

Leer donde las máquinas también leen. Contexto de producción y distribución de contenidos mediados por motores de búsqueda en la web*

Elena María Gómez-Rico¹ 

<https://doi.org/10.18046/recs.i45.04>

Cómo citar: Gómez-Rico, Elena María (2025). Leer donde las máquinas también leen. Contexto de producción y distribución de contenidos mediados por motores de búsqueda en la web. *Revista CS*, 45, a04. <https://doi.org/10.18046/recs.i45.04>

Resumen: Los motores de búsqueda (MB), y sus algoritmos PageRank y RankBrain, actúan como mediadores entre los textos, sus proveedores y sus lectores en el espacio de la web. El objetivo de este artículo es aportar a la construcción de un contexto de producción y distribución de dichos textos en un espacio específico y prominente: la web, cuyas particularidades no necesariamente son abordadas cuando se habla de alfabetización digital. Construir tales contextos, desde la creación de los textos hasta su lectura, es un método transdisciplinar utilizado en este artículo para identificar los efectos de la mediación de los MB en la capacidad de interpretación de estos contenidos. Los contextos descritos evidencian, entre otros efectos, la opacidad de las tecnologías de búsqueda, los sesgos presentes en los resultados y la visibilidad que alcanzan gracias a técnicas específicas de escritura.

Palabras clave: motores de búsqueda, alfabetización digital, alfabetización tecnocientífica, algoritmos *PageRank* y *RankBrain*, optimización para motores de búsqueda

* Artículo de revisión de literatura interdisciplinar, resultado de diversos espacios de discusión en el contexto del *Plan de lectura, escritura y oralidad* de la Secretaría de Cultura de Medellín y el *hackerspace* Unloquer de la misma ciudad. Financiado con recursos propios. Artículo de investigación recibido el 28.02.24 y aceptado el 28.02.25.

1. Universidad EAFIT (Medellín, Colombia)



How to read where there are reading machines. A context for content production and distribution mediated by search engines on the web

Abstract: Search engines (SEs), and their PageRank and RankBrain algorithms, act as mediators between texts, their providers, and their readers in the web space. The aim of this article is to contribute to the construction of a context for the production and distribution of such texts in a specific and prominent space: the web whose particularities are not necessarily addressed when discussing digital literacy. Building such contexts, from the creation of texts to their reading, is a transdisciplinary method used in this article to identify the effects of MB mediation on the ability to interpret these contents. The contexts described evidence, among other effects, the opacity of search technologies, the biases present in the results and the visibility they achieve thanks to specific writing techniques.

Keywords: Search Engines, Digital Literacy, Techno Scientific Literacy, Pagerank and RankBrain, Search Engine Optimization

Introducción

Las competencias digitales para la lectura crítica son un campo de estudio y de desarrollo de herramientas para abordar los retos del contexto de la sociedad de la información (Castells, 2004). La alfabetización ha sumado varios adjetivos que la concentran en los temas de virtualidad. Es así que esta se califica como tecnocientífica (Goetz, 2020), digital (Gilster, 1997) o informacional (Després; Cotte, 2007). En los últimos años se ha dicho que, “aunque la falta de conectividad es un obstáculo en América Latina para un aprendizaje basado en la tecnología, no es el único factor. La falta de competencias y de habilidades digitales (...) siguen siendo las principales barreras” (Organización para la Cooperación y el Desarrollo Económico - OCDE, 2020: 12). Hasta ahora no se ha encontrado un marco evaluativo de competencias digitales que integre dimensiones o capacidades para la comprensión de la web como contexto de producción y distribución de información, más allá de la capacidad de encontrar fuentes de información confiables, contrastar datos de distintas fuentes y otras capacidades relacionadas con aprender a usar de manera acrítica programas del mercado digital.

En Colombia “no se ha definido un marco específico de competencias digitales” (Contreras; Piedrahita; Ramírez, 2019: 1), aunque los investigadores resaltan iniciativas como Ciudadanía Digital del Ministerio de Tecnologías de la Información y Comunicaciones (MinTIC). En los distintos marcos existentes, entre los que se encuentran los de la Unión Europea, los norteamericanos y los propuestos por organismos de la cooperación internacional se proponen

diversas dimensiones. Por ejemplo, el Marco de Competencias Digitales para la Ciudadanía (DIGCOMP) del Joint Research Centre de la Comisión Europea cuenta con uno de los marcos que se reconocen como útiles y ya validados para evaluar las competencias digitales y se usa en distintos contextos con ese fin.

En este, según la compilación realizada por Contreras *et al.* (2019), la lectura crítica hace parte de la dimensión de información dentro de la cual se reconocen dos capacidades relacionadas: la navegación, búsqueda y filtrado de la información, y la evaluación de esta. En ninguna de las dos se integra la capacidad de interpretar los resultados de las búsquedas desde alguna comprensión del funcionamiento de las tecnologías implicadas en el proceso, entre ellas los motores de búsqueda (MB), lo que “lleva a una gran complejidad semiótica” (Després; Cotte, 2007: 5). Esto tiene implicaciones en los comportamientos de lectura.

En los inicios de internet se planteaba que este transformaría todo y que se tendría el poder de participar en el proceso de construcción del conocimiento del mañana (Després; Cotte, 2007). En efecto, internet ha cambiado la manera de dialogar con el mundo dado que, además de fundar un espacio interconectado, introdujo nuevas maneras, mediadores y comportamientos que deben interpretarse como una situación comunicativa particular, es decir, el participante actúa como usuario para responder sus preguntas y como proveedor de contenido para contestarlas (Groys, 2014).

Lo anterior define un contexto de producción que tiene un efecto profundo en nuestro comportamiento de lectura y escritura, dadas las mediaciones tecnológicas necesarias para ello. Para alimentar la reflexión, revisamos el rol de los MB.

Los resultados de una búsqueda o las respuestas a nuestras preguntas se nos presentan como muestras de contextos que la tecnología de los MB puede encontrar (Groys, 2014), lo que evidencia una diferencia fundamental, ya que, en este escenario, las palabras no funcionan bajo las reglas de la gramática y el ejercicio de lenguaje queda reducido a parámetros de aparición de palabras en distintos contextos y a operaciones con un comportamiento metalingüístico (Groys, 2014).

Así, se revisó la literatura en varios campos de investigación¹ para articular elementos que nutrieran una postura de interpretación activa y crítica para lectores o usuarios de la web y evidenciaran hasta dónde estas tecnologías permitían adoptarla. Estos campos son los de la sociometría de los MB (Mager; Norocel; Rogers, 2023), el capitalismo de vigilancia (*surveillance capitalism*) (Zuboff, 2019),

1. Dada esta interdisciplinariedad se hace uso de las citas de citas como recurso de fidelidad disciplinar buscando conservar la relación del autor con su campo.

la economía de la atención (O'Reilly; Strauss; Mazzucato, 2024), la sociopolítica de las búsquedas (*sociopolitics of search*) y sus sesgos (Díaz, 2008), la tecnología oscura (Jasanoff, 2020; Winner, 1980) y la optimización para MB (*search engine optimization*) (Sheffield, 2020), entre otros. Para el análisis de dicha situación comunicativa, estos se integraron con los estudios del discurso, la filosofía de la tecnología y la historiografía del texto. No se encontraron estudios sobre cómo los hallazgos de los diversos campos ayudan a formar usuarios/lectores y consumidores más críticos, asunto al que este artículo pretende sumar.

