
Special issue

Rubén Alcaraz-Martínez
<https://orcid.org/0000-0002-7185-0227>
ralcaraz@ub.edu
Universitat de Barcelona

Mari Váñez
<https://orcid.org/0000-0002-3284-2590>
marivallez@ub.edu
Universitat de Barcelona

Carlos Lopezosa
<https://orcid.org/0000-0001-8619-2194>
lopezosa@ub.edu
Universitat de Barcelona

Recibido
5 de febrero de 2024
Aprobado
5 de marzo de 2024

© 2024
Communication & Society
ISSN 0214-0039
E ISSN 2386-7876
www.communication-society.com

2024 – Vol. 37(2)
pp. 279-291

Cómo citar este artículo:
Alcaraz-Martínez, R., Váñez, M. & Lopezosa, C. (2024). Informar sobre inteligencia artificial: el papel de los medios digitales de la Unión Europea, británicos y estadounidenses en la visibilidad de la IA generativa, *Communication & Society*, 37(2), 279-291.
doi.org/10.15581/003.37.2.279-291

Informar sobre inteligencia artificial: el papel de los medios digitales de la Unión Europea, británicos y estadounidenses en la visibilidad de la IA generativa

Resumen

La inteligencia artificial se ha convertido en uno de los temas del año 2023. Los medios de comunicación se han hecho eco de las tecnologías y cuestiones más relevantes relacionadas con este asunto. En un contexto de extrema competitividad entre medios digitales, el posicionamiento en buscadores (SEO por sus siglas en inglés) se erige como la principal estrategia con la que cuentan los cibermedios para conseguir la visibilidad necesaria para obtener más lectores. El objetivo de este trabajo es analizar la visibilidad de los contenidos sobre inteligencia artificial publicados por los medios de comunicación para una selección de palabras clave relacionadas, y cómo esta ha repercutido en la visibilidad tanto de estas tecnologías como de las cabeceras analizadas. Para ello, se han estudiado un total de 69 medios de 12 países de la Unión Europea, Estados Unidos y Reino Unido. Los resultados muestran un claro predominio de los medios de comunicación estadounidenses, seguidos por los españoles. En la mayoría de los países analizados se observa una distribución desigual entre medios, ostentando dos o tres de ellos el grueso de la visibilidad. Las consultas de búsqueda que han aportado más visibilidad a los medios analizados responden a una intención informacional, son de tipo *long-tail*, están relacionadas con las tecnologías de la empresa OpenAI, sobre todo con ChatGPT, y mayoritariamente publicadas en noticias ubicadas en las secciones dedicadas a ciencia y tecnología. Los resultados permiten constatar, por un lado, el interés creciente por el tema y, por otro lado, el buen trabajo de ciertos medios de comunicación en materia de SEO.

Palabras clave

Inteligencia Artificial, medios de comunicación digitales, posicionamiento en buscadores, IA, SEO, visibilidad, medios digitales.

1. Introducción

En el 2023, la inteligencia artificial generativa (en adelante, IAG) ha causado un gran impacto en la sociedad y esto continuará a corto y medio plazo (Llaneras *et al.*, 2023). Tras un año desde la irrupción de esta tecnología en todo tipo de sectores y contextos, los medios de comunicación parece que la han incorporado en la *agenda setting* como un tema relevante dentro de la comunicación sobre ciencia y tecnología.

La IAG es un tipo de inteligencia artificial capaz de generar textos y contenidos creativos a partir de la consolidación de datos de diversas fuentes (Dasborough, 2023). Entre las tecnologías de IAG, la que mayor grado de popularidad ha alcanzado ha sido el denominado Chat Generative Pre-trained Transformer (ChatGPT) de OpenAI, despertando el interés por esta tecnología entre multitud de sectores y profesionales. ChatGPT se enmarca en los conocidos como transformadores generativos preentrenados (*generative pre-trained transformer* o GPT), un tipo de modelo de lenguaje grande (*large language model* o LLM), basado en técnicas de aprendizaje profundo (*deep learning*) mediante ingentes cantidades de datos (Casella *et al.*, 2023). Las técnicas de aprendizaje profundo son un conjunto de algoritmos de aprendizaje automático (*machine learning*) basados en asimilar representaciones de datos, a partir de modelos estrechamente relacionados con teorías del desarrollo cerebral humano y, por tanto, que buscan simular la forma en que las personas desarrollamos habilidades (Franganillo, 2023).

Bing Chat es una funcionalidad integrada en el motor de búsqueda Bing, introducida por Microsoft en febrero de 2023. Bing Chat integra una personalización de la tecnología desarrollada por OpenAI orientada a la búsqueda (Mehdi, 2023) y entrenada para entender la intención de búsqueda, tono y contexto de los usuarios de Bing (Lopezosa, 2023). El *chatbot* ofrece resultados informacionales generados para satisfacer las consultas de los usuarios, como complemento a los resultados web. Además, destaca por incorporar, a diferencia de ChatGPT, las fuentes de información que utiliza para generar el contenido, así como una lista de preguntas relacionadas con la consulta planteada.

Google Bard, recientemente rebautizado como Gemini, es, como ChatGPT, un *chatbot* conversacional de inteligencia artificial, en este caso basado en el modelo de lenguaje grande PaLM 2 (*Pathways Language Model 2*), sucesor de LaMDA (*Language Model for Dialogue Applications*), lanzado por Google en 2017 (Aydin, 2023). Google Bard se lanza en abierto en marzo de 2023 como respuesta al éxito de ChatGPT y a la aparición de Bing Chat.

También a partir de 2022 comenzó a popularizarse la creación de imágenes a partir de *prompts* textuales mediante herramientas como Dall-e (también de OpenAI) o Midjourney. De acuerdo con Borji (2022), Midjourney destaca por su capacidad para generar imágenes muy relevantes de acuerdo con los *prompts* utilizados, así como por su capacidad para integrar estilos y elementos de diseño. Por su parte, Dall-e destaca por su capacidad para generar imágenes a partir de la combinación de diversos elementos y conceptos. Midjourney y Dall-e son las dos aplicaciones más populares para la generación de imágenes con IA, hasta el punto de que medios de comunicación como *The Economist* o las revistas *Cosmopolitan* y *The Atlantic*, ya las han utilizado para ilustrar sus portadas (Franganillo, 2023).

Por otro lado, Internet se ha consolidado como uno de los escenarios predominantes para el acceso a la información. Ello ha provocado que los buscadores generalistas, principalmente Google, se hayan convertido en fuente de información de referencia para muchos usuarios.

