

Periodismo y algoritmos: de la era de la información a la era del algoritmo

Lluís Codina¹, Carlos Lopezosa², Cristòfol Rovira³

Resumen. Presentación del número monográfico dedicado al periodismo y algoritmos. Se ofrece la visión panorámica de algunos de los problemas de investigación relacionados, como son los algoritmos que deciden la visibilidad de las noticias, los que afectan a los sistemas de agregación y recomendación de noticias, y algunos aspectos del así llamado robot periodista. Por último, se presentan las investigaciones que forman parte de este monográfico. Se concluye reclamando líneas de investigación sólidas alrededor de estos temas, no solo desde las ingenierías, sino desde las ciencias sociales y las humanidades. Dado el impacto para el futuro del periodismo y de las sociedades libres de epidemias de desinformación, es imprescindible aportaciones desde la comunicación social.

Palabras clave: Periodismo, algoritmos, buscadores, SEO periodístico, sistemas de recomendación de noticias, sistemas de agregación de noticias, robot periodista

[en] Journalism and Algorithms: from the information era to the algorithm era

Abstract. Presentation of the monographic issue dedicated to journalism and algorithms. The panoramic vision of some of the related research problems is offered, such as the algorithms that decide the visibility of the news, those that affect the news aggregation and recommendation system, and some aspects of the so-called journalist robot. Finally, the investigations that are part of this monograph are presented. It concludes by claiming solid lines of research around these issues, not only from engineering, but from the social sciences and humanities. Given its impact on the future of journalism and of societies free of disinformation epidemics, contributions from social communication are essential.

Keywords: Journalism, algorithms, search engines, journalistic SEO, news recommendation systems, news aggregation systems, journalist robot

Sumario. Introducción. ¿Por qué decimos que el siglo XXI es el de los algoritmos? Interpretar la realidad a través de las páginas de resultados. ¿Un nuevo programa de investigación? Las redes sociales y sus escándalos. Robot periodista, la fuerza que no podemos ignorar. La sección monográfica. Referencias.

Cómo citar: Codina, L.; Lopezosa, C.; Rovira, C. (2022): Periodismo y algoritmos: de la era de la información a la era del algoritmo, en *Documentación de Ciencias de la Información* 45 (1), 1-5.

Introducción

A pocos sorprenderá si decimos que estamos en la era de los algoritmos, ya que las noticias sobre su influencia en nuestra vida cotidiana aparecen casi a diario en los medios. Por otro lado, la propia empresa periodística, como el sector económico y social de primera magnitud que es, no escapa a su influencia. Algo que, además, solo podemos esperar que se incremente con el tiempo.

En cambio, en el pasado siglo, se repetía con frecuencia que estábamos en la era de la información. Cabía matizar que la información siempre fue crucial para la humanidad, al menos en cuanto esta dejó atrás su pasado más remoto de la prehistoria. Entonces, lo

que justificaba caracterizar al siglo XX como la era de la información era el peso creciente de esta, así como su presencia y dominio constante en todos los campos de la actividad humana.

¿Por qué decimos que el siglo XXI es el de los algoritmos?

Ahora corresponde hacer una aclaración similar. En puridad, la humanidad se encuentra inmersa en la era de los algoritmos desde que la automatización tomó el control de cada vez más procesos y actividades sociales. Y lo cierto es que esto empezó al menos desde los años 60 del siglo pasado. Mucho más acusada-

¹ Universitat Pompeu Fabra. Profesor titular del Departamento de Comunicación.
E-mail: lluis.codina@upf.edu.
orcid: [0000-0001-7020-1631](https://orcid.org/0000-0001-7020-1631)

² Universitat Pompeu Fabra. Profesor asociado del Departamento de Comunicación.
E-mail: carlos.lopezosa@upf.edu.
ORCID: [0000-0001-8619-2194](https://orcid.org/0000-0001-8619-2194)

³ Universitat Pompeu Fabra. Profesor titular del Departamento de Comunicación.
E-mail: cristofol.rovira@upf.edu.
Orcid: [0000-0002-6463-3216](https://orcid.org/0000-0002-6463-3216)

mente desde la invención de los ordenadores personales en los años 90, que fue lo que puso el poder de la computación al alcance de todo tipo de empresas y actividades sociales.