De acuerdo con esto, el objetivo de este estudio es aportar a la construcción del contexto de producción y distribución de los textos en la web mediada por los algoritmos *PageRank* y *RankBrain*, que hacen parte de los motores de búsqueda (MB) de Google y de otras plataformas que tienen los mismos principios de funcionamiento, para abrir posibilidades de investigación interdisciplinar, fortalecer los ejercicios de lectura crítica y alfabetización digital o tecnocientífica y enfrentar los distintos tipos de ignorancia (Mager *et al.*, 2023) asociados a esta situación comunicativa.

Metodología

Para hacer una lectura crítica o interpretativa de un texto/contenido o de un autor debe describirse el contexto en el que habla a un lector desde sus condiciones materiales de producción y distribución (Bourdieu, 2005; Chartier, 2006) y sus contextos discursivo, institucional e histórico-social (Lopes, 1999). Esto permite interpretar la condición histórica y la situación comunicativa de lo dicho (Bajtín, 1982) y efectuar una recepción activa de los mismos.

En este sentido, primero se explicará el recurso metodológico para describir la situación comunicativa, luego se hará un recuento del contexto de distribución de textos/contenidos en internet. Con los conceptos que este primero aporta, se abordará una descripción del contexto de producción y los cuestionamientos para la formación de la lectura crítica como componente de la llamada alfabetización digital o tecnocientífica.

Se pretende entonces explorar lo que significan las decisiones de diseño y ejecución de los MB y sus algoritmos *PageRank* y *RankBrain* para un lector (receptor o consumidor de contenidos) y un escritor (o un productor de los mismos) cuando la relación está mediada por máquinas virtuales que leen, indexan y optimizan los textos para ser encontrados en ciertas condiciones materiales y sociales particulares.

Situación comunicativa a analizar para el ejercicio de lectura crítica

Para describir la situación comunicativa en la que se encuentra un lector/usuario “es necesario comprender en primer lugar la producción, el campo de producción y la relación entre el campo en el cual ha sido producido y el campo en el cual es recibido o, más precisamente, la relación entre las posiciones del autor y lector”, ya que una lectura crítica se basa en que “los textos tienen implicaciones, que están insertos en redes de problemas, que es necesario reconstituir” (Bourdieu, 2005: 199).

La situación comunicativa puede describirse de la siguiente manera: el espacio específico de distribución textual es la web, la cual genera unas condiciones propias de producción discursiva —que se llamarán producción de contenidos o textos, aunque también se esté hablando de videos, imágenes y demás formatos—, se distribuyen, en gran medida, a través de medios particulares, los MB, y están social e históricamente situadas en redes de problemas. El que aquí se quiere abordar es el de la generación de una cantidad inmensa de textos que no sabemos leer e interpretar por los elementos contextuales que, por un lado, ya hemos naturalizado de manera crítica y, por otro, no podemos conocer.

Por lo anterior, los lectores reciben textos/contenidos indexados de manera personalizada por una máquina que sabe leerlos (los MB). Estos son producidos por autores (desde hace relativamente poco también por modelos masivos de lenguaje natural - LLM) en diversos formatos (audiovisual, escrito o sonoro), después de consultar las preguntas que los lectores le hacen a la que aquí se nombrará *máquina lectora* y que se refiere, en primera instancia, a los MB, aunque esto podría expandirse a otras tecnologías mediadoras. Luego, realizan varias iteraciones de indexación con otras máquinas simuladoras de la primera máquina lectora (programas de optimización) y liberan el contenido que entra en una competencia de posicionamiento con otros creados de la misma manera. Los textos/contenidos esperan una buena clasificación en el *ranking* (*Search Engine Results Page* - SERP), logran una posición determinada y llegan eventualmente al lector humano. Así mismo, continúan evaluándose, refinándose y optimizándose para obtener una mejor puntuación frente al MB.

En esta situación comunicativa hay una especie de género discursivo que muestra “unas formas típicas para la estructuración de la totalidad, relativamente estables” (Bajtín, 1982: 267), y que “se define por la especificidad de una esfera discursiva dada, por las consideraciones del sentido del objeto o temáticas, por la situación concreta de la comunicación discursiva, por los participantes de la comunicación” (Bajtín, 1982: 267). Así, en la situación comunicativa descrita se presenta una voluntad discursiva con una mediación tecnológica

de la comunicación tal que, como planteó Martín-Barbero (2003: 13), “deja de ser meramente instrumental para espesarse, densificarse y convertirse en estructural”, lo que cataliza “nuevos modos de percepción y de lenguaje, nuevas sensibilidades y escrituras”.

Dado que “la comprensión de cualquier texto es necesariamente dependiente del conocimiento de las formas materiales que ha tomado” (Chartier, 1995: 14), hay que enfatizar que, para la especificidad de la esfera discursiva, el conocimiento común del funcionamiento de la virtualidad como forma material de los textos/contenidos es limitado, como ya se ha dicho. Pero no es solo esto, también hay que considerar lo que esta materialidad significa en la larga historia del libro y de la lectura, como lo describió Chartier (1995: 24):

Nuestra revolución actual es obviamente más extensiva que la de Gutenberg. Modifica no solamente la tecnología de reproducción del texto sino también la materialidad del objeto que comunica el texto a los lectores. Hasta la digitalización del texto ninguna innovación tecnológica en la historia había transformado las herencias del manuscrito en su organización por folios y páginas, su jerarquía de formatos (...) y sus ayudas para la lectura (concordancias, índices y tablas)².

A continuación, se remitió a varios momentos de cambio radical en la historia de los escritos en distintas corrientes culturales (de Oriente y Occidente) para dilucidar lo que puede significar la transformación de los textos físicos a los digitales: “la sustitución de la pantalla por el códex es una transformación mucho más radical porque esta cambia los métodos de organización, estructura, consulta y hasta la apariencia de la palabra escrita. Tal revolución requiere otros términos de comparación”. Otros autores contemporáneos se refieren a dos componentes importantes cuando el soporte de lectura es una pantalla; dicen que esta “debe ser organizada para dar información y metainformación sobre la situación comunicativa” (Després; Cotte, 2007: 5), sin embargo, algunos de ellos son invisibles para un lector desprevenido.

En dicho contexto, este artículo describirá los elementos nombrados anteriormente desde el ejercicio de revisión interdisciplinar para evidenciar las redes de problemas en las que estos contenidos/textos están implicados. Esto se realizará en dos partes. La primera se refiere a las variables relacionadas en el

2. Nótese que Chartier hizo referencia a las ayudas para la lectura, entre las que incluyó los índices. Se profundizará en esto más adelante cuando se hable de la manera en que los algoritmos de los MB ejecutan estas mismas ayudas para la lectura en la web y los efectos que tienen para su interpretación.

contexto de distribución de los contenidos. Empieza con un breve contexto de internet, sus navegadores y de los MB; luego se describirán su funcionamiento y algunos aspectos históricos de su diseño. Después, se analizarán los resultados del ejercicio de indexación que se organiza en las páginas de los resultados del motor de búsqueda SERP —que se conocen como tráfico orgánico— y se revisarán los agentes o grupos de interés que la literatura identifica para el contexto.

El lector o usuario final, que no es un receptor pasivo necesariamente, será el nodo de la segunda parte del artículo en la que se describirán elementos del contexto de producción. En esta parte, la literatura encontrada indagó sobre los sesgos presentes en los algoritmos de los MB y nombró algunas estrategias comunes de organización de motores de búsqueda (OMB). El artículo cierra con una reflexión sobre los vacíos que esto muestra en la alfabetización digital, tecnocientífica o tecnológica.

