 2023). Fotógrafo humano versus Inteligencia Artificial. *Fundación Gabo*. <https://fundaciongabo.org/es/etica-periodistica/blogs/fotografo-humano-versus-inteligencia-artificial>
- Freixa, P., & Redondo-Arolas, M. (2022). Mirando al mar. Tropos visuales de resiliencia y superación en la COVID-19. *Arte, Individuo y Sociedad*, 34(2), 777-798. <https://doi.org/10.5209/aris.75517>
- Freixa, P., & Redondo-Arolas, M. (2023). IA y fotoperiodismo: un escenario por definir. *XV Congreso Internacional de Comunicación Digital y Ciberperiodismo. Desinformación en la era de la inteligencia artificial*. 27 y 28 de noviembre de 2023 UPV/EHU. <https://hdl.handle.net/10230/68388>
- Freixa, P., & Redondo-Arolas, M. (2024). IA, contenido visual y periodismo: fundamentos para un debate. *IX Congreso Internacional de la AE-IC: Comunicación e Innovación Sostenible*. Murcia, 28 mayo - 1 junio 2024.
- Freixa, P., Redondo-Arolas, M., Codina, L., & Lopezosa, C. (2025). AI, Stock Photography, and image banks: Gender biases and stereotypes. *Hipertext.net*, (30), 197-214. <https://doi.org/10.31009/hipertext.net.2025.i30.05>
- Frosh, P. (2003). *The image factory: Consumer culture, photography and the visual content industry*. Berg Publishers.
- Frosh, P. (2015). Beyond the image bank: Digital Commercial Photography . En: M. Lister (Ed.), *The photographic image in digital culture* (pp. 131-148). Routledge.
- Frosh, P. (2020). Is commercial photography a public evil? Beyond the critique of stock photography. In *Photography and its publics* (pp. 187-206). Routledge.
- Goodnow, T. (2010). Visual bias in time's "the great divide": A semiotic analysis of Clinton and Obama photographs. *American Behavioral Scientist*, 54(4), 406–416. <https://doi.org/10.1177/0002764210381865>
- Gynnild, A. (2017). The Visual Power of News Agencies. *Nordicom Review*, 38(2), 25-39. <https://doi.org/10.1515/nor-2017-0412>

- Hariman, R., & Lucaites, J. L. (2007). *No caption needed: Iconic photographs, public culture, and liberal democracy*. University of Chicago Press
- Hariman, R., & Lucaites, J. L. (2016). *The public image: Photography and civic spectatorship*. University of Chicago Press.
- Hughes, C. (January 1, 2024). AI-generated images vs. Stock Photography: A comprehensive comparison. *BlogHerald*. <https://blogherald.com/seo/ai-generated-images-vs-stock-photography-a-comprehensive-comparison/> [entrada discontinuada]
- Kamin, D. (2023). *Picture-work: How libraries, museums, and stock agencies launched a new image economy*. MIT Press.
- Kwak, H., & An, J. (2021). Revealing the hidden patterns of news photos: Analysis of millions of news photos through GDELT and deep learning-based Vision APIs. *Proceedings of the International AAAI Conference on Web and Social Media*, 10(2), 99-107. <https://doi.org/10.1609/icwsm.v10i2.14840>
- Luebke, B. F. (1989). Out of focus: Images of women and men in newspaper photographs. *Sex Roles*, 20(3-4), 121-133. <https://doi.org/10.1007/BF00287986>
- Machin, D. (2004). Building the world's visual language: The increasing global importance of image banks in corporate media. *Visual Communication*, 3(3), 316-336. <https://doi.org/10.1177/1470357204045785>
- Machin D., & Polzer L. (2015). *Visual Journalism*. Palgrave Macmillan.
- May, T. (19 de marzo 2024). The best stock photo sites in 2024. *Digital Camera World*. <https://www.digitalcameraworld.com/buying-guides/best-stock-photo-sites> [entrada discontinuada]
- Miller, S. H. (1975). The content of news photos: Women's and men's roles. *Journalism and Mass Communication Quarterly*, 52(1), 70. <https://doi.org/10.1177/107769907505200112>
- Mortensen, T. M., & Gade, P. J. (2023). Does photojournalism matter? News image content and presentation in the middletown (NY) Times Herald-record before and after layoffs of the photojournalism staff. *Journalism & Mass Communication Quarterly*, 95(4), 990-1010. <https://doi.org/10.1177/1077699018760771>
- Mortensen, T. M., McDermott, B. P., & Ejaz, K. (2023). Measuring photo credibility in journalistic contexts: Scale development and application to staff and stock photography. *Journalism Practice*, 17(6), 1158-1177. <https://doi.org/10.1080/17512786.2021.1976073>
- Mortensen, T. M., Piacentine, C., Wen, T., Bost, N., & McDermott, B. P. (2024). True or false? Viewer perceptions of emotional staff and stock photos in the news. *Journal of Media Ethics*, 39(1), 16-32. <https://doi.org/10.1080/23736992.2023.2295942>
- Novaes-Cirjanic, J. (2017). *Aproximación deontológica al fotoperiodismo: Creación de estereotipos en la visión occidental del "tercer mundo"* [Tesis Doctoral, Universitat Autònoma de Barcelona]. <https://ddd.uab.cat/record/187161>
- Panofsky, E. (1979). *El significado en las artes visuales*. Alianza Editorial

- Perdices-Castillo, L., & Perianes-Rodríguez, A. (2011). Sistemas de búsqueda y visualización en bancos de imágenes comerciales. *Profesional de la información*, 20(4), 439–443. <https://doi.org/10.3145/epi.2011.jul.11>
- Piironen, J. (5 de noviembre 2022). Side-by-Side Comparison of The Most Popular Sites To Sell Stock Photos. *Jamo Images*. <https://jamoimages.com/side-by-side-comparison-of-the-most-popular-sites-to-sell-stock-photos/>
- Piironen, J. (12 de noviembre 2023). The 3 best websites to sell stock photos and a bonus. *Jamo Images*. <https://jamoimages.com/the-3-best-websites-to-sell-stock-photos/>
- Quin, R., & McMahon, B. (1997). *Historias y estereotipos*. Grafo.
- Reporters Sans Frontières, RSF. (10 de Noviembre 2023). RSF and 16 partners unveil Paris Charter on AI and Journalism. <https://rsf.org/en/rsf-and-16-partners-unveil-paris-charter-ai-and-journalism>
- Rodgers, S., & Thorson, E. (2000). "Fixing"; stereotypes in news photos: A synergistic approach with the *Los Angeles Times*. *Visual Communication Quarterly*, 7(3), 8–7. <https://doi.org/10.1080/15551390009363436>
- Rodgers, S., Kenix, L. J., & Thorson, E. (2007). Stereotypical portrayals of emotionality in news photos. *Mass Communication & Society*, 10(1), 119–138. <https://doi.org/10.1080/15205430709337007>
- Rönnback, R., Emmery, C., & Brighton, H. (2025). Automatic large-scale political bias detection of news outlets. *PLoS One*, 20(5), e0321418. <https://doi.org/10.1371/journal.pone.0321418>
- Sojit-Pejcha, C., & Crabapple, M. (11 mayo 2023). Artists and journalists call for the restriction of AI illustration in publishing. *Document Journal*. <https://www.documentjournal.com/2023/05/ai-art-generators-molly-crabapple-open-letter-copyright-law-stability-ai-midjourney-dall-e>
- Thurlow, C., Aiello, G., & Portmann, L. (2020). Visualizing teens and technology: A social semiotic analysis of stock photography and news media imagery. *New Media & Society*, 22(3), 528-549. <https://doi.org/10.1177/1461444819867318>
- Tsang, K. J. (1984). News photos in *Time* and *Newsweek*. *Journalism Quarterly*, 61(3), 578-723. <https://doi.org/10.1177/107769908406100314>
- Vrabič-Dežman, D. (2024). Promising the future, encoding the past: AI hype and public media imagery. *AI and Ethics*, 4, 743–756. <https://doi.org/10.1007/s43681-024-00474-x>
- Waldman, P., & Devitt, J. (1998). Newspaper photographs and the 1996 presidential election: The question of bias. *Journalism & Mass Communication Quarterly*, 75(2), 302–311. <https://doi.org/10.1177/107769909807500206>

Pensamiento crítico e inteligencia artificial en la academia: procedimiento de análisis matricial cualitativo para evaluar sistemas de IA

Lluís Codina

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0001-7020-1631>

Elisenda Aguilera-Cora

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0003-0923-9192>

Carlos Lopezosa

Universitat de Barcelona, España

<https://orcid.org/0000-0001-8619-2194>

Pere Freixa

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0002-9199-1270>

Codina, L., Aguilera-Cora, E., Lopezosa, C., & Freixa, P. (2025). Pensamiento crítico e inteligencia artificial en la academia: procedimiento de análisis matricial cualitativo para evaluar sistemas de IA. En J. Guallar, M. Váñez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 171-183). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuicom.12.esp>

Resumen

Este capítulo presenta el Procedimiento Matricial de Análisis de Sistemas de IA (MASIA), un método cualitativo basado en matrices, diseñado para evaluar el rendimiento y la calidad de los sistemas de inteligencia artificial generativa en entornos académicos. MASIA se centra en el análisis de tres componentes clave en las respuestas generadas por IA: síntesis narrativa, uso de fuentes y formulación de nuevos prompts. De esta forma, fomenta el pensamiento crítico en los usuarios de la IA y ofrece valiosas herramientas tanto para la docencia como para la investigación. El procedimiento define variables y parámetros analíticos que permiten la comparación de diferentes sistemas de IA, lo que facilita la toma de decisiones informada en entornos académicos y de investigación. Además, MASIA integra consideraciones éticas, como la trazabilidad, la atribución correcta y la prevención del plagio, lo que lo convierte en un instrumento flexible y adaptable a diversas necesidades y proyectos académicos. El capítulo concluye que MASIA es una herramienta sencilla pero potente para potenciar el pensamiento crítico, optimizar los procesos de enseñanza y aprendizaje y sentar las bases para la investigación comparativa sobre inteligencia artificial en el ámbito académico.

Palabras clave

Inteligencia artificial generativa; Evaluación cualitativa; Pensamiento crítico; Matrices de análisis; Ética académica; Sistemas de IA en el ámbito académico; Métodos de evaluación.

1. Introducción

El objetivo de este trabajo es presentar un procedimiento de análisis para evaluar el rendimiento y la calidad de los sistemas de inteligencia artificial generativa en entornos académicos.

El procedimiento, al que denominamos *Procedimiento Matricial de Análisis de Sistemas de IA* o MASIA, está concebido para evaluar sistemas de IA que, como parte de su respuesta no solamente proporcionan una síntesis narrativa, sino que también incluyen citas y las fuentes bibliográficas que han usado para generar el contenido.

El método de análisis favorece el pensamiento crítico de los usuarios de sistemas de IA, aporta elementos para procesos de enseñanza-aprendizaje y puede ser la base para desarrollar la toma de datos en procesos de investigación.

La utilización de fuentes como parte de la respuesta es una necesidad en el contexto académico porque permite verificar y ampliar la información proporcionada por la IA, así como mantener la cadena de atribuciones (High-Level Expert Group on Artificial Intelligence, 2019; Crompton y Burke, 2023; Kaebnick et al., 2023; Lund et al., 2023; Tilie et al., 2023; Gundersen et al., 2018; World Commission on the Ethics of Scientific Knowledge and Technology, 2019; Dwivedi et al., 2021; Bianchini et al., 2022). Esto último es doblemente conveniente, porque además de incrementar la calidad de la respuesta de las IA, previene el plagio o la inadecuada atribución, siendo ambos aspectos esenciales en trabajos académicos.

La base del procedimiento presentado aquí consiste en matrices de análisis basadas a su vez en esquemas formados por una serie de variables (Codina y Pedraza, 2016). La determinación de las variables procede de la actividad docente y de experiencias previas de los autores en el análisis y la utilización de sistemas de IA para trabajos académicos, muy especialmente, por la necesidad de aportar protocolos del uso crítico de sistemas de IA a estudiantes universitarios e investigadores predoctorales.

Desde la aparición de ChatGPT a finales de 2022, los autores han incorporado el uso de la inteligencia artificial en sus actividades docentes e investigadoras (Lopezosa y Codina, 2023; Lopezosa et al., 2023a; 2023b; Aguilera-Cora et al., 2024a; 2024b; Codina, 2025). Esta integración evidenció la necesidad de contar con una herramienta intelectual que permitiera formar a estudiantes e investigadores predoctorales en el uso adecuado de la IA (Codina y Garde, 2023). También, la necesidad de disponer de un instrumento que permitiera llevar a cabo estudios comparativos de la eficiencia de los sistemas de inteligencia artificial susceptibles de uso en la academia. Esto último, puede ser útil a efectos de investigación, o para que decisores económicos de las universidades, por ejemplo, puedan tener elementos de evaluación para adquirir sistemas de IA (Bhatia, 2023; Whitfield & Hofmann, 2024; Elsevier 2024).