Asimismo, por lo general estos usuarios suelen consumir únicamente aquellos resultados que aparecen en la primera página de las SERPs (*Search Engine Results Page*) de los buscadores, obviando aquellos que se presentan en las siguientes páginas (Kulshrestha *et al.*, 2018) e incluso aquellos que aparecen en las posiciones más bajas del primer bloque de diez resultados, repartiéndose más de la mitad del porcentaje de clics, los tres primeros resultados orgánicos (28,5 %, 15,7 % y 11 %, respectivamente) (Beus, 2020).

Ante esta circunstancia, la importancia de los buscadores resulta esencial; de hecho, el 96,55 % de las páginas de internet no obtienen tráfico de búsqueda orgánico de Google (Soulo, 2023), por lo que menos del 5 % aprovechan en su favor la ingente cantidad de tráfico que pueden aportar los buscadores a un sitio web.

La disciplina encargada de analizar los factores que afectan a los resultados de búsqueda se denomina posicionamiento en buscadores o *Search Engine Optimization* (en adelante, SEO). Implementar una estrategia de SEO resulta fundamental para el éxito de un sitio web, ya que facilita el posicionamiento de sus páginas en lugares destacados de la SERP en respuesta a las

consultas de los usuarios. Esto, a su vez, impulsa el tráfico hacia el sitio web y aumenta la visibilidad de sus contenidos (Duong, 2019).

Si bien es cierto que el SEO es un tema central en el ámbito empresarial porque tiene una repercusión comercial directa, también comienza a ser un tema del que se ha empezado a preocupar la academia. En este sentido, han surgido en los últimos años numerosas investigaciones centradas en la visibilidad web y, más concretamente, en el análisis de las consultas formuladas en los buscadores, o en el efecto de los resultados de búsqueda en las decisiones del consumidor, por poner un par de ejemplos representativos.

Todo ello se ha estudiado desde distintas esferas, incluyendo el sector de los medios de comunicación y su visibilidad. No obstante, a día de hoy no hay trabajos específicos centrados en el estudio de la visibilidad web de los contenidos relacionados con la inteligencia artificial publicados por los medios de comunicación.

Ante esta necesidad nace este estudio que tiene como objetivo principal el análisis de la visibilidad en los resultados orgánicos de los contenidos sobre la inteligencia artificial publicados por los medios de comunicación de prensa escrita, cuáles han sido las cabeceras que han conseguido una mayor visibilidad para una selección de palabras clave relacionadas, y cómo esta ha repercutido en la visibilidad de las marcas comerciales y de las tecnologías que parecen ser sus principales protagonistas: OpenAI, Google, Microsoft y Midjourney. Con los datos obtenidos se espera lograr una visión de la situación en ámbito europeo, a la vez que se busca constatar si existen diferencias destacadas entre la situación en Reino Unido y Estados Unidos.

Los objetivos específicos que se han definido en esta investigación son los siguientes:

- O1. Determinar la visibilidad web del concepto 'inteligencia artificial' y sus principales palabras clave relacionadas en los sitios web de los principales medios de comunicación de la Unión Europea, británicos y norteamericanos.
- O2. Identificar qué medios de comunicación son más visibles cuando se realiza una consulta con las palabras clave: inteligencia artificial, inteligencia artificial generativa, ChatGPT, GPT-4, GPT-3, OpenAI, Google Bard, Bing Chat, Dall-e y Midjourney.
- O3. Comprobar si existen diferencias destacadas entre la visibilidad web de los distintos medios de comunicación de la Unión Europea y, posteriormente, frente a los medios británicos y estadounidenses.

En la siguiente sección se presenta una revisión de la literatura centrada en los principios de la visibilidad web y el SEO. Seguidamente, se describe la metodología y la herramienta utilizada para obtener los datos objeto de estudio de cada uno de los medios de comunicación seleccionados. A continuación, se presentan los resultados y se discuten. Por último, se desarrollan las conclusiones, las limitaciones y las líneas de investigación futuras.

2. Marco teórico

Los estudios académicos acerca de la visibilidad web se centran mayormente en el análisis de sectores empresariales específicos. En líneas generales, se han realizado investigaciones relacionadas con el posicionamiento en buscadores en diversos contextos tales como universidades (Nevado-Chiné *et al.*, 2021; Vállez *et al.*, 2022), portales turísticos en la web (Rovira *et al.*, 2010; Pedraza-Jiménez, 2018), bibliotecas (Onaifo & Rasmussen, 2013; Vállez & Ventura, 2020), cibermedios (Giomelakis & Veglis, 2015; Lopezosa *et al.*, 2019; Pedrosa & Morais, 2021), entre otras industrias (Mladenović *et al.*, 2023).

Asimismo, se observa una prevalencia de estudios cuantitativos centrados en herramientas de auditoría SEO, especialmente aquellas vinculadas al uso de productos como SEMrush (Lopezosa & Codina, 2018), Ahrefs (García-Carretero *et al.*, 2016), Sistrix (Vállez & Ventura, 2020) y Majestic (Orduña, 2021).

En cuanto a los temas más investigados, se observa un crecimiento en los estudios sobre la visibilidad web de cuestiones sociales. Algunos trabajos indican que el uso de motores de búsqueda como Google, Yandex o Baidu desempeña un papel fundamental en la búsqueda de

información durante los periodos electorales, lo que conlleva importantes implicaciones para el consumo de información (Metaxas & Pruksachatkun, 2016; Trielli & Diakopoulos, 2020). También se destaca que, con frecuencia, los primeros resultados provienen de medios de comunicación nacionales (Unkel & Haim, 2019), lo que desempeña un papel crucial en la imagen de los candidatos políticos (Belt *et al.*, 2012).

En este sentido, y tomando como eje central el periodismo, la visibilidad web juega un papel significativo para este sector, especialmente en el contexto de los buscadores generalistas y las redes sociales (Noppari *et al.*, 2014). El SEO ha transformado el periodismo y ha introducido nuevos desafíos y oportunidades para los profesionales del sector, así como para las agencias de noticias (Neuberger, 2014) que deben plantear estrategias para optimizar su contenido y conseguir con ello más lectores.