Parte 1. Contexto de distribución de contenidos/textos en la web

Breve contexto de la web como materialidad de los contenidos/textos y de los MB como máquinas lectoras

La web surgió en los años cincuenta, su protocolo para compartir información se estandarizó en 1982 y se creó como *World Wide Web* en 1989 (Ritchie; Mathieu; Roser; Ortiz, 2023). Con la aparición del primer anuncio *clickable* en 1993 los métodos de mercadeo empezaron a aprovechar la visibilidad que podían lograr allí (Gauri-Girish; Vitthal-Gaikwad; Dhananjay, 2023).

En sus inicios se decía que, además de las actividades de mercadeo:

Un nuevo medio ha emergido y ha prometido cambiar todo esto [influencia de los medios tradicionales en nuestras decisiones]. Descentralizado y distribuido, el internet global, y en particular la web, permite a cualquiera y a todos visibilizar sus perspectivas y conocer las de los otros. (Díaz, 2008: 2)

Aunque aún se cree que la web es intrínsecamente democrática e igualitaria (Díaz, 2008; Mager *et al.*, 2023) —dado que se participa como usuario o como creador de contenidos, lo que han demostrado diversos estudios encontrados en la revisión hecha para este artículo—, esta tiene una forma y una topología de interconexiones que organiza dicha participación de manera específica, pues allí “las interconexiones entre los textos a través del hyperlinking darán forma

a la web y encarnarán otro tipo de política” (Ridgway, 2023: 2). Evidenciando los efectos de estas, las comunidades académicas respondieron con series de revistas destinadas a la crítica a Google “cimentando, en aproximadamente dos décadas de trabajo, estudios sobre cómo los motores de búsqueda privilegian ciertas voces y marginalizan otras” (Mager *et al.*, 2023: 6).

Uno de estos estudios fue desarrollado por el profesor Barabási, desde la teoría de redes libres, quien hizo un mapa hace más de 20 años de la estructura de la web, e incluso en ese entonces, dado el diseño del MB,

descubrió que un número limitado de páginas, lo que llamó *hubs*, están enlazadas a una gran cantidad de páginas mientras que la vasta mayoría de documentos están linkeados a muy pocas o ninguna otra página; por lo tanto, las primeras tienen siempre mayor probabilidad de re-enlazarse o de ser descubiertas. (Barabási como se citó en Díaz, 2008: 14)

Otros estudios, como el de los profesores Felten *et al.* (2004), interpretaron este sistema como uno de votación, “no un oráculo misterioso de la verdad. (...) Los autores web votan creando hipervínculos y (...) Google interpreta un enlace de la página A a la página B como un voto” (Díaz, 2008: 6). Este mapeo mostró que la forma de internet como red, o como la autopista de la información (Ridgway, 2023), no refleja necesariamente la descentralización ni la redistribución que se prometía en sus inicios (Rogers, 2023), mucho menos la democratización de los medios de producción de contenido.

Entre las fuerzas que le dan esta forma a la web, los autores identificaron de manera contundente a los MB como tecnología mediadora del uso y producción de contenidos. Dos años después de su aparición, los investigadores Introna y Nissebaum plantearon de manera crítica sus efectos:

Si los mecanismos de búsqueda estrechan la visión de lo que quienes buscan pueden encontrar y los sitios que pueden ser hallados, estos van a disminuir el valor general de la web como foro público y también como una fuente de información inclusiva y amplia. (como se citó en Ridgway, 2023: 2)

Richard Rogers (2023: 2) complementó lo anterior al sostener que “lo bien enlazado tiene además otra ventaja, aparte de ser indexado. Esto lo posiciona comparativamente mejor en los rankings”.

Los autores se refirieron a dos acciones: la indexación y el ordenamiento (*ranking*) de la información que, históricamente, no han sido ajenas a los libros, como ha mostrado Chartier (1995), pero que, aplicadas a la web, tienen efectos

como los que declararon Ridgway (2023) y Rogers (2023). Para comprender mejor dichos efectos, se describirá lo que la literatura muestra sobre el funcionamiento de estas tecnologías web.

Introducción breve sobre los MB y algunos de sus criterios de diseño

En las décadas siguientes al surgimiento de la web, *surfear*, término que se impuso para describir la interacción con esta, podía hacerse a través de programas llamados graficadores de enlaces o grafos (*link graphs*), índices de alta calidad construidos y actualizados por humanos como Yahoo o Altavista (Brin; Page, 1998), o a través de los MB existentes cuyo alcance era limitado (como el *World Wide Web Worm* de 1997) y algunos otros privados (ya fueran académicos o corporativos).

En 1998, Sergei Brin y Lawrence Page (1998), entonces estudiantes de la Universidad de Stanford, publicaron un artículo sobre un motor de búsqueda de gran escala que se planteaba automatizar la navegación en la web incluyendo grandes cantidades de datos de manera eficiente. Estaba diseñado para adaptarse a los cambios tecnológicos que se proyectaban para integrar hipertextos, con posibilidades de adaptación para la capacidad de computación de distintas máquinas y, finalmente, para mejorar la usabilidad de la web, ya que para llegar a una página o contenido debía saberse la dirección exacta, lo cual era un problema de optimización. Los problemas que debían resolverse para que esta tecnología tuviera una aplicación fuera de los ámbitos académicos eran, entre otros, la cantidad de tiempo necesario para obtener resultados completos y de calidad.

Un motor de búsqueda es un programa que se compone de varios algoritmos. El primero de estos, diseñado por Brin y Page, que arrojaba los resultados de las búsquedas, se llamó *PageRank* —tomando el apellido de Lawrence Page y un verbo que hace referencia a su función principal—. Este algoritmo fue el origen de una manera distinta de acceder a la información en la web y el inicio de la compañía tecnológica Google, que ahora conocemos como Alphabet, una de las cinco empresas que monopolizan el sistema base en Estados Unidos, “la cual ahora ha penetrado el núcleo de la vida económica y civil en la mayoría de los continentes, excepto China” (Van-Dijck, 2021: 2).

El algoritmo fue patentado por sus diseñadores, en conjunto con el Massachusetts Institute of Technology (MIT) y, por esta razón, el documento de su publicación original, resultado de un proceso académico, está protegido por derechos de autor y no puede conocerse a detalle. Para conocimiento general, se liberó una versión que describe las decisiones de diseño y los problemas que buscó solucionar como testigo del evento.

Los MB previos usaban palabras sin contexto semántico ni orden sintáctico para ofrecer resultados (Brin; Page, 1998), lo que era poco eficiente para la calidad deseada de las búsquedas y su relevancia. El principio de calidad (Brin; Page, 1998), precisión o importancia (Díaz, 2008) de una página para una búsqueda definida se obtiene de la confianza que una le entrega a otra, simulando la práctica de citación académica. Así lo nombraron los diseñadores del algoritmo: “de manera general, contando la cantidad de citaciones o *backlinks* que un sitio web contiene además de la presencia de ciertas palabras de búsqueda” (Brin; Page, 1998: 4), pero, aunque el algoritmo “calcula cada página en su índice, la importancia de cada enlace difiere y no se cuenta igualmente” (Ridgway, 2023: 4).