Antes de presentar los componentes del método de evaluación debemos presentar unas precisiones terminológicas, cuya consideración forma parte del procedimiento en sí, así como deberemos considerar cuál es la composición de los resultados de una IA cuando responde a una instrucción del usuario.

2. Terminología

La terminología que se presenta a continuación se considera parte del procedimiento, por lo que resulta necesario establecer con precisión el uso de un conjunto de términos para su adecuada aplicación. Los presentamos en la tabla 1.

Tabla 1
Terminología del procedimiento de evaluación de sistemas de IA.

Término	Explicación
Fuentes bibliográficas	En un sistema de IA de tipo RAG (ver la definición más abajo) es la lista de documentos (artículos de revista, informes, páginas web, etc.) que justifican la respuesta. En el contexto del procedimiento de evaluación, este concepto de fuente es el que se tiene en cuenta si no se señala lo contrario.
Fuentes de información	En un sistema de IA de tipo RAG (ver la definición más abajo) las fuentes de información son los recursos que utiliza para localizar las fuentes en las que se fundamenta su respuesta. Fuentes de información características de las IA tipo RAG pueden ser bases de datos académicas o buscadores como Google.
Indicadores	En un procedimiento de evaluación los indicadores son las características que aportan información sobre aquello que quiere ser evaluado o comparado. P.e., en el análisis comparativo de la economía de las naciones, el índice de desempleo, la inflación o el PIB son indicadores. Por su propia naturaleza también son variables.
Matrices	Estructuras de información basadas en tablas que permiten extraer datos, presentarlos de forma visual y hacer análisis comparativos. En el procedimiento presentado aquí el uso de matrices se considera normativo. La palabra tabla es equivalente. El uso del término matriz introduce la idea de que se trata de una tabla que cumple ciertas condiciones, las más importantes son que se trata de tablas homogéneas y que se organizan de forma que las filas son entidades y las columnas son las propiedades de las entidades.

Término	Explicación
Modelo de IA	También llamado <i>Large Language Model</i> (LLM). Es la denominación técnica de las inteligencias artificiales de tipo generativo por tratarse de su base tecnológica. Un modelo de IA o una LLM no es utilizable por un usuario final, ya que su uso requiere programación y APIs, por lo que los usuarios finales en general trabajan con IA a través de sistemas de IA.
Página de resultados	La respuesta de un sistema de IA se presenta al usuario en una página de resultados que suele tener al menos tres componentes: la síntesis narrativa, una lista de fuentes y un conjunto de nuevos <i>prompts</i> sugeridos.
Parámetros	En un procedimiento de evaluación los parámetros agrupan indicadores o variables. La idea es que un grupo de variables sirve para caracterizar un aspecto significativo de cierta complejidad de aquello que se pretende evaluar. Por ejemplo, un parámetro de las naciones en análisis comparativos es su economía, otro su demografía o su sistema político, etc. Pero para caracterizar a cada parámetro es necesario utilizar indicadores desagregados como el índice de desempleo en el caso de la economía, junto con otros como el PIB, etc.
Prompt	Instrucción en lenguaje natural que utiliza el usuario para obtener la respuesta de una IA.
Prompts sugeridos	Lista de nuevos <i>prompts</i> que algunos sistemas de IA proporcionan como parte de la respuesta a un <i>prompt</i> .
Rapid review	Una modalidad de revisión en la que se omiten algunos de los controles habituales de las <i>systematic reviews</i> para obtener resultados inmediatos con valor preliminar. Las síntesis narrativas de una IA pueden equipararse a una forma de <i>rapid review</i> .
Retrieval Augmented Generation (RAG)	La generación aumentada de recuperación (RAG) consiste en mejorar las respuestas generadas por sistemas de IA al combinar su conocimiento de su base de entrenamiento con información recuperada en tiempo real de fuentes externas tales como bases de datos académicas o especializadas o de buscadores generalistas como Google o Bing. La mayor parte de los sistemas de IA especializados en uso académico son de tipo RAG. También algunas IA generalistas como Perplexity. Y se supone que debería serlo Google cuando integre de forma efectiva su buscador con su IA.
Revisión de la literatura	La revisión de la literatura es un proceso sistemático de búsqueda, selección, análisis y síntesis de la información existente sobre un tema específico. Implica evaluar críticamente los estudios previos, identificar patrones, debates y vacíos en el conocimiento actual, y presentar una visión integral y organizada del estado del arte en ese campo. En sistemas de IA aplicados a la academia, la producción de revisiones de la literatura es una de sus funciones principales.
Scratchpad / Bloc de notas	El <i>scratchpad</i> o bloc de notas es la sección que precede a la síntesis narrativa donde la IA presenta la cadena de razonamientos que ha seguido para llevar a cabo sus tareas, en nuestro caso, la tarea consiste en resolver las cuatro fases que conducen a una síntesis narrativa. La posibilidad de examinar el <i>scratchpad</i> implica la posibilidad, entre otras cosas, de comprobar si la IA ha comprendido los objetivos de la tarea y ha acertado en su enfoque.
Síntesis narrativa	Una síntesis narrativa es el resultado de analizar un conjunto de fuentes utilizando un esquema bien definido y de reunir en un resumen textual coherente los conocimientos principales derivados del análisis de estos documentos. La síntesis narrativa es tanto un producto derivado de una revisión de la literatura como parte de la respuesta de una IA generativa.
Sistema de IA	Se compone de uno o más modelos de IA (también llamados LLM por <i>Large Language Model</i>), de una o varias capas de software, p.e., para consultar bases de datos (p.e., bases de datos académicas), para gestionar referencias, etc., así como de una interfaz de usuario para facilitar el uso de todo el sistema.
Utilidades	Consisten en funciones complementarias que suelen ofrecer los programas de software además de sus funciones principales. Por ejemplo, en un sistema de IA enfocado a la academia, una utilidad puede consistir en recursos para gestionar referencias.
Variables	Una variable es una propiedad de alguna entidad que puede adoptar valores diferentes para cada entidad, o a lo largo del tiempo en la misma entidad. Registrar estos valores permite analizar, caracterizar y comparar entidades. También se pueden denominar indicadores si se desea enfatizar su carácter heurístico.

Fuente: elaboración propia

3. Composición de la página de resultados de un sistema de IA

En una página de resultados de una IA del tipo que estamos considerando analizar podemos determinar la existencia de tres componentes principales:

1. Síntesis narrativa.
2. Fuentes consultadas.
3. Nuevos *prompts*.

En lo que sigue, presentaremos cada uno de estos componentes a efectos de evaluación. Pero antes debemos señalar un cuarto elemento que, aunque no forma parte de la página de resultados, podemos encontrar formando parte de la misma interfaz:

4. Utilidades y funciones adicionales y específicas de cada sistema.

3.1. Síntesis narrativa

- **Trazables mediante un *scratchpad* visible.** El *scratchpad* o bloc de notas es una sección que precede a la síntesis narrativa en la cual la IA presenta de modo transparente y trazable la cadena de razonamientos que ha seguido para resolver la tarea. La posibilidad de examinar esta cadena de razonamientos permite detectar posibles sesgo u otra clase de errores, así como permite verificar si la IA ha entendido bien la tarea, y en todo caso facilita la trazabilidad del proceso seguido.
- **Articuladas.** Significa que la síntesis se presenta mediante alguna clase de estructura. Por ejemplo, en apartados separados y posiblemente organizados por títulos y siguiendo alguna clase de lógica en la distribución de secciones.
- **Presentan coherencia y cohesión.** La coherencia consiste en la interrelación de las oraciones que forman el texto mediante su adecuada vinculación con el tema principal del texto. Se manifiesta por la unidad temática, la ausencia (relativa) de redundancia y la progresión lógica de las ideas. La cohesión, por su parte, se manifiesta en la interrelación de las oraciones por la gramática. Se determina fundamentalmente por los conectores.
- **Presentan conectividad.** El final de cada párrafo anticipa el siguiente, y el inicio de los párrafos siguientes conectan con los anteriores esta se propiedad se hace evidente mediante el uso de conectores. La conectividad se incrementa si hay una sección que reunifica las ideas principales, o algún apartado con una función equivalente.
- **Son (relativamente) extensas.** A igualdad de los demás criterios son preferibles síntesis extensas. Ya que hablamos de un rango que puede ir de las 200 palabras a las 3.000 palabras, son preferibles las que, en caso necesario, se pueden acercar más a este último límite.
- **Son multimodales.** Además de texto incluyen algún formato adicional, p.e., tablas, fichas, mapas conceptuales, mapas mentales o diagramas.

3.2. Fuentes

En las AI de tipo académico, las fuentes suelen ser documentos, informes y artículos de revistas científicas. En definitiva, las fuentes son las que permiten atribuir las ideas y contenidos componentes de la síntesis a sus creadores originales. En el contexto académico tenemos el imperativo categórico de usar sistemas de IA que junto a la síntesis narrativa generada aporten las fuentes en las que se han basado. Este es el motivo de preferir sistemas de IA de tipo RAG.

El motivo de esta preferencia es que la ausencia de fuentes en la respuesta promovería la ruptura de la cadena de atribuciones y en consecuencias predispondría al plagio. Nótese que estamos separando la realización (o no) de plagio del hecho de que algunos sistemas de IA faciliten o promuevan, de facto, el plagio al ofrecer respuestas sin fuentes. Lógicamente debemos preferir sistemas de IA que, al menos, no promuevan el plagio.

A efectos de nuestra propuesta de análisis, son preferibles las IA que en relación con las fuentes:

- Presentan **capilaridad**. Son preferibles las AI que asignan las fuentes a nivel de frases, y en su defecto a nivel de párrafos, mejor que como una lista final que afecta a la totalidad indiferenciada de la síntesis narrativa. La capilaridad implica también conectividad, ya que cuando en un párrafo con (p.e.) tres ideas relacionadas cada fuente está vinculada a cada una de las ideas, en lugar de situar las tres fuentes al final del párrafo o al final de la totalidad de la síntesis.
- Proporcionan **formatos** de citación bien formados. Esto es, con datos referenciales completos y en su caso con enlaces viables.

Además, el usuario de una IA está obligado a verificar y revisar las fuentes, no solo para evaluar los argumentos, sino también para poder atribuir ideas y contenidos de terceros, a sus verdaderos autores mediante el sistema convencional de citación.

3.3. Nuevos *prompts*

Algunas IA ofrecen, como un tercer componente destacado de sus respuestas, una lista de nuevos *prompts* o de nuevas preguntas. Estas propuestas pueden tener escaso interés o pueden ser muy incisivas. En este último caso tienen un evidente valor heurístico. Por tanto, son preferibles las IA que generan, como parte de la página de resultados, *prompts* adicionales, de los que evaluaremos:

- **Oportunidad**. Esto es, ¿los nuevos *prompts* sugeridos son adecuados a los objetivos de la búsqueda?
- **Variedad de enfoques**. Con esta variable nos referimos a la pregunta: ¿ofrecen nuevas facetas o enfoques no considerados en el *prompt* original?

3.4. Utilidades e idiofunciones

Algunos sistemas de IA presentan una o más características que son específicas y únicas del sistema considerado. Frente a las funciones comunes podemos hablar de *idiofunciones* esto es, funciones únicas de cada sistema particular y por tanto únicamente presentes en el sistema considerado. Por ejemplo, una IA puede presentar una funcionalidad consistente en extraer conceptos u otra consistente en poder diseñar matrices de análisis de las referencias.

Se trata de *idiofunciones* porque son propias y únicas de cada IA. Aunque es posible que con el tiempo estas funciones se vayan estandarizando (y el concepto pierda sentido), actualmente son significativas estas diferencias y resulta útil considerarlas.

3.5. Matrices de análisis

Con ayuda de los conceptos anteriores, ya podemos presentar los elementos de análisis, que articulamos en parámetros y variables según muestra la tabla 2.

Tabla 2
Variables de análisis.