Entonces, nos podemos preguntar, ¿qué es lo que hace que podamos hablar ahora de una era de los algoritmos? La razón para esto es por la creciente influencia de algoritmos que toman decisiones que antes eran humanas. En su momento, se predijo que el mundo iba a estar dominado por la inteligencia artificial, y este dominio se está expresando en nuestros días a través de sus algoritmos (Dörr, 2016; Orgeret, 2020).

El caso más ubicuo de tal situación, a saber, algoritmos haciendo cosas parecidas a la toma de decisiones humanas, lo tenemos en dos grandes actores de nuestra vida cotidiana: los buscadores y los sistemas de recomendación de noticias en las redes sociales.

Examinemos esto someramente con el caso de la búsqueda. Hace tiempo que los directorios donde se clasificaban de forma manual las páginas web pasaron a mejor vida. Fueron muchas las iniciativas creadas en el pasado siglo para intentar ordenar Internet de forma “manual”. Algunas del sector privado y otras del sector académico. Ejemplo de las primeras fueron los famosos directorios de Yahoo primero y de DMoz después. Pero muy pronto, los buscadores basados en algoritmos fueron tomando su relevo. Al principio convivieron con los directorios, pero acabaron definitivamente con ellos en algún momento de la primera década del 2000. Lo que nos lleva a la situación actual.

El efecto Google

Con diferencia, el buscador más influyente en casi todo occidente es Google. Su predominio se basa en sus poderosos algoritmos de filtrado y de ordenación de resultados. En el caso de Google, los expertos aseguran (y la propia documentación oficial de Google confirma) que su algoritmo de ordenación utiliza más de 200 factores, aunque hay al menos uno de ellos que posee una fuerza especial. Se trata del análisis de hiperenlaces. Un análisis que Google lleva a cabo de forma automática. Mediante este análisis, la importancia de una página para un tema determinado se establece (entre otras cosas) computando el número y la calidad de los hiperenlaces recibidos.

Ciertamente, Google utiliza otros algoritmos. Por ejemplo, para rastrear la web, para categorizar páginas o para interpretar las intenciones de los usuarios cuando lanzan sus preguntas. De modo que estamos ante cosas que, en el pasado, solamente se podían hacer de forma intelectual, o si lo prefieren, “a mano”, por usar esta expresión con la que todos nos entendemos. En cambio, al menos desde inicios del 2000 se hacen de forma mayoritaria (sino exclusiva) de forma algorítmica, pues ya no existen (o son anecdóticos) los directorios generalistas de páginas web a los que

nos hemos referido más arriba. Es un ejemplo cotidiano, casi transparente por su ubicuidad, del paso de la era de la información a la era de los algoritmos.

Interpretar la realidad a través de las páginas de resultados

Volvamos a los buscadores. La clave del rol que los algoritmos juegan en el caso de los buscadores es la ordenación de los resultados (Schultheiß y Lewandowski, 2021). Esto, que parece totalmente inocente, tiene una importancia fundamental (Lewandowski, et al. 2021) desde el momento en que los usuarios de los buscadores consultan solamente los primeros resultados. Pensamos que aún no se ha reconocido, ni por parte de los analistas ni de la propia sociedad en su conjunto, la importancia real que tiene. Aunque en este punto, debemos declarar que los autores de esta presentación han llamado la atención sobre esto, al menos, en el ámbito del SEO académico (Rovira et al., 2019; Rovira et al., 2021).