Por otro lado, otro conjunto de trabajos no menos importantes se ha centrado en estudiar los principales pilares del SEO. Antes de abordarlos, cabe destacar que los algoritmos de Google –y del resto de buscadores– están en constante evolución y, con ello, el SEO y sus técnicas, por lo que investigaciones publicadas hasta la fecha, que hoy son muy significativas, podrían quedar obsoletas a medio y largo plazo. Tomando en cuenta esta consideración, podemos confirmar que en cuanto a técnicas y factores de posicionamiento SEO, destacan, en los últimos años, las investigaciones de Shahzad *et al.* (2020) y Mittal *et al.* (2018), que se centran en el estudio de las tres técnicas más comunes en materia de SEO (*White Hat*, *Grey Hat* y *Black Hat*), demostrando que aplicar técnicas éticas (*White Hat*) de posicionamiento es la estrategia más segura y duradera. El trabajo de Umenhofer (2019) aborda el SEO desde la perspectiva del aumento del tráfico web y señala que el uso de la etiqueta *title* del estándar HTML y ciertas metaetiquetas, junto con una estrategia de identificación de palabras clave, permite lograr un mayor impacto que otras técnicas de SEO. El análisis de Windia *et al.* (2018) destaca que las mejores técnicas SEO incluyen la optimización de la metadescripción, el uso de palabras clave de enfoque único y la aplicación del atributo *alt* en las imágenes. La propuesta de Yudasubrata *et al.* (2019) se enfoca en los métodos para identificar palabras clave, destacando la orientación geográfica como un elemento diferenciador.

En cuanto a los estudios académicos sobre SEO en medios de comunicación, estos presentan una larga trayectoria que se inicia en 2011. En conjunto, estos estudios ofrecen una panorámica completa de cómo el SEO moldea la producción y difusión de noticias en línea, desde prácticas específicas en redacciones, hasta la preparación de futuros profesionales y la importancia estratégica en cibermedios, proporcionando así un marco teórico sólido para comprender la evolución del periodismo en el contexto de la visibilidad web.

Concretamente, las líneas más destacadas sobre SEO y medios de comunicación se han enfocado, hasta el momento, en la implementación del posicionamiento en buscadores en la producción de noticias (Dick, 2011), en el impacto del SEO en el tráfico de medios de comunicación en línea (Giomelakis & Veglis, 2015), en la importancia estratégica del SEO en medios digitales (Iglesias-García & Codina, 2016), en herramientas e indicadores SEO para el análisis del tráfico web de medios de comunicación (García-Carretero *et al.*, 2016), y en el SEO como otra rutina periodística más para la producción de noticias (Prawira & Rizkiansyah, 2018).

En definitiva, dado el papel significativo de los medios de comunicación como promotores de contenidos de toda índole, su influencia en la sociedad y la opinión pública, y la importancia que tiene el posicionamiento de las noticias en su lucha por ser relevantes en un mercado extremadamente competitivo, se considera pertinente llevar a cabo un estudio específico centrado en los objetivos de este trabajo.

3. Metodología

Se realiza un análisis de la visibilidad web de los contenidos sobre inteligencia artificial publicados en los medios de comunicación de prensa escrita. Se aplica un diseño cuantitativo basado en la recopilación de datos de la visibilidad de 69 medios de comunicación de 14 países.

Para la selección de los países se han usado datos facilitados por Eurostat¹ para elegir aquellos de la UE con mayor población, concretamente los que superan los 10 millones de habitantes, en total doce: Alemania, Bélgica, Chequia, España, Francia, Grecia, Italia, Países Bajos, Polonia, Portugal, Rumanía y Suecia. No obstante, se ha considerado interesante tener una visión de lo que estaba sucediendo en lengua inglesa, por ello se han incorporado los EE. UU. y el Reino Unido.

Para la selección de los medios se han utilizado los datos que proporciona SCImago Media Rankings (SMR)². Es una base de datos de medios de comunicación de todo el mundo que analiza la reputación web de cada medio. Se seleccionaron los cinco primeros medios posicionados para cada país seleccionado. En total se han obtenido datos de 69 medios de comunicación, porque uno ha sido descartado al no tratarse realmente de un medio sino de un portal generalista, y otro medio del que no se han podido obtener datos por su pobre rendimiento a nivel de visibilidad para las palabras clave objeto de estudio. Los datos sobre la visibilidad web de los medios de comunicación se han obtenido utilizando SEMrush, una herramienta que proporciona servicios de investigación competitiva de sitios web en materia de SEO. Se trata de una compañía de reconocido prestigio internacional, con más de 10 millones de clientes. SEMrush utiliza tecnología del *big data* y cuenta con una de las bases de datos más grande del mercado que contempla 25 mil millones de palabras clave y 17 mil millones de URLs rastreadas cada día. Con esta aplicación se han monitorizado los dominios de los 69 medios de comunicación analizados con el objetivo de estudiar cómo se posicionan para diferentes palabras clave vinculadas con la inteligencia artificial: inteligencia artificial, IA generativa, ChatGPT, GPT-4, GPT-3, OpenAI, Google Bard, Bing Chat, Dall-e y Midjourney, así como su correspondencia en cada idioma para las que necesitan traducción (Tabla 1). En el conjunto de datos publicado en CORA y vinculado a esta investigación se recogen todos los medios analizados y los datos básicos recogidos con SEMrush³.

Tabla 1. Traducción de la palabra clave utilizada en los diferentes idiomas.

País	Palabras clave	
Bélgica	Kunstmatige Intelligentie	Generatieve Kunstmatige Intelligentie
Chequia	Umělá inteligence	Generativní Umělá inteligence
Francia	Intelligence Artificielle	Intelligence Artificielle Générative
Alemania	Künstliche Intelligenz	Generative Künstliche Intelligenz
Grecia	Τεχνητή Νοημοσύνη	Γεννητική Τεχνητή Νοημοσύνη
Italia	Intelligenza Artificiale	Intelligenza Artificiale Generativa
Países Bajos	Kunstmatige Intelligentie	Generatieve Kunstmatige Intelligentie
Polonia	Sztuczna Inteligencja	Generatywna Sztuczna Inteligencja
Portugal	Inteligência Artificial	Inteligência Artificial Generativa
Rumanía	Inteligentă Artificială	Inteligentă Artificială Generativă
España	Inteligencia Artificial	Inteligencia Artificial Generativa
Suecia	Artificiell Intelligens	Generativ Artificiell Intelligens
Reino Unido y EE. UU.	Artificial Intelligence	Generative Artificial Intelligence

Fuente: elaboración propia.