Parámetro	Código	Variables /Pregunta de chequeo
1. Síntesis narrativa	1.1	Scrathpad ¿Presenta un <i>scratchpad</i> con la cadena de razonamientos seguidos por la IA? ¿Este documento, se puede seguir consultando una vez completada la tarea?
	1.2	Articulación ¿La síntesis narrativa se presenta organizada o articulada en diversos apartados o se presenta como un continuo sin estructura definida?
	1.3	Coherencia y conexión ¿Se conserva la una unidad temática global en el conjunto de la síntesis narrativa y en el seno de cada párrafo? ¿Hay conexión entre los apartados, secciones o párrafos de la síntesis narrativa?
	1.4	Extensión ¿Es una síntesis de extensión adecuada a los objetivos? ¿Cuántas palabras forman la síntesis narrativa? ¿Ofrece versiones alternativas en relación con la extensión?
	1.5	Multimodalidad ¿La página de resultados incluye solo texto o incluye otras morfologías de la información, como p.e., diagramas? ¿En caso de no ofrecerlas al inicio, las propone como alternativa?
2. Fuentes	2.1	Número ¿Cuántas fuentes cita en total?
	2.2	Diversidad ¿Las fuentes exhiben diversidad adecuada a los objetivos del <i>prompt</i> ? Nota: una misma base de datos no es una limitación <i>a priori</i> de la diversidad. ¿El sistema permite diferenciar si las fuentes pertenecen a textos académicos, de prensa, literatura gris u otras fuentes no regladas?
	2.3	Capilaridad ¿Las fuentes están conectadas al menos al nivel de los párrafos o secciones?
	2.4	Bien formadas ¿Las fuentes están presentadas en un formato que resulta fácil para exportar, gestionar y citar las fuentes?
3. Prompts sugeridos	3.1	Oportunidad ¿Los nuevos <i>prompts</i> sugeridos parecen adecuados u oportunos dada la necesidad de información?
	3.2	Variedad ¿Los <i>prompts</i> son variados y ayudan a ampliar el foco del tema?
4. Idiofunciones	4.1	Funciones propias y exclusivas de cada sistema considerado Además de las funciones comunes examinadas, ¿el sistema presenta otras funciones propias?

Tabla 3

Puntuaciones teóricas.

Parámetro	Código	Variables	Puntuación teórica
Síntesis narrativa	1.1	Scratchpad	0-3
	1.2	Articulación	
	1.3	Conexión	
	1.4	Extensión	
	1.5	Multimodalidad	
Fuentes	2.1	Número	
	2.2	Diversidad	
	2.3	Capilaridad	
	2.4	Bien formadas	
Prompts sugeridos	3.1	Oportunidad	
	3.2	Variedad	
Idiofunciones	4.1	Funciones propias y exclusivas de cada sistema considerado	0-3

La escala de puntuación se ofrece a modo de ejemplo. Para cada uso, los responsables pueden determinar (de forma justificada) otras escalas. En este caso, hemos usado una escala característica de evaluaciones heurísticas en el ámbito de la usabilidad de sistemas de información, y corresponde a la siguiente estimación:

Puntuación	Interpretación
0	Ausencia de función o variable considerada
1	La función o variable aparece en una expresión mínima
2	La función o variable está correctamente implementada pero admite mejoras
3	La función o variable está plenamente implementada

En esta clase de procedimientos de puntuación, las primeras puntuaciones se asignan de modo intuitivo, y se van ajustando a medida que el examen de nuevos casos permite hacer comparaciones. Las puntuaciones se terminan de ajustar cuando se han examinado todos los casos. También es habitual que haya dos analistas encargados de asignar las puntuaciones de modo independiente, luego se comparan las puntuaciones y las discrepancias se resuelven por consenso. No obstante, la escala y el procedimiento concreto para asignar los valores se pueden establecer en cada proyecto concreto.

Tabla 4
Tabla de extracción de datos

Código	Variable	Puntuación
	Síntesis narrativa	
1.1	Scratchpad	
1.2	Articulación	
1.3	Conexión	
1.4	Extensión	
1.5	Multimodalidad	
	Fuentes	
2.1	Número	
2.2	Diversidad	
2.3	Capilaridad	
2.4	Bien formadas	
	Prompts adicionales	
3.1	Oportunidad	
3.2	Variedad	
	Idiofunciones	
4.1	Funciones propias y exclusivas de cada sistema considerado	
TOTAL		

Tabla de síntesis comparativa

Sistema	Síntesis	Fuentes	Prompts	Idiofunción	TOTAL

Las tablas anteriores son ejemplos habituales de sistemas de análisis basados en el uso de matrices. Para cada proyecto concreto, los responsables pueden modificar los aspectos que consideren convenientes.

3.6. Otros modos de evaluación

Es evidente que pueden desarrollarse diferentes modos de evaluación. Por citar algunas alternativas que nos parecen significativas, podemos citar la evaluación en base a tareas, como en Font-Julián et al. (2024) en la que se desarrolla un modelo de evaluación que combina procedimientos de análisis cualitativos y cuantitativos bajo determinadas tareas y con un grupo de dos o más usuarios como jueces que consensúan sus evaluaciones.

Además, hay métodos de evaluación mediante *benchmarking* automatizados, como los que pueden verse en el portal *Artificial Analysis* (<https://artificialanalysis.ai/>) donde se comparan periódicamente decenas de modelos de lenguaje en base a una batería de tests.

3.7. Aportación diferencial de este modo de evaluación

Cualquier modo de evaluación puede ser de utilidad, todo depende del contexto y de los objetivos de cada caso. El modo de evaluación cualitativo que proponemos aquí tiene una triple función:

- Reforzar el pensamiento crítico de los usuarios respecto a la IA.
- Aportar un método de enseñanza/aprendizaje y de adquisición de habilidades en el uso de sistemas de IA.
- Disponer de un procedimiento de evaluación y de análisis comparativo de sistemas de IA basado en el análisis matricial cualitativo de parámetros e indicadores.

3.8. Procedimiento de geometría variable

El procedimiento MASIA presentado aquí aporta esquemas de análisis que se pueden aplicar tal como se presentan, cosa para la que están concebidos. Pero, por un lado, no es imprescindible utilizar todas las variables, y por otro lado se pueden añadir nuevas variables o incluso considerar otros parámetros. La idea de este método de evaluación en su esencia es la que sigue:

- Los diseñadores del análisis, con cualquiera de los tres objetivos enunciados antes, pueden considerar la conveniencia de añadir nuevas variables o de retirar alguna de las mismas.
- A partir de una concepción de la estructura de respuestas y de utilidades de un sistema de IA se propone una serie de variables de análisis, para cuya evaluación se dispone de una guía sucinta y mínimamente viable.
- Los análisis se llevan a cabo mediante matrices como las mostradas en las tablas precedentes, lo que facilita tanto la toma de datos como las operaciones de comparación.

Si el análisis se lleva a cabo con objetivos de investigación, el equipo que desee aplicar el procedimiento MASIA debe añadir los protocolos necesarios de validez, fiabilidad, transparencia y trazabilidad que son habituales en las investigaciones de calidad.

3.9. Fases del procedimiento

Concebimos el procedimiento MASIA como parte de un flujo de trabajo donde se distinguen tres fases. Una fase propia del procedimiento que se presenta en este capítulo y dos fases más que deberían tenerse en cuenta siempre que se utilizan sistemas de IA en contextos académicos, independientemente de las especificidades que determinen su uso.

3.10. Preparación

Se trata de la fase inicial, donde el investigador determina si es conveniente usar sistemas de IA para resolver una tarea, así como seleccionar la más adecuada. En esta primera fase el investigador debe tener en cuenta el objetivo de la tarea a resolver, así como los distintos sistemas de IA que mejor pueden contribuir a su consecución, si es el caso, ya sea mediante herramientas gratuitas o de suscripción, propia o institucional. También se deben valorar aspectos relacionados con la privacidad y la protección de datos. Por último, es conveniente hacer una

prueba o pretest antes de usar formalmente un sistema de IA, especialmente cuando se usa uno por primera vez, para facilitar una evaluación inicial de la eficacia de la herramienta en un área que el investigador conozca bien. Se recomienda aplicar esta fase a cualquier modelo que integre el uso de sistemas de IA en contextos académicos.

3.11. Análisis

Esta fase comprende el procedimiento MASIA para evaluar sistemas de IA que presentan, en su página de resultados, cómo mínimo los siguientes tres elementos: síntesis narrativa, fuentes consultadas y nuevos *prompts*. El investigador evalúa distintos sistemas de IA mediante el uso de matrices y variables, a las que asigna una puntuación en base al cumplimiento de una serie de preguntas de chequeo, que le permiten identificar qué sistema de IA es el más adecuado.

3.12. Ética investigadora

Esta fase debe guiar todo el proceso, desde la concepción de la tarea o proyecto de investigación para el cual se usarán sistemas de IA, hasta su culminación. Como en el caso de la fase de preparación, la ética investigadora debe formar parte de cualquier proyecto de investigación que use sistemas de IA, independientemente del enfoque o modelo adoptado. La fase de ética investigadora contempla la documentación del proceso, mediante el registro de *prompts* utilizados y fecha y contexto de interacción, o capturas de pantalla para fomentar la transparencia y la reproducibilidad. También se deben validar los resultados y verificar las fuentes obtenidas en los sistemas utilizados. Otro aspecto importante es evitar el plagio, tanto en respuestas basadas en trabajos previos como en contenido generado por los propios sistemas de IA. Además, en contextos académicos, es esencial mantener la trazabilidad de los trabajos originales y citar adecuadamente el uso de sistemas de IA.

4. Conclusiones

Hemos presentado un procedimiento relativamente simple para analizar y comparar sistemas de IA aplicados a trabajos y entornos académicos. El procedimiento cubre diferentes funciones, la primera y tal vez más destacada es que apoya el pensamiento crítico de los usuarios de las IA, al hacerles evaluar aspectos significativos de las IA.

Otra función es que, por la misma razón enunciada más arriba, pero no reducible a ella, se puede utilizar en procesos de enseñanza aprendizaje en la formación de estudiantes universitarios e investigadores predoctorales.

Por último, es un procedimiento que puede utilizarse con fines de investigación. Para ello, los miembros del equipo disponen de esquemas sistemáticos con los que analizar y comparar diversas IA a partir de un mismo *prompt* o de un pequeño conjunto de *prompts*.

En general, entendemos que este procedimiento, como señalamos más arriba, es fácil de interpretar y de aplicar, sin perjuicio de su capacidad heurística, que puede ser notable gracias al esfuerzo que los analistas apliquen en el análisis.

5. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

6. Referencias

Aguilera-Cora, E., Lopezosa, C., & Codina, L. (2024a). *Scopus AI beta: Functional analysis and cases* (DigiDoc Editorial Series. DigiDoc Reports). Pompeu Fabra University, Department of Communication. <https://repositori.upf.edu/handle/10230/58658>

Aguilera-Cora, E., Lopezosa, C., Fernández-Cavia, J., & Codina, L. (2024b). Accelerating research processes with Scopus AI: A place branding case study. *Pan-American Journal of Communication*, 6(1). <https://doi.org/10.21555/rpc.v6i1.3088>

Bianchini, S., Müller, M., & Pelletier, P. (2022). Artificial intelligence in science: An emerging general method of invention. *Research Policy*, 51(10), 104604. <https://doi.org/10.1016/j.respol.2022.104604>

Bhatia, P. (2023). ChatGPT for academic writing: A game changer or a disruptive tool? *Journal of Anaesthesiology Clinical Pharmacology*, 39(1), 1. https://doi.org/10.4103/joacp.joacp_84_23

Codina, L. (2025). *Sistemas de inteligencia artificial para la síntesis del conocimiento: marcos teóricos y revisiones de la literatura*. [Presentación]. Facultad de Comunicación, Universidad Pompeu Fabra. <http://hdl.handle.net/10230/70242>

Codina, L., & Garde, C. (2023). Uso de ChatGPT en la docencia universitaria: fundamentos y propuestas[Informe]. *Facultad de Comunicación, Universidad Pompeu Fabra*. <http://hdl.handle.net/10230/570157>

Codina, L., & Pedraza-Jiménez, R. (2016). Características y componentes de un sistema de análisis de medios digitales: el SAAMD. R. Pedraza-Jiménez, L. Codina, & J. Guallar (Eds.), *Calidad en sitios web: método de análisis general, e-commerce, imágenes, hemeroteca y turismo* (pp. 15–40). UOC Publicaciones.