De esta manera, el motor de búsqueda es un mecanismo de indexación, es decir, es un ordenador del contenido y un moderador que arroja de manera organizada los textos dependiendo del cálculo que sitúa la Uniform Resource Locators (URL) en el *ranking* construido con unas características definidas (Mager *et al.*, 2023).

En el artículo, Brin y Page (1998: 4) también propusieron que *PageRank* podía entenderse como “un modelo del comportamiento de un usuario buscando información de confianza en internet”. Con esto, nombraron una especie de usuario arquetípico que necesitaba encontrar información que estaba dispersa en la web. Más adelante hablaremos del usuario de confianza o *trusted user* que es un elemento clave de este contexto de producción de los contenidos/textos en internet.

Aunque “todos los ingredientes para *PageRank* ya estaban disponibles desde mediados de 1990” (Ridgway, 2023: 6), el MB de Google evoluciona sin descanso. Desde 2007, varios artículos indicaron que “está usando más de 200 señales para determinar la clasificación de una página web en su indexación, con cientos de máquinas involucradas en el proceso para cada búsqueda” (Eustace como se citó en Evans, 2007: 2). Según Evans (2007), las actualizaciones anuales realizadas a los algoritmos pueden ser del orden de los miles, lo que suma a esto mayores dificultades para comprender su funcionamiento, incluso suponiendo que se pudiera conocer el diseño o el código, lo cual no es posible por las razones de propiedad intelectual ya expuestas (Ridgway, 2023). Con las actualizaciones, “*PageRank* es entonces tanto un sistema de análisis de enlaces como un sistema IR (sistema de recopilación de información - *Information Retrieval*)” (Rieder como se citó en Ridgway, 2023: 6), y con el desarrollo se ha nombrado además como máquina perfiladora (*profiling machine*). Por esta y otras razones se considera como una tecnología de caja negra (*black box technology*), he aquí la razón de usar el adjetivo en este contexto.

El algoritmo *PageRank* ha ido migrando paulatinamente a *RankBrain* desde el 2014, “un algoritmo de machine learning que está siendo implementado para

hacer consultas irreconocibles (...) y ahora es considerado la tercera señal después de los enlaces y el contenido” (Sullivan como se citó en Ridgway, 2023: 10). En 1998, los creadores del MB dicen simplemente que Google es una herramienta de búsqueda (Ridgway, 2023), sin embargo, cabría preguntarse si su función ha cambiado y analizar sus impactos, ya que “los resultados de Google se han escrutado alrededor de los años por lo que privilegian, sea la web de superficie (*the surface web*), las poderosas y optimizadas páginas web, la función de personalización y otras propiedades” (Rogers, 2023: 1) que comúnmente favorecen intereses económicos (Mager *et al.*, 2023).

Respecto a lo anterior,

Google incorpora desarrollos dentro de la disciplina de la inteligencia artificial, machine learning y LLM (*large language models* - algoritmos de lenguajes de gran escala) como BERT, MUM, PaLM, GLaM, iniciativas como la de los *1000 Languages*, LaMDA y, el chatbot Bard dentro del MB. (Ridgway, 2023: 14)

En 2024, este ha tomado un nuevo nombre, Gemini, cuyo efecto como un LLM apenas está por verse. Estas tecnologías oscurecen parte del ejercicio necesario para la interpretación de los textos en la web, ya sea que actuemos como usuarios o como creadores de contenidos.

A continuación, se describen algunas de las decisiones de diseño de los MB que hacen parte de la función que desempeñan y dan la forma centralizada y no distribuida a la web. Aquí se nombrarán, aunque se articularán a la situación comunicativa en el contexto de producción.

Algunas decisiones de diseño de los algoritmos *PageRank* y *RankBrain*

La literatura describe siete decisiones de diseño de los MB que tienen efecto en la reflexión que se propone este artículo. El primero es el usuario de confianza o *trusted user*, mecanismo del MB de retroalimentación de usuario para definir el peso correcto del enlace para la indexación y el *ranking*. Los lectores, los usuarios, o quienes reciben el contenido funcionan como evaluadores de la calidad de la búsqueda. Aunque este tipo de usuario es un evaluador de varias iteraciones del algoritmo de búsqueda en los primeros diseños, también es uno de sus principios de funcionamiento, pues tal es el nombre del nivel de confianza que los usuarios/lectores tienen con respecto a la herramienta: “Usuarios que tienen un alto nivel de confianza en los MB y en Google en particular” (Schultheiß; Lewandowski, 2021: 2). Esto ha sido demostrado por numerosos estudios, por

ejemplo, sobre la confiabilidad de los resultados, la confianza en los mismos en el *ranking*, particularmente, y en el uso de SERP como su criterio de calidad.

La siguiente herramienta ya se ha descrito de manera reiterada, pero no se ha nombrado como tal, es la estrategia de los *backlinks*, que es la actividad de cocitación entre páginas web y otros tipos de contenidos bajo una URL. Quienes estudian el comportamiento del algoritmo proponen que este es uno de los indicadores más directos para una buena clasificación en el *ranking*, como se describirá más adelante.

Otra herramienta es la personalización, en la que “usuarios de confianza se agrupan en coincidencias de audiencias con consumidores que se vuelven propiedad de Google y, adicionalmente, se correlacionan con audiencias por similitud dentro de las bases de datos de *YouTube* y *Gmail*” (Ridgway, 2023: 11). Ridgway (2023: 12), trajo reflexiones de Esposito, Mayer y Cukier, investigadores de los efectos de *big data*, para explicarlos desde otras formas de uso estadístico como generar perfiles de usuarios: “estos procedimientos digitales que guardan búsquedas constantes y calculan correlaciones luego, procesan los patrones y generan datos nuevos que reemplazan la validez estadística”. Con este ejercicio no hay necesidad de muestreo de una población para investigar sus comportamientos porque se puede obtener información casi de su totalidad.

La estrategia de búsqueda a través de palabras clave (*keyword searches*) es una característica importante desde los inicios de este MB y permitió flexibilizarlas por lo que un usuario/lector puede ingresar palabras aproximadas y aun así obtener resultados. En el caso de buscadores previos al de Google, estos exigían que se indicaran las palabras exactas con ortografía correcta para poder obtener resultados. Según Feuz (como se citó en Ridgway, 2023: 7), lo anterior generó una base de datos para la plataforma que “reflejó los pensamientos de los usuarios como si fueran consultas, impartiendo el poder de una cierta gobernabilidad semántica”. Además, Ridgway (2023) sostuvo que realizar consultas con palabras clave se ha convertido en un hábito de los nuevos medios, o sea que este aspecto es extensible a otros espacios de comunicación virtuales.

La funcionalidad de autocompletar lo que el usuario/lector busca ha evidenciado un tipo de sesgos distintos a los que dan forma a la web a través de los *backlinks* o cocitación, ya que empezó a evidenciarse que los resultados de las consultas y los perfiles creados con el procesamiento de esa gran cantidad de información presentaban sesgos de estereotipo y eran socialmente problemáticos: “Google ha oscurecido los orígenes de autocompletar para dejarlo impune a los estereotipos y otros resultados problemáticos” (Roger, 2023: 5).

El último recurso que se integra para la lectura es el de la publicidad (*advertising*). Es la línea de negocios de publicidad pagada del MB que, aunque se sabe

que hace parte de este, algunos estudios muestran que los usuarios “no están suficientemente familiarizados con los modelos de negocios de Google y tienen dificultades distinguiendo *ads* de los resultados orgánicos” (Schultheiß; Lewandowski, 2021: 2), entonces, no es tan clara la división entre el espacio dedicado a tal ejercicio con fines comerciales y el tráfico o resultados orgánicos, es decir, los que participan en el *ranking*.