¹ Eurostat. Population el 1 de enero.

https://ec.europa.eu/eurostat/databrowser/view/tps00001/default/table?lang=en&category=t_demo.t_demo_pop

² <https://www.scimagomedia.com/index.php>

³ <https://doi.org/10.34810/data1070>

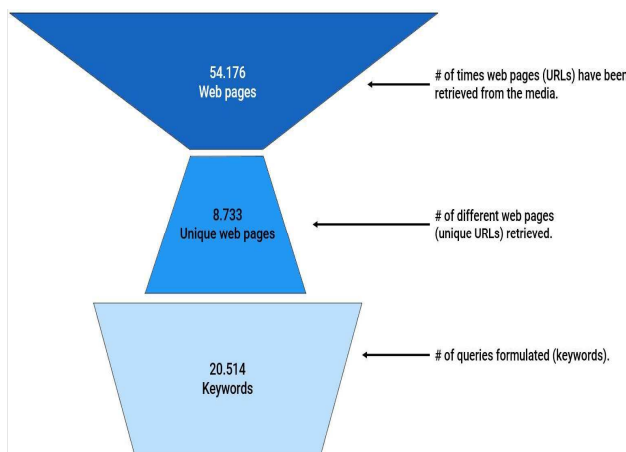
Para la obtención de la información relativa a cada país se utilizaron los términos en su propio idioma y se consultaron los resultados de posicionamiento en el buscador Google con el ccTLD (*Country Code Top-Level Domain*) correspondiente a través de la herramienta *Domain overview > Organic research* de SEMrush, que toma en consideración todas aquellas palabras clave que han generado impresiones de resultados (*snippets* de cualquier tipo) para el dominio analizado entre los 100 primeros resultados mostrados por el buscador para cada búsqueda. La extracción de datos se limitó a los resultados obtenidos por cada medio en su área geográfica (p. ej., *El País* en la versión española de Google o *Le Monde* en la versión francesa del buscador) mediante las opciones que ofrece la herramienta Semrush para limitar la consulta a un área geográfica específica. Junto con las palabras clave obtenidas, se extrajeron diferentes métricas y datos relacionados. Concretamente, la posición que corresponde al puesto obtenido en el *ranking* por el resultado que ha disparado la palabra clave; el volumen de búsqueda (SV) que corresponde al número de consultas mensuales estimado para la palabra clave; la dificultad de la palabra clave (KD) que se valora entre 0 y 100 para predecir lo difícil que es posicionar un contenido para esa palabra clave; las características de la SERP (PT), es decir, los tipos de *snippets* o bloques de resultados que la palabra clave es capaz de “disparar” (*snippet* orgánico, historias destacadas, recetas, carrusel de vídeos, bloque “más preguntas”, entre otros); y el tipo de intención de búsqueda (KI) que puede ser informacional, comercial, transaccional o navegacional. Se registraron las diferentes consultas formuladas por los usuarios, aplicando de este modo técnicas de expansión de consultas (Efthimiadis, 1996).

Este proceso se repitió para todos los medios estudiados, con lo que se creó un extenso corpus que se puede consultar en el conjunto de datos mencionado anteriormente. Los datos se recogieron en diciembre del 2023 y se corresponden con el *ranking* de octubre de 2023.

4. Resultados

Se muestran los resultados obtenidos de la visibilidad web de los medios de comunicación analizados en relación con los términos asociados a la inteligencia artificial indicados en la sección metodología. En primer lugar, el conjunto de datos recogidos con la herramienta SEMrush y que ha servido de base para el análisis y estudio posterior. Se han obtenido las páginas web de los diferentes medios de comunicación analizados que han aparecido como resultado en el buscador en las 100 primeras posiciones de las SERPs (Figura 1). Los datos obtenidos son: número de páginas web recuperadas, número de páginas web únicas, y número de consultas realizadas (*Keywords*) en cuyos resultados aparece una página web de los medios analizados.

Figura 1. Datos recogidos sobre la visibilidad web de los conceptos vinculados con la inteligencia artificial en los dominios de los medios.

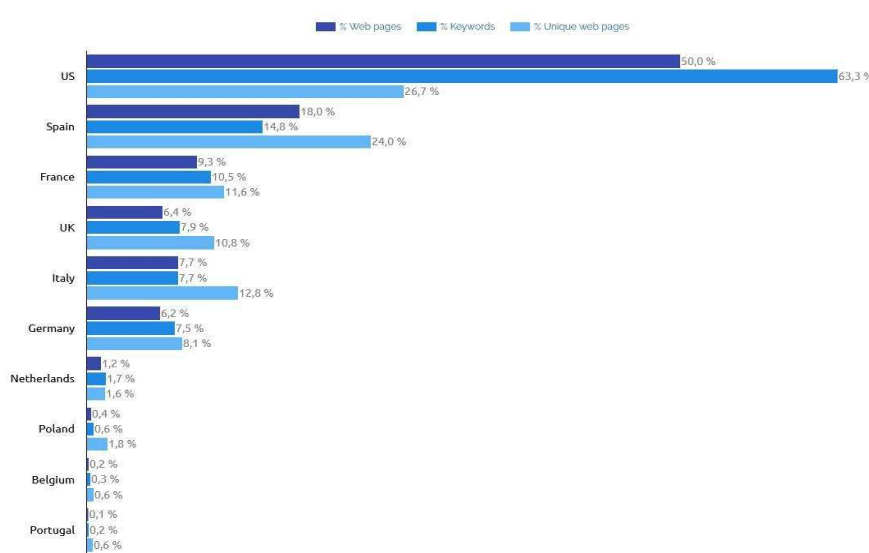


Fuente: elaboración propia a partir de datos de Semrush.

Tal como se aprecia en la figura anterior, se trata de una imagen con un formato parecido a un embudo que va de lo más general a lo más específico: total de páginas web recuperadas, páginas web únicas (diferentes) y el número de consultas formuladas por los usuarios en el buscador (*keywords*) para acceder a los contenidos sobre inteligencia artificial. Como se observa también en el gráfico, el número total de palabras clave utilizadas es mayor que el número de páginas web únicas posicionadas, lo que muestra la capacidad de esas páginas para posicionarse para diferentes combinaciones de palabras clave.

Al analizar la distribución por países, destaca un claro predominio de los medios estadounidenses en cuanto a la visibilidad de sus contenidos, como se observa en la Figura 2. En conjunto, estos medios ocupan el primer lugar tanto en páginas web, páginas web únicas, así como en palabras clave posicionadas, superando a los demás medios. Entre el primer (EE. UU.) y el segundo país (España), se observa una diferencia del 178,2 % en cuanto a páginas posicionadas, del 326,6 % en la cantidad de palabras clave y del 11,6 % en cuanto a páginas únicas posicionadas.