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00392-8>

Dwivedi, Y.K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J.S., Eirug, A., Galanos, V., Ilavarasan, P.V., Janssen, M., Jones, P., Kar, A.K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M.D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>

Elsevier. (2024, January 16). *Launch of Scopus AI to help researchers navigate the world of research*. <https://www.elsevier.com/about/press-releases/launch-of-scopus-ai-to-help-researchers-navigate-the-world-of-research>

- Font-Julián, C. I., Orduña-Malea, E., & Codina, L. (2024). ChatGPT Search as a tool for scholarly tasks: Evolution or return? *Infonomy*, 2(5). <https://doi.org/10.3145/infonomy.24.059>
- Gundersen, O. E., Gil, Y., & Aha, D.W. (2018). On reproducible AI: Towards reproducible research, open science, and digital scholarship in AI publications. *AI Magazine*, 39(3), 56–68. <https://doi.org/10.1609/aimag.v39i3.2816>
- High- Level Expert Group on Artificial Intelligence. (2019, April 8). *Ethics guidelines for trustworthy AI*. European Commission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Kaebnick, G. E., Magnus, D., Kao, A., Hosseini, M., Resnik, D. B., Dubljević, V., Rentmeester, C. A., Gordijn, B., & Cherry, M. J. (2023). Editor's statement on the responsible use of generative AI technologies in scholarly journal publishing. *Medicine, Health Care and Philosophy*, 26(4), 499–503. <https://doi.org/10.1007/s11019-023-10176-6>
- Lopezosa, C., & Codina, L. (2023). Testing Bard: Probando Bard: así funciona la Inteligencia Artificial Generativa de Google. *Anuario ThinkEPI*, 17. <https://doi.org/10.3145/thinkepi.2023.e17a25>
- Lopezosa, C., Codina, L., & Ferran-Ferrer, N. (2023a). *ChatGPT como apoyo a las systematic scoping reviews: integrando la inteligencia artificial con el framework SALSA*. Col·lecció del CRICC. Barcelona: Universitat de Barcelona.
- Lopezosa, C., Codina, L., Boté-Vericad, J. J. (2023b). *Testeando ATLAS.ti con OpenAI: hacia un nuevo paradigma para el análisis cualitativo de entrevistas con inteligencia artificial*. Barcelona: Universitat Pompeu Fabra. Departament de Comunicació. [30] p. (Serie Editorial DigiDoc. DigiDoc Reports).
- Lund, B., Wang, T., Mannuru, N. R., Nie, B., Shimray, S. R., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial intelligence-written research papers and the ethics of large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), 570–581. <https://doi.org/10.1002/asi.24750>
- Tlili, A., Huang, R., Mustafa, M. Y., Zhao, J., Bozkurt, A., Xu, L., ... Wang, H. (2023). Speaking of transparency: Are all artificial intelligence (AI) literature reviews in education transparent? *The Journal of Applied Learning & Teaching (JALT)*, 6(2). <https://doi.org/10.37074/jalt.2023.6.2.15>
- UNESCO. (2023, September 8). *Guidance for generative AI in education and research*. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Whitfield, S., & Hofmann, M. A. (2023). Elicit: AI literature review research assistant. *Public Services Quarterly*, 19(3), 201–207. <https://doi.org/10.1080/15228959.2023.2224125>
- World Commission on the Ethics of Scientific Knowledge and Technology. (2019) *Preliminary study on the ethics of artificial intelligence*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000367823>

Web of Science Research Assistant: análisis funcional y recomendaciones de uso

Carlos Lopezosa

Universitat de Barcelona, España

<https://orcid.org/0000-0001-8619-2194>

Elisenda Aguilera-Cora

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0003-0923-9192>

Lluís Codina

Universitat Pompeu Fabra, España

<https://orcid.org/0000-0001-7020-1631>

Juan-José Boté-Vericad

Universitat de Barcelona, España

<https://orcid.org/0000-0001-9815-6190>

Lopezosa, C., Aguilera-Cora, E., Codina, L., & Boté-Vericad, J. J. (2025). Web of Science Research Assistant: análisis funcional y recomendaciones de uso. En J. Guallar, M. Váñez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 184-200). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvi-com.13.esp>

Resumen

Este capítulo ofrece un análisis funcional integral de Web of Science Research Assistant, una herramienta de inteligencia artificial generativa integrada en la base de datos académica de Clarivate. Diseñada para mejorar los flujos de trabajo de investigación, la herramienta apoya tareas como revisiones de literatura, exploración temática, identificación de expertos y selección de revistas. A través de un examen detallado de su interfaz y funcionalidades, que incluyen la síntesis temática, el mapeo de co-citaciones y la visualización de tendencias, este capítulo destaca el potencial de la herramienta para agilizar el descubrimiento científico. Se presentan varios casos prácticos que demuestran sus capacidades para optimizar búsquedas, identificar artículos seminales y sugerir revistas de publicación relevantes. A pesar de sus características innovadoras, el uso de la herramienta debe guiarse por el pensamiento crítico, la transparencia y las consideraciones éticas. Se anima a los investigadores a ver al asistente no como un sustituto, sino como un complemento de la labor académica. El capítulo concluye enfatizando el valor de una interacción reflexiva y estratificada con la IA para integrar estas tecnologías de forma responsable en la práctica académica.

Palabras clave

Inteligencia artificial generativa; Herramientas de búsqueda académica; Web of Science; Web of Science Research Assistant; Automatización de la revisión de literatura; Ética y transparencia en la investigación.

1. Introducción

Web of Science es una de las bases de datos académicas más importantes del mundo, reconocida por su carácter multidisciplinar y la calidad de sus criterios de selección. Se trata de una base de datos que abarca, desde las Ciencias Experimentales a las Humanidades, pasando por las Ciencias Sociales. Web of Science contiene información sobre documentos publicados desde 1900 hasta la actualidad en casi 23 mil revistas científicas, sumando unos 180 millones de artículos, así como sobre unos 150 mil libros (Clarivate, 2025a).

La cuestión es que, paralelamente a la importancia de bases de datos como Web of Science, la importancia de las herramientas de inteligencia artificial generativa ha ido crecido exponencialmente y han trascendido a todos los ámbitos de la sociedad, incluidos especialmente el científico y el académico. Este capítulo presenta Web of Science Research Assistant, la inteligencia artificial integrada en la base de datos Web of Science.

En lo que sigue, se describe la interfaz de la herramienta y se analizan sus funciones. Como se verá, Web of Science Research Assistant tiene capacidad para transformar el acceso, la exploración, el análisis y la síntesis de la información científica, aunque su uso siempre debe ir acompañado de pensamiento crítico y con perspectiva ética, lo que incluye a su vez, la transparencia.

2. Qué es y cómo funciona Web of Science AI Research Assistant

Web of Science AI Research Assistant, es una herramienta creada por Clarivate que utiliza modelos de IA generativa entrenados con la base de datos de Web of Science Core Collec-

tion para dar soporte a los investigadores y estudiantes que quieran avanzar en sus investigaciones académicas (Clarivate, 2025b).

Se trata de un servicio que va más allá de una base de datos académica tradicional ya que funciona como un asistente que puede ayudar a (1) encontrar artículos seminales y relevantes sobre una disciplina o sobre un tema en cuestión de segundos; (2) agilizar tareas avanzadas como revisiones de literatura, identificación de expertos o selección de revistas para publicar; (3) explorar conexiones entre conceptos académicos, autores y artículos científicos mediante visualizaciones interactivas en forma de mapas de tendencias y redes de co-citación; e incluso (4) realizar búsquedas de documentos científicos en lenguaje natural y en varios idiomas, facilitando el acceso a la información científica sin necesidad de utilizar operadores de búsqueda avanzados.

Cuando el usuario realiza una consulta, la herramienta recupera los documentos más relevantes utilizando algoritmos de similitud semántica y búsqueda por palabras clave. Seguidamente, organiza los resultados según su relevancia y genera respuestas o resúmenes usando los contenidos de los artículos seleccionados. Web of Science Research Assistant permite, además, realizar tanto búsquedas documentales como preguntas de síntesis, adaptando, así, el tipo de respuesta a la necesidad del usuario.

Al igual que otras herramientas como Scopus AI (Aguilera-Cora et al., 2024a; 2024b); Elicit (Arroyo-Machado, 2024); Epsilon (Ren et al., 2025); Perplexity (Torres-Salinas y Arroyo-Machado, 2025); ChatGPT (Torres-Salinas et al., 2024; Boté-Vericad et al., 2024); Copilot (Lopezosa, 2023a; Boté-Vericad, 2024), o Scite (Codina, 2024), Web of Science Research Assistant reduce el esfuerzo manual en la construcción de búsquedas complejas y facilita la identificación de tendencias, vacíos de investigación y conexiones no evidentes entre trabajos científicos.

En lo que sigue, realizamos un análisis funcional de la herramienta para mostrar, con ejemplos, las posibilidades que brinda Web of Science Research Assistant para la investigación.

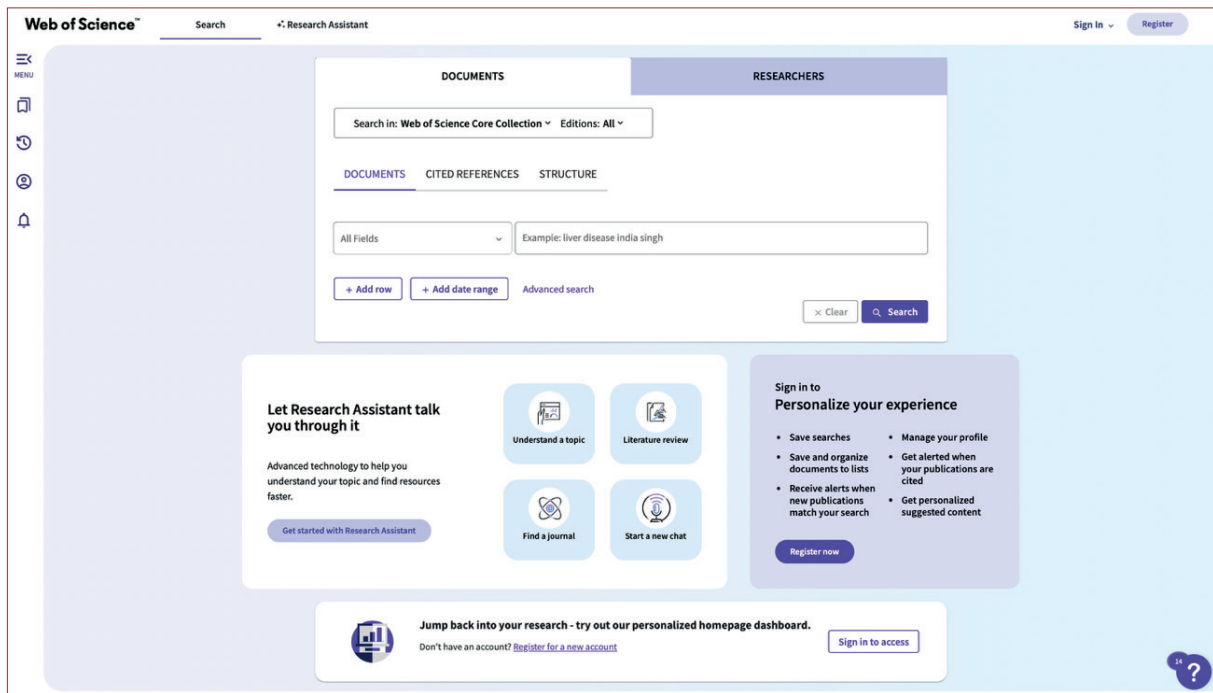
3. Casos prácticos de Web of Science Research Assistant

El asistente de inteligencia artificial de Web of Science ofrece una variedad de informes y visualizaciones diseñados para facilitar la investigación académica. Estas funcionalidades permiten a los usuarios analizar información de manera eficiente y obtener perspectivas visuales sobre los datos y resultados relevantes.

En primer lugar, cuando se ingresa a la plataforma general de Web of Science (figura 1) se observa la interfaz principal, centrada en su función de búsqueda académica. En la parte superior se encuentra un menú que permite acceder a funciones como "Search" y "Research Assistant", además de opciones para iniciar sesión o registrarse. En nuestro caso, para activar el servicio de IA de Web of Science debemos seleccionar la pestaña "Research Assistant".

Figura 1

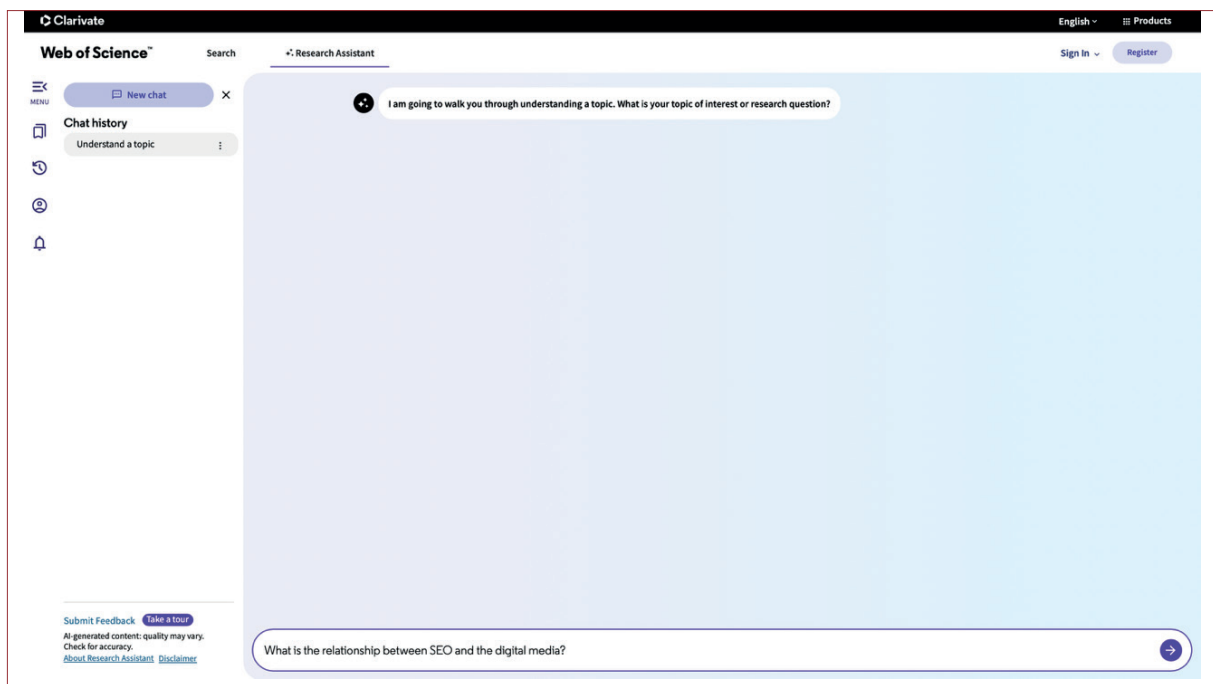
Interfaz de búsqueda principal de Web of Science con opciones avanzadas y asistencia para la investigación.