Pero vamos al punto que nos interesa aquí: aunque un resultado en Google arroje 100, 200 o 200 mil resultados, incluso 200 millones, esto carece totalmente de importancia. La razón es que es sabido que los usuarios examinarán apenas unos pocos resultados. ¿Cuáles? Los diez primeros (Ong, 2020). He aquí la cuestión. Dado que los ciudadanos van a examinar únicamente los diez primeros resultados, es imposible exagerar la importancia del algoritmo de ordenación, puesto que van a ser exactamente los diez primeros los únicos resultados que van a ser tomados en cuenta, sin importar si la búsqueda arrojó en realidad 200 o 2000 resultados (o dos millones).

De modo que es el algoritmo de ordenación entonces el que determina a qué información se expondrán, de facto, los ciudadanos de medio planeta, cada vez que busquen información sobre temas que les afectan. A veces estos temas serán triviales, como encontrar vuelos baratos de un lugar a otro, o saber en qué cines se exhibe tal película. Pero muchas veces los temas serán, sin exagerar, de vida o muerte, como cuando un ciudadano quiere informarse sobre las vacunas, y en estas situaciones críticas, la visibilidad se vuelve un elemento clave para la relevancia social del periodismo. A su vez, solamente un aumento de la relevancia del periodismo puede desplazar la intoxicación informativa en la, por otro lado, hipersaturada oferta informativa de internet (Castells et. al 2022).

SEO periodístico

Todo lo anterior se puede extrapolar con todas sus consecuencias al caso del periodismo y su visibilidad en el contexto de la hiper saturada oferta informativa de la web (Giomelakis y Veglis, 2015; 2016). Cuando las consultas de los ciudadanos tienen un componente de actualidad, Google activa los elementos de su

algoritmo que hará que la mayor parte de la página de resultados, o al menos, la parte superior de la página de resultados quede poblada por noticias. Es decir, otras clases de fuentes quedarán desplazadas para dar espacio a contenidos producidos por medios de comunicación en el limitado espacio de su página de resultados.

A su vez, solamente los contenidos de medios de comunicación que hayan sido optimizados para los buscadores podrán ocupar este espacio. Mientras que los medios de comunicación se cuentan por varias decenas de miles en todo el mundo, o por varios cientos en cada país, lo cierto es que solamente unos pocos podrán colocar cada vez parte de su producción, en este limitado espacio que es la página de resultados de un buscador.

Creo que con esto podemos apreciar hasta qué punto es importante que los algoritmos de los buscadores en general, y el de Google en particular, sean conocidos por parte de los profesionales del periodismo, y de modo que esto lo dice casi todo sobre la importancia y la necesidad del SEO periodístico para las empresas de comunicación (Lopezosa et al. 2020a; 2020b).

¿Un nuevo programa de investigación?

La cuestión es que la importancia creciente de los algoritmos que gobiernan la búsqueda o los sistemas de recomendación de las plataformas digitales podría sugerir la necesidad de un nuevo programa de investigación del que, entonces, este número que el lector tiene en sus manos es un necesario primer paso.

Aunque este posible programa de investigación podría hacer muchas preguntas. De momento, nosotros sugerimos estas dos: ¿es democrático que el buscador más utilizado del mundo siga manteniendo en secreto las claves de su algoritmo cuando se aplica en sectores vitales para la democracia y la convivencia social como es la información de actualidad? ¿Hasta qué punto la visión de la (supuesta) realidad que ofrecen los resultados ordenados algorítmicamente afectan las decisiones de los ciudadanos en momentos clave tales como las elecciones o las recomendaciones de campañas de salud?

Las redes sociales y sus escándalos

Todo o buena parte de lo que hemos señalado hasta ahora, *mutatis mutandis*, se puede trasladar al mundo de las redes sociales, que junto con los buscadores, forman las tan traídas y llevadas plataformas digitales. Las que, por cierto, han arrebatado al periodismo una parte de su anterior exclusividad a la hora de informar sobre la actualidad, con resultados no precisamente positivos para la salud democrática de los países.

Como es sabido, las redes sociales, notablemente en este caso, Facebook, YouTube, Twitter y otras

actúan de facto como sistemas de agregación y recomendación de informaciones de actualidad.