Figura 2. Distribución por países de la visibilidad web de los medios de comunicación.



Fuente: elaboración propia a partir de datos de Semrush.

A continuación, se analizan los tipos de búsquedas formuladas. La mayor parte de las consultas son de tipo informacional, concretamente el 68,8 % de las palabras clave. Le siguen las relacionadas con la investigación comercial (*commercial*) (26,1 %), las navegacionales (8,5 %) y, por último, las transaccionales (1,8 %)⁴. En referencia al tipo de resultado que se muestra en la SERP, el grueso del tráfico sobre el total de las páginas analizadas procede de resultados orgánicos (96,4 %), seguido por los bloques de tipo *People also ask* (2,7 %), historias destacadas (0,8 %) y los carruseles de vídeo (0,1 %).

Otro elemento que se ha analizado es el nivel de especificidad de las palabras clave, es decir, si se pueden considerar *short-tail* (formadas por uno o dos términos), *mid-tail* (formadas por tres o cuatro términos) o *long-tail* (formadas por más de cuatro términos). En este sentido, las más frecuentes son las consultas formuladas utilizando tres o más términos para confeccionar las palabras clave (38,4 %), aunque la diferencia no es excesivamente grande (11,8 %) respecto a las palabras de tipo *short-tail* que representan el 26,6 % sobre el total. Finalmente, las palabras clave *mid-tail* suponen el 34,9 %.

⁴ SEMrush clasifica algunas palabras clave bajo más de una de las categorías anteriores, razón por la cual la suma de los porcentajes es superior al 100 %.

Si atendemos al resultado global, los diez medios internacionales que mayor visibilidad ostentan para las palabras clave objeto de estudio son, de mayor a menor, *The New York Times*, *The Washington Post*, *The Wall Street Journal*, AP News, *La Vanguardia*, *El Español*, *El País*, *Le Monde*, *Le Figaro* y *El Mundo*. Entre los diez primeros, encontramos 4 medios estadounidenses liderando la clasificación del ranking, seguidos por 4 medios españoles y 2 franceses. La Figura 3 presenta los datos de los seis países mejor posicionados, ordenados por el número total de páginas web que ocupan en los primeros 100 resultados de la SERP.

Figura 3. Distribución de la visibilidad por los medios de comunicación analizados de los seis países con más visibilidad (EE. UU., España, Francia, Reino Unido, Italia y Alemania).

News outlet	Web pages	Keywords	Unique web pages	News outlet	Web pages	Keywords	Unique web pages
nytimes.com	10.764	10.764	713	lavanguardia.com	2.953	2.406	590
washingtonpost.com	6.748	6.748	515	elespanol.com	2.399	2.068	521
wsj.com	5.004	5.004	711	elpais.com	2.083	1.782	494
apnews.com	2.598	2.598	220	elmundo.es	1.285	1.164	232
usatoday.com	2.021	2.021	186	abc.es	1.034	956	264
News outlet	Web pages	Keywords	Unique web pages	News outlet	Web pages	Keywords	Unique web pages
lemonde.fr	1.687	1.453	312	theguardian.com	1.251	1.053	298
lefigaro.fr	1.277	1.178	302	reuters.com	986	882	256
ouest	1.129	1.054	239	independent.co.uk	616	584	174
leparisien.fr	933	857	164	dailymail.co.uk	368	356	129
afp.com	221	221	31	telegraph.co.uk	246	241	87
News outlet	Web pages	Keywords	Unique web pages	News outlet	Web pages	Keywords	Unique web pages
repubblica.it	1.547	1.300	325	zeit.de	1.136	972	236
corriere.it	1.316	1.183	349	spiegel.de	1.039	953	183
ansa.it	725	668	250	welt.de	494	472	128
lastampa.it	385	364	125	bild.de	392	358	81
ilmessaggero.it	223	218	72	focus.de	318	305	83

Fuente: elaboración propia a partir de datos de Semrush.

Al analizar los datos generales de estos seis países, podemos observar el dominio de los medios estadounidenses. De esta manera, los diarios de Estados Unidos posicionan 721 páginas web únicas en la primera página de la SERP (entre las posiciones 1 y 10), lo que representa un 30,1 % del total de páginas que ocupan estas primeras posiciones. Le siguen los españoles (558 páginas web únicas y 23,3 % sobre el total), los italianos (329 páginas web únicas y 13,7 % sobre el total), los franceses (254 páginas web únicas y 10,6 % sobre el total), los británicos (227 páginas web únicas y 9,5 % sobre el total) y, finalmente, los alemanes (190 páginas web únicas y 7,9 % sobre el total).

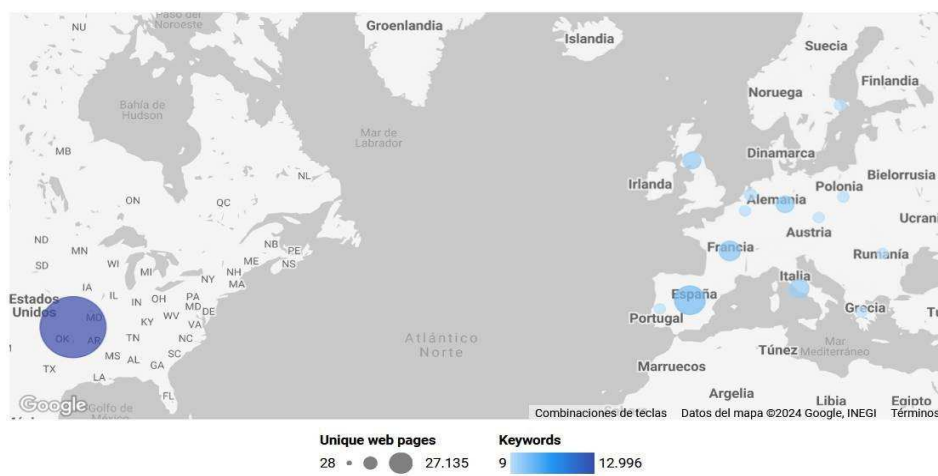
Para realizar una primera aproximación al tipo de contenidos que se posicionan de cada medio, se analizan las URLs de las páginas web recuperadas. Se han obtenido los subdirectorios (en la mayoría de los casos, secciones del medio), lo que permite ver que las secciones dedicadas a la ciencia y tecnología son las que acaparan el grueso de la visibilidad para las palabras clave analizadas, seguidas por otras como sociedad, economía y cultura.