Una vez hemos ingresado en el recurso "Research assistant" se muestra una nueva interfaz en forma de chat interactivo (figura 2). En el centro de la pantalla, el asistente virtual inicia la conversación con el mensaje: "I am going to walk you through understanding a topic. What is your topic of interest or research question?", invitando al usuario a introducir su tema de interés o pregunta de investigación.

Figura 2

Interfaz del asistente de investigación de Web of Science, que guía al usuario en la comprensión de un tema a partir de una pregunta inicial.



A la izquierda se encuentra un panel lateral con el menú y la sección de "Chat history", donde aparece la conversación actual bajo el título "Understand a topic". También se observan iconos para navegación y opciones del sistema.

En la parte inferior izquierda, hay enlaces para enviar comentarios, hacer un tour o consultar información sobre el asistente de investigación.

Una vez descrita la interfaz de "Research assistant" probamos una instrucción que dice: "What is the relationship between SEO and the digital media?", indicando que el tema a explorar es la relación entre el posicionamiento web (SEO) y los medios digitales.

El resultado de esta instrucción (figura 3) muestra la respuesta generada por la herramienta. En la parte superior aparece un apartado titulado "Overview", donde se explican los principales temas identificados en los resúmenes de los artículos consultados. En nuestro caso, estos temas incluyen técnicas de optimización en buscadores (SEO), características del contenido y metadatos, así como el impacto de los motores de búsqueda en los medios digitales. Asimismo, se destaca la importancia del SEO para mejorar la visibilidad web, el papel del contenido en los rankings de búsqueda y cómo estas dinámicas afectan al periodismo y la investigación académica.

Figura 3

Resumen temático generado por el asistente de Web of Science sobre la relación entre SEO y medios digitales, acompañado de artículos fundamentales sobre técnicas de optimización y visibilidad web.

The screenshot displays the Web of Science Research Assistant interface. At the top, a search bar contains the query "What is the relationship between SEO and the digital media?". Below the search bar, the "Overview" section provides a summary of the main themes identified in the provided abstracts: "search engine optimization techniques", "content characteristics and metadata", and "impact of search engines on digital media". It highlights strategies for improving web visibility, the role of content and metadata in search rankings, and the influence of search engines on journalism and academic research.

Under the "Search engine optimization techniques:" section, there are three bullet points:

- Various techniques, such as PageRank and the use of in-links, are crucial for achieving high rankings in search engines like Google, as demonstrated in a study of optimized web pages ¹.
- Successful search engine optimizers recognize the importance of directories and social bookmarking sites in enhancing page visibility ².
- Professional communicators are encouraged to analyze keywo... [See full response](#)

Below the bullet points, there are icons for liking, sharing, and copying text.

The "Here are some seminal papers:" section lists five papers:

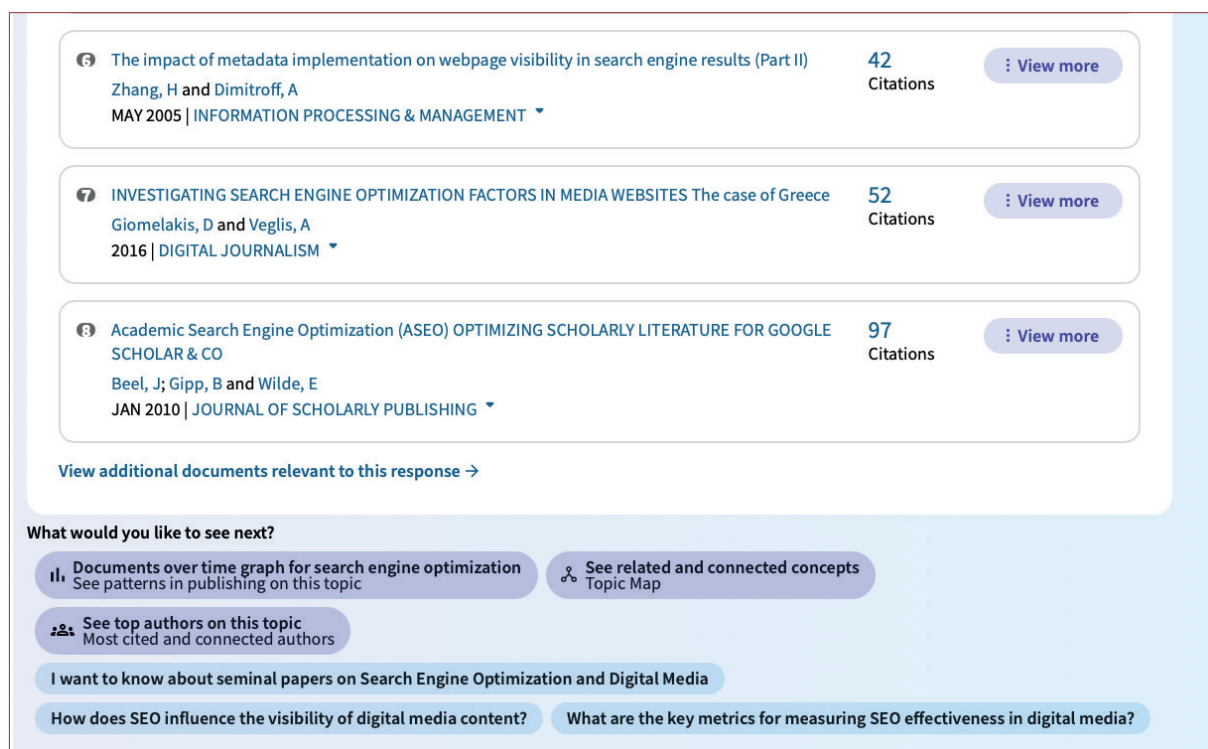
Paper Title	Author(s)	Year	Journal	Citations	View more
The anatomy of a large-scale hypertextual Web search engine	Brin, S and Page, L	APR 1998	COMPUTER NETWORKS AND ISDN SYSTEMS	6,942	View more
Analysing Google rankings through search engine optimization data	Evans, MP	2007	INTERNET RESEARCH	92	View more
The impact of webpage content characteristics on webpage visibility in search engine results (Part 1)	Zhang, J and Dimitroff, A	MAY 2005	INFORMATION PROCESSING & MANAGEMENT	54	View more
How to Use Search Engine Optimization Techniques to Increase Website Visibility	Killoran, JB	MAR 2013	IEEE TRANSACTIONS ON PROFESSIONAL COMMUNICATION	57	View more
What is search engine optimization: SEO?	Yalçin, N and Köse, U	2010	WORLD CONFERENCE ON LEARNING, TEACHING AND ADMINISTRATION PAPERS	47	View more

Debajo del resumen, se ofrece una lista de artículos fundamentales (“Seminal papers”) relacionados con el tema. Entre ellos se incluyen trabajos clave de Brin y Page (1998), con más de 6.900 citas, y otros estudios sobre rankings en Google, características del contenido, y técnicas para aumentar la visibilidad en línea (figuras 3 y 4).

Cada entrada muestra el título del artículo, los autores, fuente de publicación, fecha, número de citas y un botón para ver más detalles. Asimismo, debajo de esta lista aparece un botón para ver documentos adicionales y una sección interactiva titulada “What would you like to see next?”, donde se ofrecen opciones de exploración complementaria. Estas incluyen analizar la evolución de las publicaciones a lo largo del tiempo, ver conceptos relacionados mediante un mapa temático, consultar los autores más citados, y acceder a nuevos *prompts* específicos, que permiten obtener un enfoque guiado y exploratorio del asistente.

Figura 4

Segunda parte de resumen temático generado por el asistente de Web of Science sobre la relación entre SEO y medios digitales.



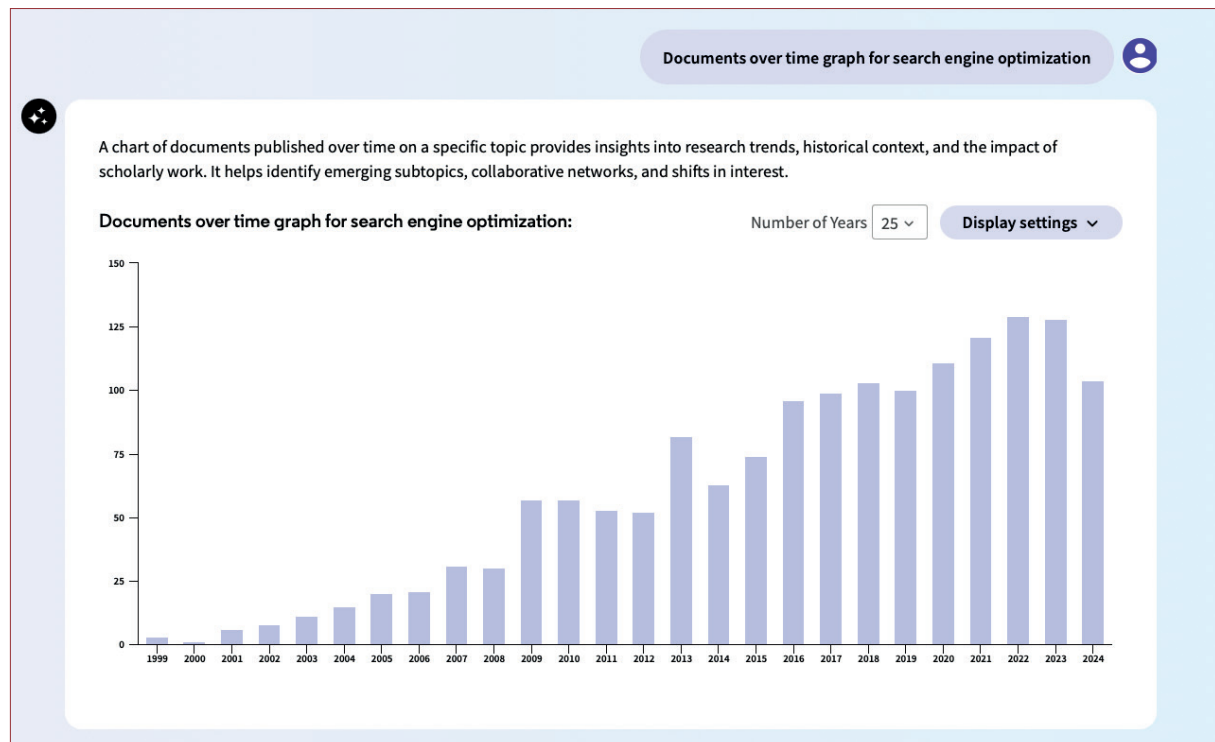
Además de estas opciones, Research Assistant propone otras instrucciones como por ejemplo “Documents over time graph for search engine optimization” (figura 5). Este recurso muestra una visualización gráfica. En la parte superior, el *chatbot* de Clarivate ofrece un breve texto que explica que este tipo de gráfico permite observar las tendencias de publicación sobre un tema específico a lo largo del tiempo, proporcionando contexto histórico, identificando subtemas emergentes y revelando cambios en el interés académico. Justo al lado, se pueden ajustar los parámetros del gráfico mediante un menú desplegable de años (en este caso, 25) y otro de configuraciones de visualización.

En el ejemplo que mostramos, el gráfico de barras ilustra la evolución del número de documentos publicados sobre “search engine optimization” desde 1999 hasta 2024, donde se observa un crecimiento progresivo desde el año 2000, con un notable incremento a partir de

2008. El pico de publicaciones se da en los años 2022 y 2023, superando los 125 documentos anuales, mientras que en 2024 se percibe una leve disminución. Este tipo de análisis permite detectar momentos clave de interés en la investigación sobre SEO, posiblemente relacionados con avances tecnológicos, cambios en los algoritmos de los buscadores o el crecimiento de la comunicación digital.

Figura 5

Gráfico de publicaciones sobre optimización en motores de búsqueda (SEO) a lo largo del tiempo, que muestra el crecimiento sostenido del interés académico desde 1999 hasta 2024.