Los diseñadores del MB argumentaron que esta funcionalidad era un apoyo para que los consumidores siempre accedieran a los mejores productos (en una de las versiones del artículo publicado en 1998), sin embargo, el diseño del MB se basa en la confianza que un usuario tiene en su neutralidad y las estrategias usadas para la funcionalidad argumentada, se vuelven casi invisibles, incluso para un ojo entrenado pues el objetivo es justo este. Lo que logran las técnicas de optimización para MB es diseñar estrategias para disminuir lo que en el campo llaman fricción— para avanzar en el *funnel* o túnel de conversión (término del mercadeo digital) en el que un lector se convierte en comprador.

Estos solo son algunos de los procesos del funcionamiento de este motor de búsqueda, aunque no se puede saber exactamente cómo trabaja. Por esta razón, los MB se caracterizan como tecnologías de caja oscura o *black box technologies* (Ridgway, 2023), lo que representa parte de la dificultad que termina en imposibilidad de hacer lectura crítica de los contenidos en la web. Esto se argumenta desde distintas perspectivas en la crítica al capitalismo cognitivo o al capitalismo de vigilancia. Sin subvalorarlas, aquí interesa resaltar que el adjetivo –el de caja oscura– indica que el ejercicio de conocimiento “es inicialmente desestructurado y amenazado por esa información que no puede alcanzarse” (Zuboff, 2019: 2).

La base de este argumento está en las operaciones de traducción de la teoría de las redes sociotécnicas de Bruno Latour que propone que dichas operaciones sean realizadas por todos los agentes del sistema y así estos podrían “interpretar el significado de los artefactos tecnológicos” (Feenberg como se citó en Ridgway, 2023: 3). En este artículo, se problematizó el hecho de no poder realizar la operación de traducción necesaria para la lectura crítica dado el oscurecimiento de los medios para acceder a los textos en internet.

Se concluye entonces que los MB no son transparentes y no pueden serlo, ya sea por estar protegidos por el secreto empresarial o porque no pueden ser de dominio común, debido a la visibilidad que logran otorgar bajo sus mecanismos centralizadores y organizadores de los resultados personalizados de búsqueda.

A continuación, se describirá lo que la literatura muestra sobre los agentes participantes en esta situación comunicativa.

Agentes (*stakeholders*) en el contexto de producción y distribución de contenidos en la web

En la situación comunicativa descrita, distintos autores han reconocido varios agentes o grupos de interés que influyen en las diversas etapas del procedimiento, ya que “el proceso de publicación cualquiera que sea su modalidad siempre es un proceso colectivo, que implica a numerosos actores para la materialidad del texto de la textualidad del libro” (Chartier, 2006: 14).

Para acceder a información en internet hay varias maneras. La primera es conocer la URL en la que está la información (porque distintos moderadores de contenido humanos o sistematizados lo provean o simplemente porque se sabe). La segunda es usar un MB, ya sea a través de un asistente o de manera directa. También existen “los *metasearch engines - metacrawlers, niche search engines* - como *Deadliner*, portales de información - como *Yahoo*, y motores de comparación de compras - *mySimon*” (Bhargava; Feng, 2002: 117), algunos ya desaparecidos. Desde el 2021, esta mediación también está empezando a ser realizada por las llamadas inteligencias artificiales basadas en los LLM que se vienen relacionando.

Aunque existen proveedores distintos de MB, “Google domina alrededor del 92 % del mercado de las búsquedas mundiales, así era hasta noviembre 2022” (Ridgway, 2023: 1). Esto quiere decir que este contexto tiene nombre propio y, dadas las ventajas que ofrece, se considera para los efectos como primer recurso: “consultar principalmente empieza en un sitio de motor de búsqueda, lo que significa que *cualquier* página puede convertirse en la frontal” (Després; Cotte, 2007: 5) para obtener la atención del usuario. Esto “les da [a los MB] una influencia tremenda en la adquisición de conocimiento de los usuarios de internet” (Schultheiß; Lewandowski, 2021: 1) y sus comportamientos de lectura.

Se encontraron pocos estudios que describieran el contexto en este sentido, los hallados nombraron los siguientes grupos de interés: 1) los creadores o proveedores de contenidos (*content providers*) que agrupa agencias y servicios de profesionales de diversas disciplinas y formatos dedicados a crear textos/ contenidos para la web que se entregan al usuario/lector a través de servicios de otros profesionales que diseñan programas para gestionar contenidos (unos mejor reputados para la indexación que otros). Estos trabajan de la mano con 2) los optimizadores para motores de búsqueda (*search engine optimizers - SEO*) que son quienes están estudiando la competencia de los textos/contenidos en los MB y sus SERP, y diseñan estrategias para escalar allí que luego ejecutan los creadores o proveedores de contenido, quienes “son forzados a diseñar sus páginas de acuerdo a los algoritmos de búsqueda para poder obtener visibilidad” (Duka; Sikora; Strzelecki, 2023: 3).

Para ellos, “hacen que las cosas sean notablemente más complejas” (Després; Cotte, 2007: 5), lo que además tiene un impacto en la escritura por la baja importancia dada en estas técnicas a criterios de periodismo de calidad. Los criterios de optimización incluyen diferencias para distintos dispositivos (tabletas, celulares y computador), mecanismos para la accesibilidad (con textos y audios para minorías con discapacidad), normalización del estilo o tono (con una fuerte tendencia a los que se conoce como llamado a la acción - *call to action*), transformaciones en el *ethos* del escritor y en el orden en el que se dispensa la información, entre otros, para crear el valor percibido del texto/contenido en la web (Sheffield, 2020).

Describen como agente también a 3) la página de resultados SERP como un grupo de interés en el marco de la teoría actor-red de Bruno Latour y, por lo tanto, como una asociación de humanos y no humanos que sería la proveedora del MB. Finalmente, están 4) los usuarios de quienes se piensa, desde el campo de producción, que, “probablemente no saben sobre OMB ni cómo OMB funciona” (Schultheiß; Lewandowski, 2021: 2), es decir, desconocen los procesos de optimización de contenidos para la web. Este resultado se obtuvo con una muestra europea y, aunque no se encontraron trabajos similares en Latinoamérica ni en Colombia, según la Organización para la Cooperación y el Desarrollo Económicos (OCDE, 2020), el nivel de alfabetización digital es mayor en Europa, entonces podemos afirmar que sería importante hacer un estudio para conocer la relación de la región con el tema.

En los análisis de las prácticas y estrategias de creación de contenidos también se nombraron otros agentes del proceso como 5) las empresas que generan programas que evalúan los contenidos (las hemos llamado máquinas que simulan a las máquinas lectoras) intentando emular y predecir la manera en la que el MB de Google los leerá y los indexará. Estas se dedican a hacer el ejercicio inverso al que hace Google, ya que tratan de comprender el funcionamiento del MB y hacen simulaciones iterativas para mejorar los indicadores del contenido en la verdadera competencia en los SERP. Luego de esto, convierten las comprensiones de dichos ejercicios en indicadores, datos y herramientas que los generadores de contenido usan en el momento de definir las temáticas, los enfoques y todo lo relacionado con el contenido/texto.