En este caso, los problemas políticos, democráticos y sociales en general de los algoritmos han sido ampliamente tratados incluso en los medios de comunicación, a diferencia del caso de los algoritmos de Google, cuyo trato ha sido mucho más en términos técnicos (léase en términos de SEO).

Poco más diremos aquí a propósito de los algoritmos y las redes sociales, salvo reiterar en la necesidad de persistir en su estudio, y en la permanente necesidad de vigilar y denunciar las posiciones abusivas de las respectivas redes. Muy especialmente cuando, de forma expresa, modelan sus algoritmos para favorecer la polémica, vinculada casi siempre con episodios de desinformación.

Robot periodista, la fuerza que no podemos ignorar

No queremos dejar de mencionar el fenómeno denominado a veces con la expresión “robot periodista”. Como seguramente no ignoran los lectores de esta revista, con la expresión robot periodista se menciona una realidad muy consolidada y que parece que seguirá ganando importancia en el futuro (Carlson, 2014). En realidad, el robot periodista es parte del fenómeno más amplio del llamado periodismo computacional (Linden, 2017; Váñez et al., 2018; De la Hoz et al., 2020).

Se trata de la tendencia a incorporar algoritmos en el corazón de las redacciones de cada vez más medios de comunicación de todo el mundo. Aunque se denominan robots, obviamente se trata de software. Es una clase de software capaz de producir noticias en las que el componente central son datos tales como resultados deportivos o datos de tipo económico o financiero. También pueden producir noticias sobre eventos, tales como accidentes o catástrofes naturales o artificiales, en los que se requiere reunir de forma rápida datos e informaciones puntuales procedentes de muy diversas fuentes y cuentas de redes sociales (Dörr, 2016).

En estos casos, el debate más intenso tiene relación con el futuro de la profesión periodística. Esto es, en este caso, el miedo predecible es por la posible sustitución de periodistas humanos mediante robots -léase, mediante software y algoritmos-. Por suerte, el diagnóstico más repetido por los expertos es que el robot periodista sustituye tareas repetitivas e ingratas por naturaleza, pero no puede sustituir a los periodistas humanos (Linden, 2017).

La razón es que lo mejor del periodismo no está en esta clase de labores repetitivas o mecánicas consistentes en agregar datos puntuales. Lo mejor del periodismo, y aquello que los ciudadanos esperan del mejor periodismo, tiene mucho más que ver con actividades genuinamente inteligentes que ninguna inteligencia artificial puede ni siquiera acercarse. Tiene

que ver con la capacidad para entender e interpretar el entorno, y con la capacidad de ejercer la vigilancia del poder, entre otras muchas funciones que exigen creatividad e inteligencia real, no simulada. La idea principal es que la inteligencia artificial impulse el periodismo, en lugar de amenazarlo.

En concreto, los expertos consideran que la inteligencia artificial debe ser parte de los esfuerzos para la sostenibilidad de las empresas periodísticas (De la Hoz et al., 2020; López, 2021). Se puede citar en este sentido el reciente anuncio de la influyente Knight Foundation de dotar con 3 millones de dólares un fondo para introducir la inteligencia artificial en las redacciones de medios locales (Cheung, 2021). No en vano, el artículo donde informan de esta iniciativa abre con esta expresión definitiva: “Artificial intelligence is a force we can’t afford to ignore” (Cheung, 2021).

La sección monográfica

Por último, corresponde presentar los contenidos del apartado monográfico de este (dejénnos decirlo), oportuno número, ya que incluye una diversidad de trabajos que trazan un arco sobre los temas que afectan al periodismo y a los algoritmos.