En algunos pocos casos, se observa también como esa visibilidad se consigue a través de otros sitios web vinculados al medio, disponibles en un subdirectorio o subdominio, pero de naturaleza independiente. Este es el caso, por ejemplo, del blog *Andro4all* que aporta 82 URLs a la *La Vanguardia*; El blog *Pixels* de *Le Monde*, con 35 URLs; o los medios *Vandal*, *El Androide libre* y *Omicron* vinculados a *El Español*, los cuales aportan 28, 26 y 12 URLs, respectivamente.

Si se analizan el tipo de palabras clave utilizadas para las consultas relacionadas con los nombres comerciales de los servicios basados en tecnologías vinculadas a la inteligencia artificial, los que han devuelto una mayor cantidad de palabras clave y de páginas posicionadas para estos términos son, de mayor a menor, ChatGPT (22 574 palabras clave), Midjourney (2 359 palabras clave), Google Bard (1 013 palabras clave), Bing Chat (976 palabras clave), Dall-e (945 palabras clave), GPT-4 (794 palabras clave) y GPT-3 (742 palabras clave). Si sumamos las tres variantes de GPT, el resultado incrementa aún más la distancia entre la visibilidad de esta tecnología y el resto (24 110 palabras clave).

Por último, se ha recogido en la Figura 4 un mapa de la distribución por países de la visibilidad estudiada desde el punto de vista de dos variables: el número de páginas web únicas (identificadas por el tamaño de la marca) y el número de palabras clave que posicionan (identificadas por la intensidad del color).

Figura 4. Distribución geográfica del número de páginas web únicas y de las palabras clave utilizadas.



Fuente: elaboración propia a partir de datos de Semrush.

El 16,19 % de las palabras clave posicionadas por los medios analizados muestran resultados de búsqueda entre las diez primeras posiciones del *ranking*. En el caso de España, la cifra asciende hasta el 25,19 %, por encima de otros países como Italia (22,88 %), Alemania (15,18 %), Francia (14 %), el Reino Unido (13,04 %) o los EE. UU. (12,88 %). Algunos de los conceptos que más se repiten entre las palabras clave posicionadas entre los diez primeros resultados para los medios españoles son, al margen de las estudiadas en esta investigación, “gratis” (6,92 %), “chat” (6,3 %), “app” (4,43 %), “imágenes” (4,19 %), “fotos” (2,68 %), “preguntas” (1,91 %), “noticias” (1,75 %), “Elon Musk” (1,63 %), “peligros” (1,3 %) y “dibujos” (1,26 %). Esto denota un importante interés por la capacidad de las IA para generar contenido visual de diversa naturaleza (8,13 %).

5. Discusión y conclusión

Como se ha comentado en la introducción, no hay dudas sobre la relevancia del SEO para la visibilidad de los sitios web, especialmente para los sitios basados en contenidos (Charlton *et al.*, 2008). En el contexto español, donde los medios de comunicación impresos consolidados experimentan un declive y nuevos medios nativos digitales ganan terreno en audiencia (Vara *et al.*, 2023), el SEO se convierte, para estas empresas, en una práctica esencial para continuar manteniendo su relevancia como actores de peso en un sector en el que los medios de noticias nativos digitales obtienen los mejores resultados entre la franja de edad de los lectores del futuro (18 a 24 años) (Negredo & Kaufmann, 2021).

Los resultados obtenidos revelan un claro dominio de los medios estadounidenses por lo que respecta a la visibilidad web para las consultas formuladas en inglés, para conceptos relacionados con la IA, superando significativamente a otros países. España se sitúa al frente de Europa, por delante del resto de los países analizados del continente, incluido el Reino Unido. Aunque el éxito de los medios de estos dos países se deba también en buena medida a las mayores comunidades geográficas y lingüísticas que ostentan, no es menos cierto que la supremacía de los medios estadounidenses resalta la influencia global de los medios de este país en la difusión de la información sobre la IA a nivel global, específicamente en el contexto lingüístico angloparlante.

Si atendemos a los tipos de búsquedas más habituales, destacan las informativas, algo lógico teniendo en cuenta la naturaleza de los sitios web analizados. Esto también denota el gran interés del público a lo largo de 2023 por obtener información sobre la IA y sus tecnologías relacionadas, siendo los medios de comunicación un posible referente en este sentido.

El análisis de las URLs permite observar cómo las secciones dedicadas al tema “ciencia y tecnología” y similares son las más prominentes. Esto da buena cuenta del enfoque específico con el que la IA se ha tratado por los medios analizados. En este sentido, a medida que su uso se vaya generalizando, se presenten nuevos retos relacionados, barreras en el acceso, así como posibles problemas y conflictos sociales derivados de su uso, es de esperar que los medios se hagan eco de noticias que quizá se podrían encajar en secciones como “Cultura”, “Pensamiento” o “Sociedad”, entre otras.

Para concluir, se realiza un repaso de los objetivos definidos, a la vez que se identifican las limitaciones del estudio y las futuras líneas de investigación.

Respecto al primer objetivo, los resultados de este estudio muestran un interés por parte de todos los medios de comunicación analizados, menos uno, en las noticias relacionadas con la inteligencia artificial. El concepto ‘inteligencia artificial’ es el que aporta una mayor cantidad de tráfico, pues también es sobre el que se han generado una mayor cantidad de contenidos.

En relación con el segundo objetivo, se confirma que la tecnología sobre la que más se ha escrito, así como la que más visibilidad ha dado a los medios analizados durante el periodo objeto de estudio, es ChatGPT. Su lanzamiento en abierto en noviembre de 2022, así como su enorme popularización a lo largo de 2023, ha quedado reflejada en los resultados obtenidos, situando este *chatbot* casi como un sinónimo de inteligencia artificial en el imaginario colectivo.