Parte 2. Contexto de producción - contextos de búsquedas

Hemos descrito elementos de la web, como el contexto de distribución de información, centrando la atención en el ejercicio mediador de los “usualmente considerados como intermediarios neutrales que clasifican los resultados objetivamente” (Schultheiß; Lewandowski, 2021: 2): los MB.

Como señaló Ridgway (2023), los usuarios y los contenidos se influyen mutuamente en temas de índole social, económica y tecnológica con los datos de su comportamiento. Para la descripción del argumento de la vigilancia capitalista (*surveillance capitalism*) que defienden varios autores, entre ellos, Zuboff (2019: 3), la situación comunicativa definida se nombra en los términos que siguen:

Google organiza la información del mundo y la hace asequible y usable al entregar resultados relevantes de manera gratuita. En cambio, los usuarios de confianza - *trusted users* están ahora encerrados en, lo que Zuboff considera como un pacto faustiano del siglo XXI, abasteciendo y proveyendo datos *ad infinitum* que perpetúan al Gran Otro.

Aunque este tratamiento puede discutirse, el comportamiento de los usuarios en la web responde a los patrones de este contexto de producción. Un estudio reciente encontró “en alrededor del 50 % de los casos, [que] los usuarios hicieron clic en el primer resultado, con el 97 % de todos los clics hechos en la primera página de los resultados arrojados” por el SERP (Urman; Makhortykh, 2023: 1). Este no es el único estudio con conclusiones similares para describir el comportamiento de lectura en la web.

Así pues, se han descrito algunos elementos del contexto de distribución, ahora se verán los mismos a través de la producción de textos/contenidos y sus procedimientos de optimización.

Optimización para motores de búsqueda: máquina que lee trata de simular el comportamiento de la máquina lectora

Las estrategias de los creadores de contenidos para competir en el espacio digital se denominan optimización para MB o OMB. La posibilidad de aplicar distintas técnicas y obtener un posicionamiento abrió todo un espectro de indicadores y un mercado de productos, servicios y plataformas para competir por la visibilidad que los MB han ofrecido.

Las estrategias de OMB se pueden comprender en dos grandes grupos: el primero es la optimización dentro de la página (*on page optimization*), que consiste en organizar el sitio web desde su naturaleza (si es del sector tecnológico o de las artes tendrá unas estructuras de información distintas y puntajes distintos) y desde sus componentes estructurales:

La estructura html, los títulos, los cuerpos de texto, los hipervínculos, las URL, la frecuencia de ciertas *keywords* y sinónimos, uso de robots txt, de mapas ordenados en la URL, de un ejercicio de investigación extensiva de los competidores para apoyar el trabajo de posicionamiento de las palabras clave, que llevan al usuario a imágenes y textos que tienen metadatos articulados de manera cuidadosa. (Vinit-Kumar; Kumari; Kumar; Appa-Rao, 2012: 3)

Se escribe con un objetivo: que la máquina lectora encuentre el contenido, lo evalúe y le asigne una posición en su *ranking* (SERP). Su tarea es estar siempre explorando la red para agregar al índice un contenido u otro en una posición puntual, y como esto cambia constantemente, esta es continua.

El segundo grupo de estrategias se conoce como optimización fuera de la página (*off page optimization*) y se concentra en hacer las gestiones necesarias y construir los contenidos cuidadosamente organizados para lograr el posicionamiento a través de páginas mejor posicionadas que la propia. Como se ha dicho antes, es una manera de transferir la confianza de una URL a otra y así los contenidos de medios reconocidos se ven más neutrales y orgánicos, menos publicitarios, buscando no romper la experiencia de lectura del usuario que es su objetivo. Con esto quieren decir, además, que el usuario no percibe el ejercicio publicitario o ideológico del que está siendo objeto (Zuboff, 2019).

Este es el medio sobre el cual se efectúa la votación que se nombró antes, en palabras del profesor Felten *et al.* (2004). Se trata de una estrategia para aumentar la probabilidad de indexación de los contenidos (o de ser encontrado por el motor de búsqueda) y luego de mejorar sus posibilidades de escalar hasta ocupar los primeros lugares en sus listas SERP, puesto que allí tienen mayor visibilidad. Esta estrategia específica se llama *link building* o construcción de enlaces tipo *backlink*, como se nombró antes, y consiste en lograr que se genere un enlace de páginas web cuyo cálculo de depósito y generación de confianza sea alto hacia otra web que lo tiene más bajo. En el mercado, esta variable es denominada como autoridad de dominio (*domain authority*). La página con mayor autoridad de dominio puede decidir si transfiere o no su autoridad a la página que le enlaza. Con el tiempo, esto se convirtió en otra estrategia de comercialización para la

transferencia de confianza entre páginas, incluso hay compraventa de estos enlaces en plataformas especializadas.

Aunque esta contabilidad de enlaces no es un indicador directo, “los algoritmos de los motores de búsqueda buscan principalmente comprender el contexto y la intención de búsqueda del usuario” (Duka *et al.*, 2023: 6). Por eso, el usuario ve que el entorno de búsqueda cambia, por ejemplo, a veces muestra elementos que antes no o viceversa, lo que dificulta también el ejercicio de interpretar la información que recibe en sus consultas.

Esta y otras técnicas de OMB surgen de hipótesis, pruebas, experimentos y concursos de posicionamiento de textos/contenidos que realizan investigadores del campo y algunas empresas que han encontrado allí un mercado nuevo (Evans, 2007). Estas pruebas generan una especie de principios heurísticos que no garantizan una *performance* perfecta de los contenidos, pero funcionan.

Además del mercado de empresas que se dedican a esto, se genera un campo de estudios con complejos modelos estadísticos que analizan los algoritmos de MB. Así como hay *big data* recopilada directamente de los comportamientos de los usuarios, también hay gran cantidad de datos sobre el comportamiento de los MB a través de los cuales se mueve la información en la web. Esto ha permitido identificar lo que llaman los sesgos de los MB, los cuales caracterizan el campo de producción de estos contenidos/textos.

Sesgos en los algoritmos: ¿cómo se habla de estos con relación a los MB?

Existen diversos estudios relacionados, algunos son propiamente sobre las subjetividades incluidas en el diseño de los algoritmos de los MB, y otros que, con mayor profundidad, implican la programación misma. El estudio de los resultados en términos sociales y políticos es un campo reciente que se ha nombrado como sociopolítica de los MB (*search engine sociopolitics*) al que se ha sumado otro conocido como ideología algorítmica (Mager *et al.*, 2023). Este último se preocupa por analizar los efectos de los MB en la sociedad y en los comportamientos de las personas.

De acuerdo con lo anterior, las decisiones de diseño del algoritmo *PageRank* y *RankBrain* pueden evaluarse como exitosas desde un punto de vista técnico, sin embargo, estas también fueron, y siguen siendo, espacios de experimentación con los comportamientos de búsqueda de los usuarios y sus formas de obtener información.

Estas decisiones generan las condiciones materiales que influyen las situaciones comunicativas en cuestión y dan lugar a interpretar e identificar

sesgos en los MB. Pero, ¿qué significa que un MB presente sesgos? Se trata de la exposición de los usuarios a contenidos que no quieren o deben recibir, o a contenidos controversiales, estereotipados y contradictorios. Gerhart (como se citó en Díaz, 2008) planteó que los usuarios usan un MB que se alinea con sus propios sesgos, pero Zuboff (como se citó en Ridgway, 2023) propuso que es el usuario quien aprende los sesgos del MB que usa por la relación asimétrica de poder que se da en la interacción. Por esto, se habla de evaluar dichos sesgos en el marco de modelos deliberativos y democráticos en los que las ideas de todos, populares o no, tengan la misma probabilidad de ser encontradas, aunque otras posturas justifiquen los sesgos como resultado de las mismas búsquedas del usuario (Mager *et al.*, 2023).