En primer lugar, la sección monográfica incluye el trabajo de García-Marín, titulado “Modelos algorítmicos y fact-checking automatizado. Revisión sistemática de la literatura”. El fact-checking algorítmico o automatizado consiste, tal como nos explica el autor, en la comprobación de la veracidad de una información aplicando tecnologías de inteligencia artificial. En este trabajo se presenta por primera vez un análisis sobre la evolución y los países de procedencia de la investigación en estas técnicas. Asimismo, estudia la precisión de las soluciones basadas en algoritmos y el impacto de los trabajos analizados. Cabe señalar que el trabajo consiste en una revisión de la literatura, lo que constituye una forma idónea de abrir el número.

El siguiente trabajo, debido a Trillo y Gallego, “SEO para periodistas: nuevo perfil profesional y estrategias de publicación desde la redacción”, presenta importantes claves sobre este punto crítico del modelo de sostenibilidad de la prensa, que significa el SEO periodístico, con la enorme ventaja para los lectores, que está hecho desde dentro de una redacción periodística real. El trabajo de estas autoras establece, en-

tre otras cosas, que el impacto de los algoritmos está transformando el perfil del periodista, ya que ahora debe poseer competencias en producción, edición y publicación SEO. Pero, como explican las autoras, también está modificando la propia *agenda setting* de los medios, al verse obligados a incorporar las tendencias y temas de interés de las audiencias. Las autoras realizan, así mismo, un acercamiento al posicionamiento en buscadores con el objetivo de explorar las oportunidades que supone este perfil emergente para los diferentes perfiles profesionales afectados.

Por su parte, Boté-Vericad y Váñez, en su trabajo titulado “Aplicaciones de deepfakes. Manipulación de contenido audiovisual y riesgos para los usuarios basados en las políticas de privacidad” presentan otro de los temas clave de la desinformación debida a manipulaciones algorítmicas. Se trata del llamado “deepfake”, que se centra en la manipulación de registros de vídeo, consiguiendo vídeos falsos con una profunda sensación de realidad, y que en manos de actores de la intoxicación informativa, pueden causar un enorme daño, como es fácil de comprender.

Por último, la sección se completa con el trabajo de Pedraza-Jiménez titulado “Evaluación de la calidad de los agregadores de noticias: propuesta de un nuevo instrumento de análisis basado en las características de los agregadores de noticias chinos de referencia”. En este caso, el trabajo aborda el tema de los agregadores, que utilizan sistemas de recomendación. El trabajo en concreto propone un sofisticado sistema de análisis de esta clase de instrumentos, y lo refiere al caso de los agregadores de noticias chinos. Se trata de una interesante oportunidad de conocer algo más sobre el país que está llamado a tener cada vez un papel de referencia más importante en todo el mundo, y en este sentido cualquier oportunidad de conocerlos mejor es muy bienvenida.

Solo nos resta recomendar a todos los interesados por el futuro del periodismo una revisión atenta de los temas que nos aporta este número. Nuestra motivación es que el estudio de los algoritmos en el periodismo ya sea con este o con otro nombre, como el de periodismo computacional, se transforme, cuanto antes mejor, en una muy sólida y muy crítica corriente de estudios de orientación crítica. Las ingenierías son imprescindibles en este campo, pero la visión social, humanista y crítica, desde los estudios de comunicación cuando el tema afecta al periodismo, son igualmente imprescindibles.

Referencias

- Carlson, Matt (2015) “The Robotic Reporter”. *Digital Journalism*, 3:3, 416-431, DOI: [10.1080/21670811.2014.976412](https://doi.org/10.1080/21670811.2014.976412)
- Cheung, Paul. (2021). “AI for local news: advancing business sustainability in newsrooms”. *Knight Foundation*, May 12, 2021, <https://knightfoundation.org/articles/ai-for-local-news-advancing-business-sustainability-in-newsrooms/>
- De la Hoz, Karen; Coelho, Florencia (2020) “IA para periodistas. Una herramienta por explotar”. *Prodigioso Volcán*. <https://www.prodigiosovolcan.com/sismogramas/ia-periodistas/prodigioso-volcan-IA-para-periodistas.pdf>