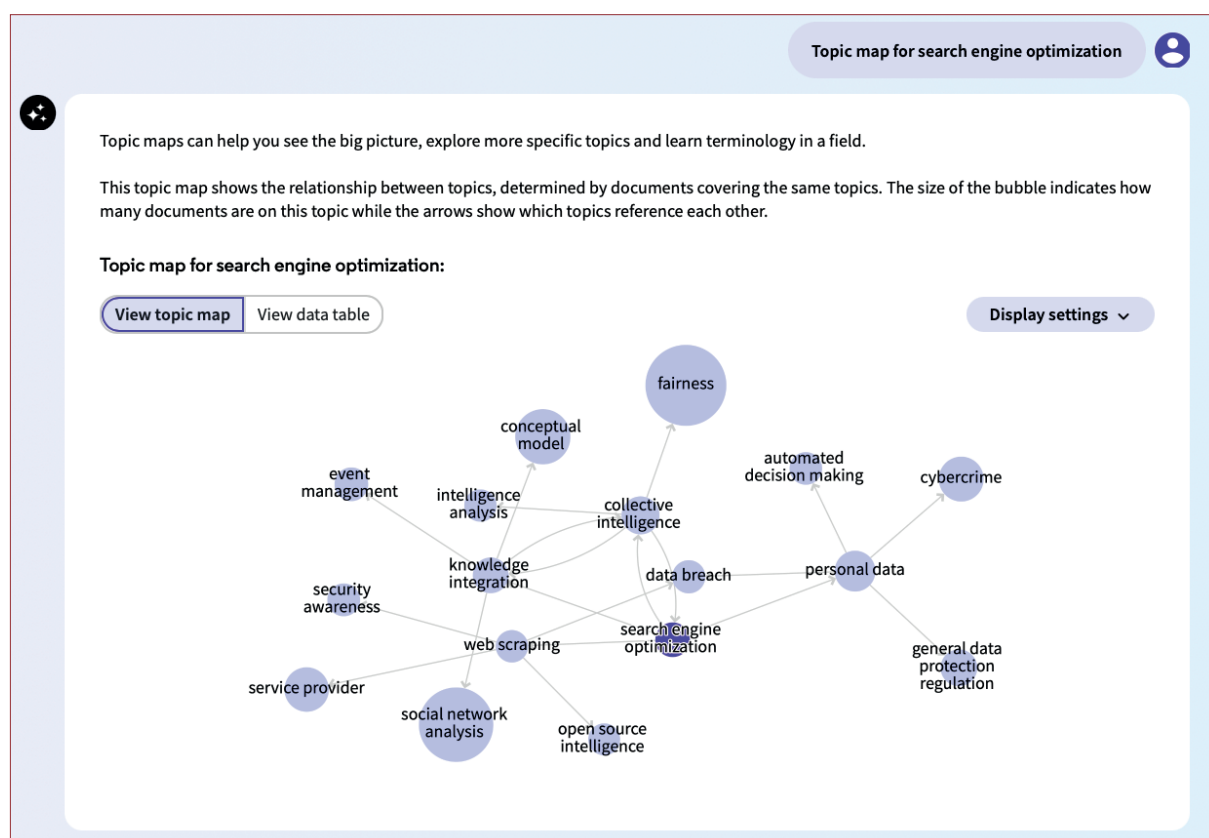
Adicionalmente podemos crear un mapa temático interactivo (*topic map*). En el ejemplo que proponemos, utilizamos de nuevo el tema *search engine optimization* (figura 6). En la parte superior se explica que este tipo de visualización permite ver la relación entre conceptos en un campo de investigación, basándose en los documentos que cubren esos temas. Las burbujas representan distintos temas, y su tamaño indica la cantidad de documentos relacionados con cada uno, mientras que las flechas muestran cómo se referencian entre sí los conceptos.

Respecto a nuestro ejemplo, en el centro del mapa aparece el nodo principal: *search engine optimization*, conectado a términos como *data breach*, *collective intelligence*, *web scraping*, *knowledge integration*, *open source intelligence* o *personal data*. También se visualizan otros temas relevantes vinculados, como *cybercrime*, *general data protection regulation*, *automated decision making*, *fairness*, *social network analysis*, *security awareness* e *intelligence analysis*.

Este recurso es muy interesante, puesto que ofrece un tipo de representación que facilita la exploración de subtemas, la identificación de conexiones interdisciplinarias y el descubrimiento de tendencias emergentes en el campo del SEO.

Figura 6

Mapa temático sobre SEO que visualiza conceptos relacionados y su interconexión, según la coocurrencia de temas en los documentos analizados.



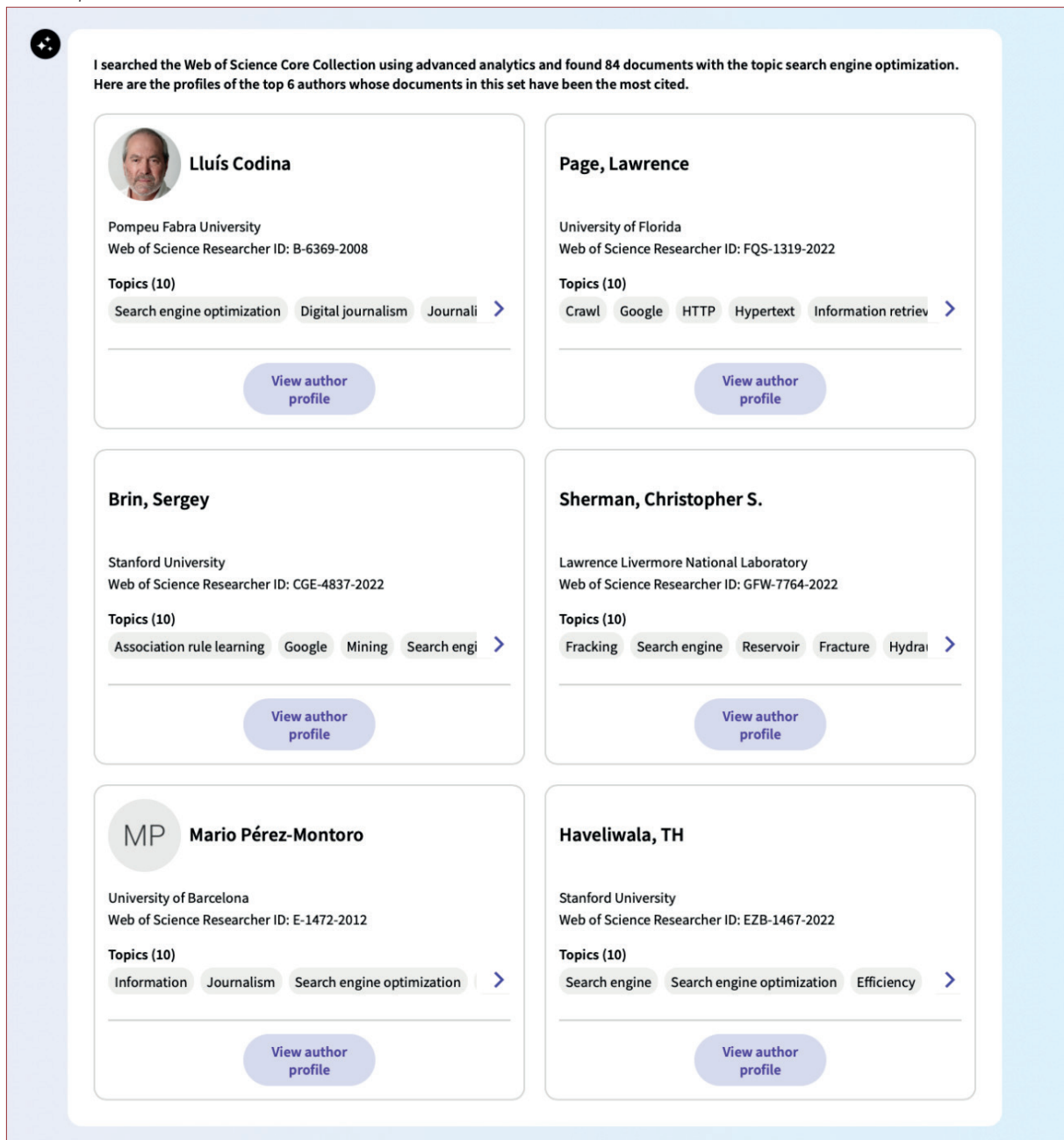
Otro recurso valioso con el que cuenta Research Assistant es el de los autores más citados (figura 7). En el caso de nuestro ejemplo, para el tema *search engine optimization*, se observa que se identificaron 84 documentos relevantes, y se presentan los perfiles de los seis autores cuyos trabajos han sido los más citados dentro de ese conjunto. Cada autor está representado en una tarjeta individual que incluye su nombre, afiliación institucional, identificador de investigador en Web of Science, los temas más frecuentes en sus publicaciones (hasta 10) y un botón para ver su perfil completo. Este panel ofrece una visión clara de quiénes son los principales investigadores en este ámbito y permite profundizar en sus contribuciones.

Otra instrucción muy útil que se le puede pedir al asistente de IA de Web of Science es que identifique artículos seminales de temas específicos. En nuestro caso (figura 8) aplicamos la siguiente instrucción: "I want to know about seminal papers on Search Engine Optimization and digital media".

El resultado es un apartado titulado "Overview", donde se sintetizan tres temas clave extraídos de los resúmenes de los documentos analizados: (1) técnicas de optimización en buscadores, (2) características del contenido y metadatos, y (3) el impacto de los buscadores en los medios digitales. El texto ofrecido por Web of Science AI Research Assistant destaca la importancia de técnicas como PageRank, la optimización de metadatos, la colaboración entre creadores de contenido y la adaptación continua a los cambios algorítmicos para mejorar la visibilidad web.

Figura 7

Perfiles de los seis autores más citados en Web of Science en relación con el tema de la optimización en motores de búsqueda (SEO).



El resultado de esta instrucción también analiza cómo el SEO ha influido en el periodismo y en los medios digitales, subrayando el uso creciente de estas estrategias por parte de los profesionales para aumentar el alcance de sus contenidos. Además, se menciona el surgimiento del concepto de ASEO (*Academic Search Engine Optimization*) para mejorar la visibilidad de publicaciones académicas. La respuesta concluye con un resumen que refuerza la relevancia de estos tres ejes en los ámbitos del periodismo y la investigación académica. Finalmente, se listan artículos seminales, con su título, autores, fuente de publicación, año, número de citas y un botón de acceso a más información (figura 8 y 9).

Figura 8

Análisis temático y bibliográfico sobre SEO y medios digitales, con un resumen de técnicas clave, impacto en el periodismo y referencias a artículos fundamentales.

 I want to know about seminal papers on Search Engine Optimization and Digital Media 



Overview

The main themes identified in the provided abstracts are "search engine optimization techniques", "content characteristics and metadata", and "impact of search engines on digital media". These themes highlight the strategies for improving web visibility, the role of content and metadata in search rankings, and the influence of search engines on journalism and media practices.

Search engine optimization techniques:

- Various techniques, such as PageRank and the use of in-links, are crucial for achieving high rankings in search engines like Google, as demonstrated in a study of optimized web pages .
- Successful search engine optimizers recognize the importance of directories and social bookmarking sites in enhancing page visibility .
- Continuous adaptation to changing algorithms and practices is essential for effective search engine optimization .

Content characteristics and metadata:

- Key content characteristics, including keyword placement and frequency, significantly impact webpage visibility in search engine results .
- The implementation of metadata, particularly the subject field, plays a vital role in improving a webpage's visibility, with effective keywords derived from both the title and full text .
- A combination of content strategies, including keyword analysis and engagement with other web content creators, can enhance search engine rankings .

Impact of search engines on digital media:

- The evolution of digital technologies has transformed journalism, making search engine visibility critical for driving traffic to news websites .
- Media professionals are increasingly utilizing SEO practices to enhance the reach and effectiveness of their content .
- The concept of academic search engine optimization (ASEO) has emerged, focusing on optimizing scholarly literature for better visibility in academic search engines .

Conclusion

In summary, the abstracts collectively emphasize the significance of "search engine optimization techniques", "content characteristics and metadata", and "impact of search engines on digital media". These themes illustrate the strategies and considerations necessary for improving web visibility and the evolving role of search engines in various fields, particularly in journalism and academia.

[See less](#)

  Copy Text

Here are some seminal papers:

 The anatomy of a large-scale hypertextual Web search engine
Brin, S and Page, L
APR 1998 | COMPUTER NETWORKS AND ISDN SYSTEMS ▾

6,942
Citations

 View more

 Analysing Google rankings through search engine optimization data
Evans, MP
2007 | INTERNET RESEARCH ▾

92
Citations

 View more

 The impact of webpage content characteristics on webpage visibility in search engine results (Part 1)
Zhang, J and Dimitroff, A
MAY 2005 | INFORMATION PROCESSING & MANAGEMENT ▾

54
Citations

 View more

Figura 9

Continuación del resultado del análisis temático y bibliográfico sobre SEO y medios digitales, con un resumen de técnicas clave, impacto en el periodismo y referencias a artículos fundamentales.