Finalmente, por lo que respecta al tercer objetivo, de entre los países de la Unión Europea analizados, los medios españoles en su conjunto son los que han generado más contenido y han obtenido también mayor visibilidad para un mayor conjunto de palabras clave. Francia, Italia y Alemania, en este orden, son los tres siguientes países europeos con medios con una mayor visibilidad. En los tres países se observa una distribución desigual entre medios. En el caso de España, *La Vanguardia* (30,27 %) y *El Español* (24,59 %) alcanzan más de la mitad de la visibilidad conseguida por el conjunto de medios españoles analizados, seguidos por *El País* (21,35 %), *El Mundo* (13,17 %) y *ABC* (10,6 %). Esta situación se pone de relieve aún más en Alemania, donde *Zeit* supone por sí solo el 36,54 % de las palabras clave posicionadas; y en Italia, en la que *Il Corriere della Sera* (31,36 %) y *La Repubblica* (36,87 %) aportan el 68,23 % de los resultados a su país. La situación descrita se da también en EE. UU. y Reino Unido. En el primer caso, el *New York Times* (39,67 %) y el *Washington Post* (24,87 %) suponen el 64,54 % de la visibilidad. En el caso británico, *The Guardian* (36,08 %) y la agencia Reuters (28,44 %), el 64,52 %. Si hablamos de cifras globales, tal y como se ha comentado en los resultados y la discusión, los medios estadounidenses son los que consiguieron la mayor visibilidad, situándose incluso los dos medios con peores métricas (*AP news* y *USA Today*) por encima de todos los medios del resto de los países analizados, con la excepción de *La Vanguardia*, que supera a ambos, y de *El Español* y *El País*, que superan a *USA Today*, pero no a *AP News*.

Una de las limitaciones principales que se pueden atribuir al estudio es que el análisis de la visibilidad tan solo aporta información acerca las impresiones obtenidas por los medios de

comunicación en Google Search; esto es, la cantidad de veces que han conseguido aparecer en la SERP de Google para unas determinadas palabras clave y su posición en el *ranking*. No obstante, la foto completa para comprender el valor de esta posición en el *ranking* y del tráfico potencial que ha podido aportar a estos medios solo está al alcance de cada uno de ellos a través de sus herramientas de analítica. Ese estudio pormenorizado puede aportar información sobre la calidad de esas visitas (porcentaje de rebote, de interacciones, tiempo de permanencia en página, páginas visitadas por sesión, entre otras).

Finalmente, este trabajo puede ser la base para futuras líneas de investigación centradas en el análisis del impacto de la IA en otros tipos de sitios web de contenidos, sectores productivos, educativos, entre otros. También se podría ampliar el alcance de la investigación hacia otros entornos geográficos, o centrarse en un análisis más pormenorizado o cualitativo de los tipos de contenidos que se publican.

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)”. PID2021-123579OB-I00 (MICINN), Ministerio de Ciencia e Innovación (España).

Referencias

- Ambarsari, E. W., Yulianingsih, Prasetya, R., Cholifah, W. N. & Rahim, R. (2018). Applying C-FDT as making decision for the content of SEO media online. *Journal of physics: conference series*, 1363, 1-7. <https://www.doi.org/10.1088/1742-6596/1363/1/012059>
- Aydın, Ö. (2023). Google Bard generated literature review: metaverse. *Journal of AI*, 7(1), 1-14.
- Belt, T. L., Just, M. R. & Crigler, A. N. (2012). The 2008 media primary: handicapping the candidates in newspapers, on TV, cable, and the Internet. *The international journal of press/politics*, 17(3), 341-369. <https://www.doi.org/10.1177/1940161212444124>
- Beus, J. (2020). *Why (almost) everything you knew about Google CTR is no longer valid*. Retrieved from <https://www.sistrix.com/blog/why-almost-everything-you-knew-about-google-ctr-is-no-longer-valid>
- Borji, A. (2022). Generated faces in the wild: quantitative comparison of stable diffusion, Midjourney and DALL-e 2. *ArXiv*, Cornell University. <https://www.doi.org/10.48550/arXiv.2210.00586>
- Cascella, M., Montomoli, J., Bellini, V. & Bignami, E. (2023), Evaluating the feasibility of ChatGPT in healthcare: An analysis of multiple clinical and research scenarios. *Journal of medical systems*, 47(1), 1-5. <https://www.doi.org/10.1007/s10916-023-01925-4>
- Charlton, C., Blank, R., Otuome, H. & Gonzalez, O. (2008). *AdvancED Flex Application Development*. New York: Apress.
- Dasborough, M. T. (2023). Awe-inspiring advancements in AI: the impact of ChatGPT on the field of organizational behavior. *Journal of organizational behavior*, 44(2), 177-179. <https://www.doi.org/10.1002/job.2695>
- Dick, M. (2011). Search engine optimisation in UK news production. *Journalism practice*, 5(4), 462-477. <https://www.doi.org/10.1080/17512786.2010.551020>
- Duong, V. (2019). *SEO management: methods and techniques to achieve success*. London: Wiley & Sons.
- Franganillo, J. (2023). La inteligencia artificial generativa y su impacto en la creación de contenidos mediáticos. *methaodos. revista de ciencias sociales*, 11(2), m231102a10. <https://www.doi.org/10.17502/mrcs.v11i2.710>
- García-Carretero, L., Codina, L., Díaz-Noci, J. & Iglesias-García, M. (2016). Herramientas e indicadores SEO: características y aplicación para análisis de cibermedios. *El Profesional de la Información*, 25(3), 497-504. <https://www.doi.org/10.3145/epi.2016.may.19>