Quienes han estudiado el comportamiento frente a los MB mostraron que, “sumando a la gran confianza depositada por los usuarios en los motores de búsqueda hay un gran desconocimiento sobre cómo funcionan y de dónde vienen sus ganancias” (Schultheiß; Lewandowski, 2021: 2). Este desconocimiento se suma a otros estudios que se refieren al tipo de ignorancia socialmente construida o de una lógica de la ignorancia que emerge a la par de una del conocimiento y que se obtiene como resultado de las lógicas algorítmicas y comerciales, como explicaron Haider y Rödl (como se citaron en Mager *et al.*, 2023), y que en este artículo se referencian como decisiones de diseño, ya que el ángulo monopolista y corporativo no es el enfoque de este.

De esta forma se va esbozando una subjetividad oculta (Groys, 2014) en estos mediadores de conocimiento, contraria a una objetividad neutral y supuesta en los procesos de las máquinas que leen. Es fundamental saber esto para hacer una interpretación correcta de los resultados que se obtienen en una consulta realizada en este espacio de producción y distribución de información.

Para cerrar, se conectará lo anterior con lo que significa un ejercicio de lectura. Leer ha sido objeto de reflexión desde la aparición del texto. Gadamer (2003: 218), de la corriente de la filosofía hermenéutica, disciplina abocada a la interpretación de estos, propone que debe ejercitarse la comprensión del sentido de todo texto “y es necesario saber comprender así”, evocando una manera de leer que se sabe hacer o se aprende previamente, en la que el lector debe haberse alfabetizado y que tiene una manera, una forma. Así que, para concluir sobre la posibilidad de hacer lectura crítica en este contexto, hay que articular las disciplinas anteriores con la hermenéutica filosófica del autor nombrado y con la alfabetización, en este caso, digital.

Conclusiones: lectura crítica - alfabetización digital o tecnocientífica

El concepto de alfabetización digital fue propuesto por Paul Gilster (1997: 1) quien “considera que la alfabetización digital engloba principalmente la habilidad de adquirir, comprender e integrar información digital (específicamente las habilidades como buscar en internet, navegar el hipertexto y evaluar y componer información digital)”. Así, estableció la diferencia con el manuscrito o el códex, como nos planteó Chartier (1995) con elementos particulares.

Bhatt (2023), indicó que la navegación o el viaje por medios digitales es profundamente personal y diferenciada. Esto implica que las formas de alfabetización deben adaptarse a esta experiencia, así como lo han hecho los MB. Además, sostuvo que la digitalización —como nombró al proceso de creación de la materialidad digital de la que se componen los contenidos/textos en la web— impacta nuestra relación con el lenguaje en todos los sentidos. Por su parte, Groys (2014) planteó un comportamiento extra o metalingüístico que libera a las palabras de sus contextos. Todo esto significa que “la alfabetización de la información en la era del internet no será solo la habilidad de encontrar la información apropiada sino, sobre todo, dominar la interpretación de los marcos de los enunciados” (Després; Cotte, 2006: 5).

En algún momento, los MB fueron considerados portales (Ridgway, 2023), se han nombrado máquinas filosóficas (Groys, 2014), sistemas de recopilación de información (*information retrieval* - IR) y de generación de plusvalía comportamental (*behavioral surplus*) (Zuboff, 2019). Se suman otras interpretaciones de estos como mecanismos de votación (Felten *et al.*, 2004), moderadores (Mager *et al.*, 2023) o como sencillos buscadores (Brin; Page, 1998). Nombres que intentan comprender su papel en la sociedad.

Sean efectivamente uno u otro, su presencia cambia las reglas de lectura e interpretación como las hemos comprendido en el contexto del códex y esto significa, según la invitación de Gadamer, *saber comprender así*. Este es un artículo que pretende arrojar elementos para una lectura crítica, activa, creativa y propositiva que tiene la capacidad de interpretar o asociar exitosamente los contextos de una consulta en un marco de mediaciones tecnológicas útiles, pero que implican un efecto en el significado de lo que se lee.

Internet hoy tiene una forma que será distinta en el futuro, su materialidad es finita, aunque la virtualidad parezca mostrarla infinita, tiene una gobernabilidad distinta de los contenidos presentes y accesibles. Así que, para poder hacer lectura crítica de la información, hay que comprender primero cómo es ese espacio

contenedor de tantas respuestas y cuáles son las fuerzas le dan su forma, es decir, “un lector tiene que aprender a leer y a descifrar estratos superpuestos de signos luego a relacionarlos con el contexto correcto para poder interpretar la situación con la que está siendo confrontado” (Després; Cotte, 2007: 5). En *The Odyssey of Pedagogies of Technoscientific Literacies*, la investigadora Greta Goertz (2021: 8) citó a Gadamer (2003) y planteó la posibilidad de comprender el contexto así:

El peso de lo que encontramos al comprender algo en un evento lingüístico, está en las palabras jugando alrededor y sobre lo que se quiso decir. Los juegos de lenguaje existen donde nosotros, como aprendices [y nunca dejamos de serlo] ascendemos a la comprensión del mundo.

De manera que no se trata de sumar formas a este intrincado de tecnologías, sino de integrar herramientas para comprenderlas aprovechando su utilidad, sus posibilidades al jugar con el lenguaje (Gadamer, 2003) y sus límites. Por esto, dicho contexto de producción y distribución de textos en la web evidencia que, para participar en el juego de visibilidad de contenidos que ofrecen los MB, como componente faltante en la formación o en la alfabetización digital no es suficiente saber usar las herramientas tecnológicas, pues ellas son una experiencia creativa y participativa (Goetz, 2021) que llaman a un usuario activo y crítico en el ejercicio de interpretación en el que siempre se encuentra.

Sumado a lo anterior, debe considerarse que dicha alfabetización no es realizada solamente a través de iniciativas educativas, ya que “los criterios de diseño de las herramientas digitales y las plataformas —sus algoritmos y modelos de negocio— generan formas específicas de alfabetización” en sí mismas (Bhatt, 2023: 9). Esto implica examinar, en los sentidos expuestos por este artículo, los marcos de medición de las capacidades de hacer lectura crítica en el espacio digital, dado que no se encontraron capacidades de interpretación de la información similares a las que se plantean en este contexto en las dimensiones de navegación, búsqueda y filtrado de los marcos de alfabetización propuestos al inicio ni en las relacionadas con su evaluación.