The screenshot displays a list of five research articles related to Search Engine Optimization (SEO) and digital media. Each article entry includes a numbered icon, the title, authors, publication date, journal name, citation count, and a 'View more' button. At the bottom, there is a link to view additional documents relevant to the response.

Icon	Title	Authors	Date	Journal	Citations	Action
4	How to Use Search Engine Optimization Techniques to Increase Website Visibility	Killoran, JB	MAR 2013	IEEE TRANSACTIONS ON PROFESSIONAL COMMUNICATION	57	View more
5	What is search engine optimization: SEO?	Yalçin, N and Köse, U	2010	WORLD CONFERENCE ON LEARNING, TEACHING AND ADMINISTRATION PAPERS	47	View more
6	The impact of metadata implementation on webpage visibility in search engine results (Part II)	Zhang, H and Dimitroff, A	MAY 2005	INFORMATION PROCESSING & MANAGEMENT	42	View more
7	INVESTIGATING SEARCH ENGINE OPTIMIZATION FACTORS IN MEDIA WEBSITES The case of Greece	Giomelakis, D and Veglis, A	2016	DIGITAL JOURNALISM	52	View more
8	Academic Search Engine Optimization (ASEO) OPTIMIZING SCHOLARLY LITERATURE FOR GOOGLE SCHOLAR & CO	Beel, J; Gipp, B and Wilde, E	JAN 2010	JOURNAL OF SCHOLARLY PUBLISHING	97	View more

[View additional documents relevant to this response →](#)

Por otro lado, Web of ScienceResearch Assistant cuenta con una función que sugiere revistas científicas adecuadas para publicar un trabajo académico específico (figura 10). Cuando accedemos a este recurso, observamos en el centro de la pantalla, un mensaje del asistente que indica: "Let's find journals that could be a good fit for you to publish in by matching your document title and abstract to relevant journals." A continuación, se solicita al usuario que proporcione el título de su documento para comenzar el proceso.

Figura 10

Asistente de Web of Science para sugerencias de revistas, que recomienda publicaciones adecuadas en función del título y resumen del documento.

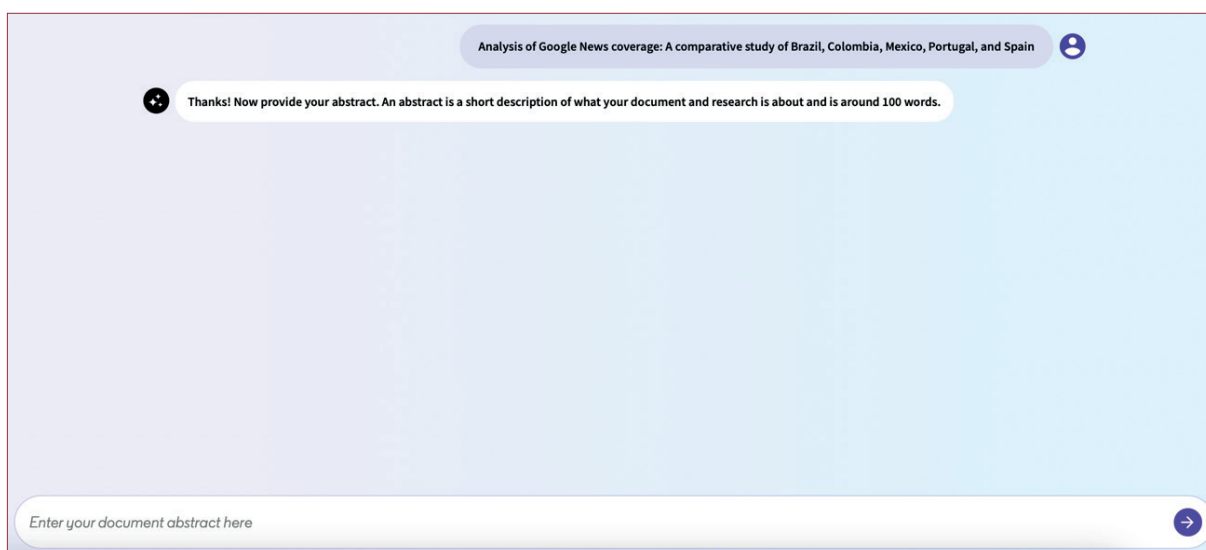
The screenshot shows the 'Research Assistant' interface. At the top, there are links for 'Sign In' and 'Register'. The main content area features a message from the assistant: "Let's find journals that could be a good fit for you to publish in by matching your document title and abstract to relevant journals." Below this message is a text input field with the placeholder text "To get started, please provide the title of your document."

En nuestro caso, ingresamos el título de un manuscrito: “Analysis of Google News coverage: A comparative study of Brazil, Colombia, Mexico, Portugal, and Spain” (figura 11). En respuesta, el asistente solicita el resumen del documento, indicando que debe ser una breve descripción del contenido y la investigación, de aproximadamente 100 palabras (figura 12).

En la parte inferior se encuentra el campo de entrada de texto donde el usuario debe escribir el resumen, acompañado de un botón con flecha para enviarlo.

Figura 11

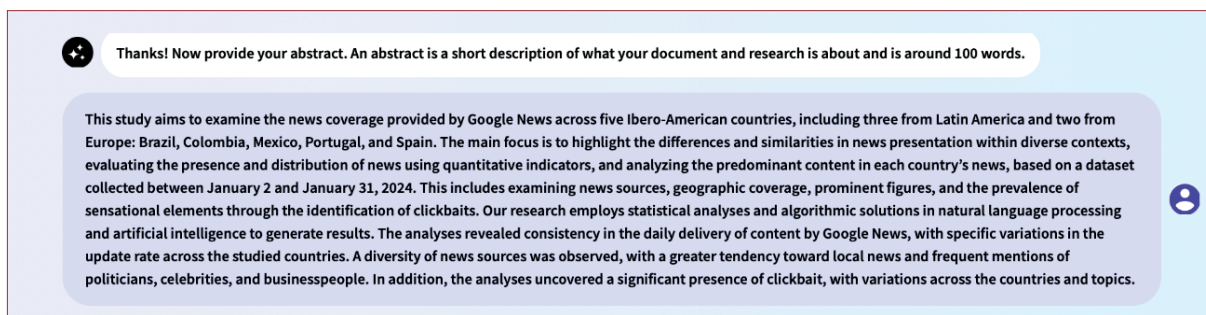
Paso del asistente de Web of Science para sugerir revistas, en el que se solicita el resumen del documento tras ingresar el título del estudio.



The screenshot shows a chat interface with a light blue background. At the top, a message bubble from the assistant says: "Analysis of Google News coverage: A comparative study of Brazil, Colombia, Mexico, Portugal, and Spain". Below it, a system message bubble says: "Thanks! Now provide your abstract. An abstract is a short description of what your document and research is about and is around 100 words." At the bottom, there is a text input field with the placeholder "Enter your document abstract here" and a blue arrow button to the right.

Figura 12

Ejemplo de resumen introducido en el asistente de Web of Science para obtener recomendaciones de revistas, centrado en un análisis comparativo de la cobertura informativa en Google News en cinco países iberoamericanos.



The screenshot shows the same chat interface as Figure 11. The system message bubble is still present. Below it, a message bubble from the user contains an example abstract: "This study aims to examine the news coverage provided by Google News across five Ibero-American countries, including three from Latin America and two from Europe: Brazil, Colombia, Mexico, Portugal, and Spain. The main focus is to highlight the differences and similarities in news presentation within diverse contexts, evaluating the presence and distribution of news using quantitative indicators, and analyzing the predominant content in each country's news, based on a dataset collected between January 2 and January 31, 2024. This includes examining news sources, geographic coverage, prominent figures, and the prevalence of sensational elements through the identification of clickbaits. Our research employs statistical analyses and algorithmic solutions in natural language processing and artificial intelligence to generate results. The analyses revealed consistency in the daily delivery of content by Google News, with specific variations in the update rate across the studied countries. A diversity of news sources was observed, with a greater tendency toward local news and frequent mentions of politicians, celebrities, and businesspeople. In addition, the analyses uncovered a significant presence of clickbait, with variations across the countries and topics." At the bottom right, there is a blue arrow button.

Como resultado de ofrecer el título y el resumen, la herramienta de IA de Web of Science presenta una lista de las 5 revistas científicas que mejor coinciden con el contenido del documento, junto con sus respectivos “match scores”, que indican el nivel de coincidencia temática según palabras clave y áreas de indexación (figura 13). Por último, este paso concluye el proceso de recomendación de revistas, ofreciendo al usuario opciones de publicación según afinidad temática, visibilidad académica y cobertura disciplinaria.

Figura 13

Resultados del asistente de Web of Science con las cinco revistas más compatibles para publicar, según el título y resumen del documento analizado.

Thank you! I'm now searching for journals that match your title and abstract based on keywords.

Here are the top 5 journals that match your documents.

SUSTAINABILITY

Match Score 0.50

Publisher MDPI,ST ALBAN-ANLAGE 66, BASEL, SWITZERLAND, CH-4052

ISSN / eISSN 2071-1050 / 2071-1050

Web of Science Core Collection Science Citation Index Expanded | Social Sciences Citation Index

Additional Indexes Current Contents Agriculture, Biology & Environmental Sciences | Current Contents Social And Behavioral Sciences | Essential Science Indicators

ENVIRONMENTAL SCIENCE AND POLLUTION RESEARCH

Match Score 0.25

Publisher SPRINGER HEIDELBERG,TIERGARTENSTRASSE 17, HEIDELBERG, GERMANY, D-69121

ISSN / eISSN 0944-1344 / 1614-7499

Additional Indexes Biological Abstracts | BIOSIS Previews

FRONTIERS IN PSYCHOLOGY

Match Score 0.22

Publisher FRONTIERS MEDIA SA,AVENUE DU TRIBUNAL FEDERAL 34, LAUSANNE, SWITZERLAND, CH-1015

ISSN / eISSN 1664-1078 / 1664-1078

Web of Science Core Collection Social Sciences Citation Index

Additional Indexes Current Contents Clinical Medicine | Current Contents Social And Behavioral Sciences | Essential Science Indicators

INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS

Match Score 0.19

Publisher SCIENCE & INFORMATION SAI ORGANIZATION LTD,19 BOLLING RD, BRADFORD, WEST YORKSHIRE, ENGLAND, 00000

ISSN / eISSN 2158-107X / 2156-5570

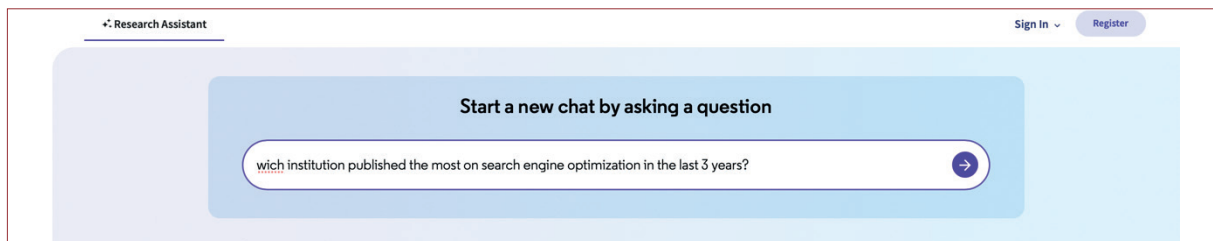
Web of Science Core Collection Emerging Sources Citation Index

Otra instrucción interesante que se puede pedir a Research Assistant es saber qué instituciones han investigado más sobre un tema específico (figura 14). Para ello podemos escribir el siguiente prompt: "Which institution has published the most on search engine optimization in the last 3 years?".

En este caso pretendemos saber qué institución ha publicado más sobre SEO en los últimos tres años. Esta función permite al asistente recuperar y analizar publicaciones recientes para identificar institucionales líderes en investigación sobre ese tema específico.

Figura 14

Consulta inicial en el asistente de Web of Science para identificar qué institución ha publicado más sobre SEO en los últimos tres años.



El resultado (figura 15) es una gráfica de barras. En la parte superior, se explica que la búsqueda se realizó con publicaciones entre el 29 de noviembre de 2021 y el 29 de noviembre de 2024, usando una combinación de términos relacionados con la optimización en buscadores. Y, además, se indica que el resultado arrojó 137,295 documentos.