- Giomelakis, D. & Veglis, A. (2015). Investigating search engine optimization factors in media websites, the case of Greece. *Digital Journalism*, 4(3), 379-400.
<https://www.doi.org/10.1080/21670811.2015.1046992>
- Iglesias-García, M. & Codina, L. (2016). Los cibermedios y la importancia estratégica del posicionamiento en buscadores (SEO). *Opción*, 32(9), 929-944. Retrieved from <http://produccioncientificaluz.org/index.php/opcion/article/view/21784>
- Kulshrestha, J., Eslami, M., Messias, J., Bilal Zafar, M., Ghosh, S., Gummadi, K. P. & Karahalios, K. (2018). Search bias quantification: investigating political bias in social media and web search. *Information retrieval journal*, 22(1), 188-227. <https://www.doi.org/10.1007/s10791-018-9341-2>
- Llaneras, K., Rizzi, A. & Álvarez, J. A. (2023). ChatGPT es solo el principio: la inteligencia artificial se lanza a reorganizar el mundo. *El País*. Retrieved from <https://elpais.com/sociedad/2023-01-29/chatgpt-es-solo-el-principio-la-inteligencia-artificial-se-lanza-a-reorganizar-el-mundo.html>
- Lopezosa, C. & Codina, L. (2018). *Análisis de posicionamiento en medios de comunicación con herramientas SEO: cobertura informativa de los premios Oscar 2017*. Barcelona: Universitat Pompeu Fabra. Retrieved from <http://hdl.handle.net/10230/33632>
- 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. Retrieved from <http://hdl.handle.net/10230/36234>
- Lopezosa, C. (2023). Bing chat: hacia una nueva forma de entender las búsquedas. *Anuario ThinkEPI*, 17. <https://www.doi.org/10.3145/thinkepi.2023.e17a04>
- Mehdi, Y. (2023). Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web. *Official Microsoft Blog*. Retrieved from <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>
- Mladenović, D., Rajapakse, A., Kožuljević, N. & Shukla, Y. (2023). Search engine optimization (SEO) for digital marketers: exploring determinants of online search visibility for blood bank service. *Online information review*, 47(4), 661-679. <https://www.doi.org/10.1108/OIR-05-2022-0276>
- Metaxas, P. T. & Pruksachatkun, Y. (2016). Manipulation of search engine results during the 2016 US Congressional Elections. In *International Conference on Internet and Web Applications and Services (ICIW)* (pp. 22-26). Retrieved from <https://repository.wellesley.edu/object/ir264>
- Mittal, M. K., Kirar, N. & Meena, J. (2018). Implementation of search engine optimization: through white hat techniques. In *International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)* (pp. 674-678). <https://www.doi.org/10.1109/ICACCCN.2018.8748337>
- Negredo Bruna, S. & Kaufmann-Argueta, J. (2021). Las audiencias de los medios nativos digitales de noticias: competir entre las grandes marcas periodísticas. In R. Salaverría & M. P. Martínez-Costa (Eds.), *Medios nativos digitales en España* (pp. 123-142). Salamanca: Comunicación Social.
- Neuberger, C. (2014). The journalistic quality of Internet formats and services: results of a user survey. *Digital journalism*, 2(3), 419-433. <https://www.doi.org/10.1080/21670811.2014.892742>
- Nevado-Chiné, N., Alcaraz-Martínez, R. & Navalón, J. A. (2021). Análisis de la implementación Schema.org en el repositorio RODERIC e impacto en el posicionamiento en Google y Google Scholar. *Revista española de documentación científica*, 44(3). <https://www.doi.org/10.3989/redc.2021.3.1797>
- Noppari, E., Heinonen, A. & Vainikka, E. (2014). Critical but co-operative: netizens evaluating journalists in social media. *Observatorio (OBS*)*, 8(4), 1-16. <https://www.doi.org/10.15847/obsOBS842014782>

- Onaifo, D., & Rasmussen, D. (2013). Increasing libraries' content findability on the web with search engine optimization. *Library hi tech*, 31(1), 87-108.
<https://www.doi.org/10.1108/07378831311303958>
- Orduña-Malea, E. & Costas, R. (2021). Link-based approach to study scientific software usage: the case of VOSviewer. *Scientometrics*, 126(9), 8153-8186.
<https://www.doi.org/10.1007/s11192-021-04082-y>
- Prawira, I. & Rizkiansyah, M. (2018). Search engine optimization in news production online marketing practice in Indonesia online news media. *Pertanika journal of social sciences & humanities*. Retrieved from <http://www.pertanika.upm.edu.my/pjtas/browse/regular-issue?article=JSSH-T0727-2018>
- Pedraza-Jiménez, R. (2018). Analysis of destination search in Google. *IPBA*, 90, 63-65.
- Pedrosa, L. & Morais, O. J. (2021). Visibilidade web nos buscadores: fatores algorítmicos de SEO *on-page* (FAOPs) como técnica e prática jornalística. *Estudios sobre el mensaje periodístico*, 27(2). <https://www.doi.org/10.5209/esmp.71291>
- Rovira, C., Fernández-Cavia, J., Pedraza-Jiménez, R. & Huertas, A. (2010). Posicionamiento en buscadores de las webs oficiales de capitales de provincia españolas. *El Profesional de la Información*, 19(3), 277-283. <https://www.doi.org/10.3145/epi.2010.may.08>
- Shahzad, A., Jacob, D. W., Nawi, N. M., Mahdin, H. & Saputri, M. E. (2020). The new trend for search engine optimization, tools and techniques *Indonesian journal of electrical engineering and computer science*, 18(3), 1568-1583.
<https://www.doi.org/10.11591/ijeecs.v18.i3.pp1568-1583>
- Soulo, T. (2023). 96.55% of content gets no traffic from Google: here's how to be in the other 3.45% [new research for 2023]. *Ahrefs blog*. Retrieved from <https://ahrefs.com/blog/search-traffic-study>
- Trielli, D. & Diakopoulos, N. (2020). Partisan search behavior and Google results in the 2018 U.S. midterm elections. *Information, communication & society*, 25(1), 1-17.
<https://www.doi.org/10.1080/1369118X.2020.1764605>
- Umenhofer, L. (2019). Gaining ground: search engine optimization and its implementation on an indie book press. *Publishing research quarterly*, 35(2), 258-273.
<https://www.doi.org/10.1007/s12109-019-09651-x>
- Unkel, J. & Haim, M. (2019). Googling politics: parties, sources, and issue ownerships on Google in the 2017 German federal election campaign. *Social science computer review*, 39(5), 1-18. <https://www.doi.org/10.1177/0894439319881634>
- Vázquez, M. & Ventura, A. (2020). Analysis of the SEO visibility of university libraries and how they impact the web visibility of their universities. *The Journal of Academic Librarianship*, 46(4), 102-171. <https://www.doi.org/10.1016/j.acalib.2020.102171>
- Vázquez, M., Lopezosa, C. & Pedraza-Jiménez, R. (2022). A study of the Web visibility of the SDGs and the 2030 Agenda on university websites. *International Journal of Sustainability in Higher Education*, 23(8), 41-59. <https://www.doi.org/10.1108/IJSHE-09-2021-0361>
- Vara, A., Negredo, S., Moreno, E., Kaufmann, J. & Amoedo, A. (2023). *Digital news report 2023*. Reuters. Retrieved from <https://reutersinstitute.politics.ox.ac.uk/es/digital-news-report/2023/espana>
- Yudasubrata, A. T. W., Damanik, R. O. P., Hidayanto, A. N. & Budi, N. F. A. (2019). Search Engine Optimization (SEO) approach in studying information demand and supply: methodology of geo-targeted keywords with case study of hospital websites in Jakarta. In *2019 International Conference on Advanced Computer Science and Information Systems (ICACSIS)* (pp. 519-524). <https://www.doi.org/10.1109/ICACSIS47736.2019.8979776>