Referencias

1. Bajtín, Mihail Mihajlovič (1982). *Estética de la creación verbal*.
2. Bhargava, Hemant; Feng, Juan (2002). *Paid Placement Strategies for Internet Search Engines*. Trabajo presentado en Proceedings of the 11th International Conference on World Wide Web, WWW, Honolulu, Estados Unidos. <https://doi.org/10.1145/511446.511462>
3. Bhatt, Ibrar (2023). Postdigital Possibilities in Applied Linguistics. *Postdigital Science and Education*, 6, 743-755. <https://doi.org/10.1007/s42438-023-00427-3>
4. Bourdieu, Pierre (2005). *Capital cultural, escuela y espacio social*. Ciudad de México: Siglo XXI.
5. Brin, Sergey; Page, Lawrence (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems*, 30(1-7), 107-117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)
6. Castells, Manuel (2004). *La era de la información: economía, sociedad y cultura* (Vol. 1 Coyoacán: Siglo XXI. s.a. de c.v.
7. Chartier, Roger (1995) *Forms and Meanings-Texts, Performances, and Audiences from Codex to Computer*. Filadelfia: University of Pennsylvania Press.
8. Chartier, Roger (2006). *¿Qué es un texto?* Madrid: Círculo de Bellas Artes.
9. Contreras-Germán, Jorge; Piedrahita-Ospina, Alberto; Ramírez-Velásquez, Iliana (2019). Competencias digitales, desarrollo y validación de un instrumento para su valoración en el contexto colombiano. *Trilogía - Ciencia Tecnología Sociedad*, 11(20). <https://doi.org/10.22430/21457778.1083>
10. Després-Lonnet, Marie; Cotte, Dominique (2007). "Letter to the editor" No more? *Editorial Roles and Models in the Web Era*. Trabajo presentado en *Information, Communication y Society 10th Anniversary Symposium*, ICS, York, UK.
11. Díaz, Alejandro (2008). Through the Google Goggles: Sociopolitical Bias in Search Engine Design. En *Web Search: Multidisciplinary Perspectives* (pp. 11-34), editado por Amanda Spink; Michael Zimmer. Berlin: Springer.
12. Duka, Mariusz; Sikora, Marek; Strzelecki, Artur (2023). From Web Catalogs to Google: A Retrospective Study of Web Search Engines Sustainable Development. *Sustainability*, 15(8), 6768. <https://doi.org/10.3390/su15086768>
13. Evans, Michael (2007). *Analyzing Google Rankings Through Search Engine Optimization Data*. *Internet Research*. Vol. 17 No. 1, pp. 21-37. <https://doi.org/10.1108/10662240710730470>

14. Felten, Edward; Rubin, Aviel; Stubblefield, Adam (2004). *Analysis of voting data from the recent Venezuela referendum*. Recuperado de <https://aceproject.org/ero-en/regions/americas/VE/venezuela-statistical-analysis-recall-referendum-2004>
15. Gadamer, Hans-Georg (2003). *Verdad y método*. Salamanca: Ediciones Sígueme S.A.
16. Gauri-Girish, Jadhav; Vitthal-Gaikwad, Shubhangi; Dhananjay, Bapat (2023). A Systematic Literature Review: Digital Marketing and its Impact on SMEs. *Journal of Indian Business Research*, 15(1), 76-91. <https://doi.org/10.1108/JIBR-05-2022-0129>
17. Gilster, Paul (1997). *Digital Literacy*. Nueva York: Wiley Computer Pub.
18. Goetz, Greta (2021). The Odyssey of Pedagogies of Technoscientific Literacies. *Postdigital Science and Education*, 3(2), 520-545. <https://doi.org/10.1007/s42438-020-00188-3>
19. Groys, Boris (2014). *Volverse público. Las transformaciones del arte en el ágora contemporánea*. Buenos Aires: Caja Negra.
20. Jasanoff, Sheila (2020). Imagined Worlds: the Politics of Future-Making in the Twenty-First Century. En *The Politics and Science of Prevision-Governing and Probing the Future* (pp. 27-44, editado por Andreas Wenger; Ursula Jasper; Myriam Dunn Caveity. Abingdon-on-Thames: Routledge.
21. Lopes, Maria-Immacolata (1999). La investigación de la comunicación: cuestiones epistemológicas, teóricas y metodológicas. *Diálogos de la Comunicación*, 56, 13-27. Recuperado de <https://gc.scalahed.com/recursos/files/r161r/w25092w/MariaVassallo-libre.pdf>
22. Mager, Astrid; Norocel, Ov-Cristian; Rogers, Richard (2023). Advancing Search Engine Studies: The Evolution of Google Critique and Intervention. *Big Data & Society*, 10(2). <https://doi.org/10.1177/20539517231191528>
23. Martín-Barbero, Jesús (2003). *La globalización en clave cultural. Una mirada latinoamericana*. *Animus. Revista Interamericana de Comunicação Midiática*, 1(2). Recuperado de <https://rei.iteso.mx/server/api/core/bitstreams/38fe51bc-b88a-4a10-890c-5898234f47b4/content>
24. O'Reilly, Tim; Strauss, Ilan; Mazzucato, Mariana (2024). Algorithmic Attention Rents: A Theory of Digital Platform Market Power. *Data & Policy*, 6, e6. <https://doi.org/10.1017/dap.2024.1>

25. Organización para la Cooperación y el Desarrollo Económicos - OCDE (2020). *Aprovechar al máximo la tecnología para el aprendizaje y la formación en América Latina*. Paris: OCDE. <https://doi.org/10.1787/ce2b1a62-en>
26. Ridgway, Renée (2023). Deleterious Consequences: How Google's Original Sociotechnical Affordances Ultimately Shaped 'Trusted Users' in Surveillance Capitalism. *Big Data & Society*, 10(1). <https://doi.org/10.1177/20539517231171058>
27. Ritchie, Hannah; Mathieu, Edouard; Roser, Max; Ortiz-Ospina, Esteban (2023). Internet. *OurWorldInData.org*. Recuperado de <https://ourworldindata.org/internet>
28. Rogers, Richard (2023). Algorithmic Probing: Prompting Offensive Google Results and their Moderation. *Big Data & Society*, 10(1). <https://doi.org/10.1177/20539517231176228>
29. 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*, 77(2), 542-557. <https://doi.org/10.1108/JD-07-2020-0127>
30. Sheffield, Jenna-Pack (2020). Search Engine Optimization and Business Communication Instruction: Interviews with Experts. *Business and Professional Communication Quarterly*, 83(2), 153-183. <https://doi.org/10.1177/2329490619890335>
31. Urman, Aleksandra; Makhortykh, Mykola (2023). You are How (and Where) You Search? Comparative Analysis of Web Search Behavior Using Web Tracking Data. *Journal of Computational Social Science*, 1-16. <https://doi.org/10.1007/s42001-023-00208-9>
32. Van Dijck, José (2021). Seeing the Forest for the Trees: Visualizing Platformization and its Governance. *New Media & Society*, 23(9), 2801-2819. <https://doi.org/10.1177/1461444820940293>
33. Vinit-Kumar, Gunjan; Kumari, Monika; Kumar, Amit; Appa-Rao, Allam (2012). Search Engine Optimization with Google. *International Journal of Computer Science Issues (IJCSI)*, 9(1), 206. <http://dx.doi.org/10.13140/2.1.1634.5288>
34. Winner, Langdon (1980). Do Artifacts Have Politics? *Daedalus*, 109(1).
35. Zuboff, Shoshana (2019). Surveillance Capitalism and the Challenge of Collective Action. *New Labor Forum*, 28(1), 10-29. <https://doi.org/10.1177/1095796018819461>

Elena María Gómez-Rico

Ingeniera de Diseño de Producto y magíster en Estudios Humanísticos de la Universidad EAFIT. Docente y gestora cultural y educativa en diversos espacios (museo, programa municipal, centro de pensamiento-think tank y colectivo tipo hackerspace). CvLAC: https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0000110729. Correo electrónico: egomezri@eafit.edu.co