La gráfica titulada "Institution bar chart" presenta las 10 instituciones con mayor número de publicaciones sobre el tema en cuestión. En el caso de nuestro ejemplo, destaca ampliamente la Egyptian Knowledge Bank (EKB) con más de 5.500 publicaciones, ocupando la primera posición. Esta visualización permite comprender de manera rápida qué instituciones lideran la producción científica en torno al SEO y temas relacionados en el período indicado. Además de esta gráfica, podemos visualizar esta instrucción en modo de "tree map" (figura 16).

Figura 15

Gráfico de barras con las instituciones que más han publicado sobre SEO en los últimos tres años, destacando al Egyptian Knowledge Bank como la más productiva.

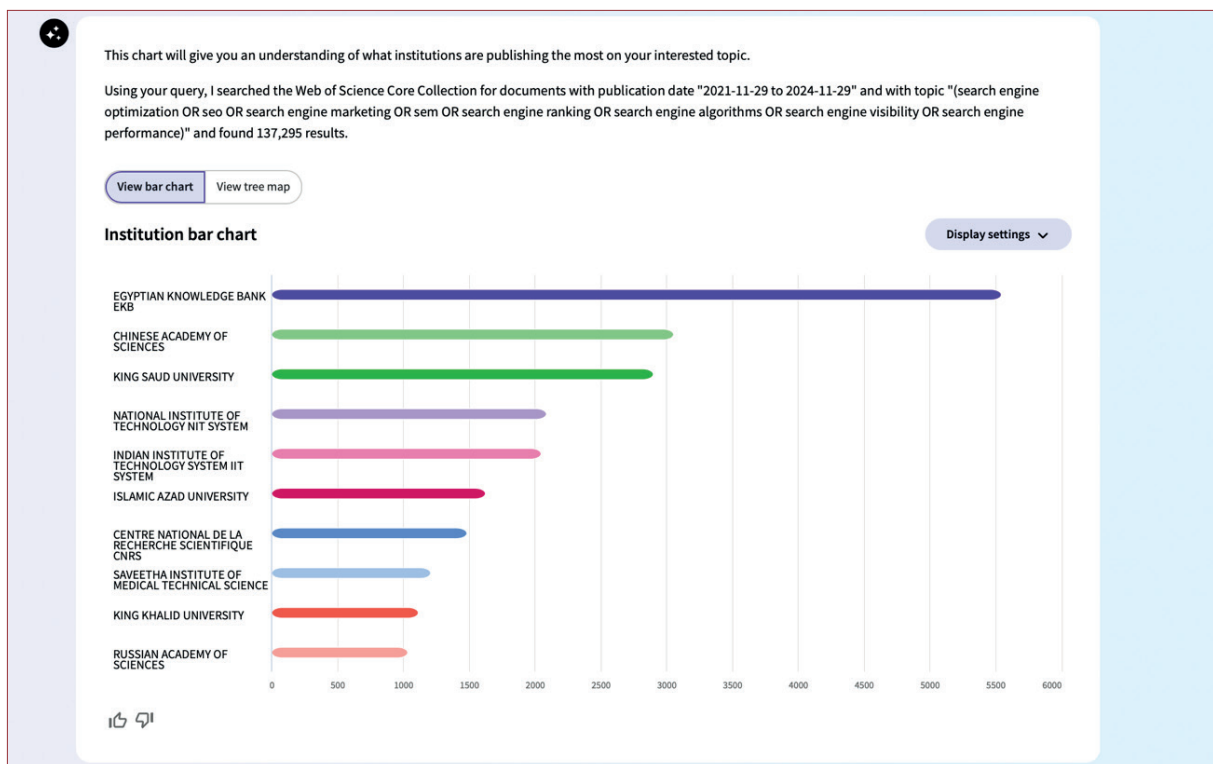
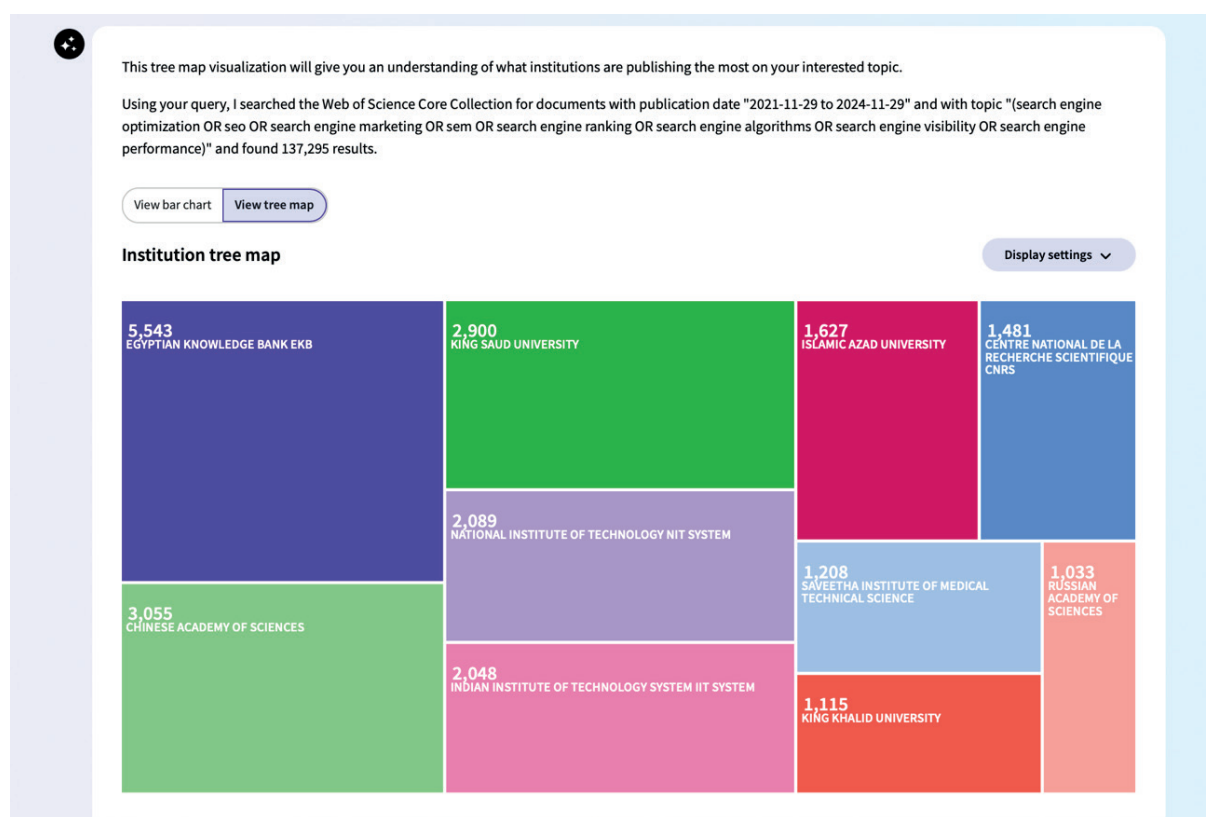


Figura 16

Tree map que muestra las instituciones con mayor volumen de publicaciones sobre SEO entre 2021 y 2024, liderado por el Egyptian Knowledge Bank.



A lo largo de este apartado hemos mostrado algunos ejemplos de lo que podemos hacer con el asistente de IA de Clarivate para la base de datos Web of Science. Aunque las posibilidades ya son muy amplias, es previsible que en el futuro los recursos que ofrece esta herramienta sigan mejorando y expandiéndose, a medida que evoluciona la tecnología de IA.

4. Conclusiones

Como hemos podido ver a lo largo de este apartado, Web of Science Research Assistant ofrece una experiencia completa e integrada que combina resúmenes y análisis de artículos, informes personalizados, visualizaciones interactivas, gestión y citación de fuentes, y feedback en tiempo real. Todo ello supone un complemento muy interesante para hacer investigación.

Se trata de una solución avanzada que puede acabar transformando la forma en la que los investigadores acceden, exploran y gestionan la información científica. Sin embargo, aunque este tipo de herramientas pueden ser muy útiles, no sustituyen el trabajo crítico del investigador. La selección de fuentes, el análisis profundo de los textos y la construcción de un marco teórico sólido siguen siendo tareas humanas.

Además, es importante tener en cuenta aspectos como la ética, la transparencia y el pensamiento crítico al usar inteligencia artificial en la investigación (Lopezosa, 2023b; Codina, 2025). No todo lo que la herramienta sugiere es siempre válido, ni toda la información que ofrece debe aceptarse sin revisión (Orduña-Malea y Cabezas-Clavijo, 2023; Font-Julia et al.,

2024). En definitiva, la existencia del asistente de Web of Science debe entenderse como una muy buena noticia, porque ayuda a los investigadores a hacer mejor su trabajo, con menos esfuerzo repetitivo y en menos tiempo. De modo que es una excelente ayuda para iniciar o avanzar en proyecto. La clave, como siempre en el uso de la IA es triple: (1) no delegar todo el trabajo en la IA, sino usarla para expandir y guiar, pero sin permitir que sustituya el trabajo del investigador; (2) verificar las aseveraciones, fuente y datos factuales o conceptuales y (3) ser transparente con su utilización. Usada de este modo, como un proceso de interacción crítica escalonada, puede ser un instrumento que modifique para siempre algunos modos de hacer ciencia.

5. Financiación

Este trabajo forma parte del proyecto “Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)”. Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

6. Referencias

Arroyo-Machado, W. (2024). Búsqueda y análisis bibliográfico con Elicit. *Infonomy*, 2(4). <https://doi.org/10.3145/infonomy.24.050>

Aguilera-Cora, E., Lopezosa, C., & Codina, L. (2024). *Scopus AI Beta Functional analysis and cases*. Barcelona: Universitat Pompeu Fabra. Departament de Comunicació. (Serie Editorial DigiDoc. DigiDoc Reports).

Aguilera-Cora, E., Lopezosa, C., Fernández-Cavia, J., & Codina, L. (2024). Accelerating research processes with Scopus AI: A place branding case study. *Revista Panamericana de Comunicación*, 6(1). <https://doi.org/10.21555/rpc.v6i1.3088>

Boté-Vericad, J.-J., Fabeiro, R., & Anglada, L. (2024). Creando un chatbot con ChatGPT como soporte a la catalogación en bibliotecas, archivos y centros de documentación. *Comparación de modelos de lenguaje en versión gratuita y premium*. <https://hdl.handle.net/2445/213600>

Boté-Vericad, J.-J. (2024). Copilot Microsoft y ChatGPT: Análisis de datos, dispersión, relación y correlaciones de variables [vídeo]. <https://hdl.handle.net/2445/213581>

Clarivate. (2025a). *Resources for Librarians and Administrators*. <https://clarivate.libguides.com/librarianresources/coverage>

Clarivate (2025b). *Web of Science Research Assistant*. <https://clarivate.com/academia-government/scientific-and-academic-research/research-discovery-and-referencing/web-of-science/web-of-science-research-assistant>

Codina, L. (2024). Inteligencia artificial generativa para trabajos académicos: método de análisis y comparación de Elicit, Epsilon, Perplexity, SciSpace y Scite. *Lluís Codina*. <https://www.lluiscodina.com/inteligencia-artificial-generativa-academia>

Codina, L. (2025). Interacción crítica escalonada con inteligencia artificial: explorando caminos para usar la IA en la academia. *Lluís Codina*. <https://www.lluiscodina.com/interaccion-critica-ia>

Font-Julián, C. I., Orduña-Malea, E., & Codina, L. (2024). ChatGPT Search as a tool for scholarly tasks: evolution or devolution?. *Infonomy*, 2(5). <https://doi.org/10.3145/infonomy.24.059>

Lopezosa, C. (2023a). Bing chat: hacia una nueva forma de entender las búsquedas. *Anuario ThinkEPI*, 17. <https://doi.org/10.3145/thinkepi.2023.e17a04>

Lopezosa, C. (2023b). La Inteligencia artificial generativa en la comunicación científica: retos y oportunidades. *Revista de Investigación e Innovación en Ciencias de la Salud*, 5(1), 1-5. <https://doi.org/10.46634/riics.211>

Orduña-Malea, E., & Cabezas-Clavijo, Á. (2023). ChatGPT and the potential growing of ghost bibliographic references. *Scientometrics*, 128(9), 5351-5355. <https://doi.org/10.1007/s11192-023-04804-4>

Ren, Y., Mohammadi, L., & Codina, L. (2025). Aplicación de sistemas de IA en un trabajo de final de máster universitario: presentación de un caso. Barcelona: Universitat Pompeu Fabra. Departamento de Comunicación. <http://hdl.handle.net/10230/70069>

Torres-Salinas, D., & Arroyo Machado, W. (2025). *Flujos de trabajo inteligentes con ChatGPT, Perplexity y NotebookLM*. <https://doi.org/10.5281/zenodo.14981646>

Torres-Salinas, D., Thelwall, M., & Arroyo Machado, W. (2024). ChatGPT for Bibliometrics: Potential applications and limitations. <https://hdl.handle.net/10481/92547>



CUVICOM

Curación, Visualización y Visibilidad en la Comunicación Digital

EPI