- Dörr, Konstantin Nicholas (2016) "Mapping the field of Algorithmic Journalism". *Digital Journalism*, 4:6, 700-722, DOI: [10.1080/21670811.2015.1096748](https://doi.org/10.1080/21670811.2015.1096748)
- Giomelakis, Dimitrios; Veglis, Andreas (2015). "Employing Search Engine Optimization Techniques in Online News Articles". *Studies in Media and Communication*. Vol. 3, No. 1, pp. 22- 33. <http://dx.doi.org/10.11114/smc.v3i1.683>.
- Giomelakis, Dimitrios; Veglis, Andreas (2016). "Investigating Search Engine Optimization Factors in Media Websites". *Digital Journalism*. 4:3, 2016 379-400, pp. 379-400. DOI: [10.1080/21670811.2015.1046992](https://doi.org/10.1080/21670811.2015.1046992)
- Lewandowski, Dirk; Sünkler, Sebastian; Yagci, Nurce. (2021). "The influence of search engine optimization on Google's results: A multi-dimensional approach for detecting SEO". In *13th ACM Web Science Conference 2021 (WebSci '21)*, June 21–25, 2021. <https://doi.org/10.1145/3447535.346247>
- Linden, Carl-Gustav (2017) "Decades of Automation in the Newsroom: Why are there still so many jobs in journalism?" *Digital Journalism*, 5:2, 123-140, DOI: [10.1080/21670811.2016.1160791](https://doi.org/10.1080/21670811.2016.1160791)
- López, César (2021). "¿Cómo están usando los medios en español el procesamiento de lenguaje por medio de inteligencia artificial?". *Fundación Gabo* (25 de junio 2021). <https://fundaciongabo.org/es/blog/laboratorios-periodismo-innovador/como-estan-usando-los-medios-en-espanol-el-procesamiento-de>
- Lopezosa, Carlos; Codina, Lluís., Díaz-Noci, Javier; Ontalba, José Antonio. (2020a) "SEO and the digital news media: From the workplace to the classroom". *Comunicar*, 2020, vol. 28, n. 63, pp. 65-75.
- Lopezosa, Carlos; Codina, Lluís; López-García, Guillermo; Corbella-Cordomi, Joan. (2020b) "Mapa de visibilidad y posicionamiento en buscadores de los principales grupos mediáticos españoles". *Profesional de la información*, v. 29, n. 2, e290203. <https://doi.org/10.3145/epi.2020.mar.03>
- Ong, Si Quan (2020). "63 SEO Statistics for 2021". *Ahrefs blog* (June 30, 2020). <https://ahrefs.com/blog/seo-statistics/>
- Orgeret, Kristin Skare (2020). "Discussing Emotions in Digital Journalism". *Digital Journalism*, 8:2, 292-297, DOI: [10.1080/21670811.2020.1727347](https://doi.org/10.1080/21670811.2020.1727347)
- Rovira, Cristòfol; Codina Lluís; Guerrero-Solé Frederic; Lopezosa Carlos (2018). "Ranking by relevance and citation counts, a comparative study: Google Scholar, Microsoft Academic, WoS and Scopus". *Future internet*. 2019;11(2):[22 p.]. DOI: [10.3390/fi11090202](https://doi.org/10.3390/fi11090202)
- Rovira, Cristòfol; Codina Lluís; Lopezosa Carlos (2021) "Language bias in the Google Scholar ranking algorithm". *Future Internet*. 2021;13(2):31. DOI: [10.3390/fi13020031](https://doi.org/10.3390/fi13020031)
- Schultheiß, Sebastian; Lewandowski, Dirk. (2021) "Outside the industry, nobody knows what we do" SEO as seen by search engine optimizers and content providers". *Journal of Documentation*, Vol. 77 No. 2, pp.542-557
- Vállez, Mari; Codina, Lluís (2018). "Periodismo computacional: evolución, casos y herramientas". *Profesional de la Información*, 27(4), 759–768. <https://doi.org/10.3145/epi.2018.jul.05>