

LAS INDUSTRIAS PERIODÍSTICAS EN LA ERA DE INTERNET: EL CASO DE LOS MEDIOS NOTICIOSOS COLOMBIANOS EN GOOGLE NEWS COLOMBIA

TANIA LUCÍA COBOS*

Universidad Tecnológica de Bolívar, Colombia

tcobos@utb.edu.co

Recibido: 5/12/2019 / Aceptado: 20/3/2020

doi: 10.26439/contratexto2020.n033.4785

Resumen. En la era de los medios noticiosos en internet, cabe analizar la forma como aparecen los medios colombianos en la edición de Google News Colombia en los aspectos de identificación, visibilidad y rol de proveedor. Google News, propiedad de Google, es un agregador de noticias con 72 ediciones internacionales que, de forma automatizada, sin editores humanos, toma fragmentos de noticias de miles de medios, sin pagar por ello, y los organiza, jerarquiza y ofrece a las audiencias a través de la web y aplicaciones. En un *web scraping* realizado desde el 1 de enero al 31 de marzo del 2015, se capturaron 1 222 320 noticias, a partir de las cuales se identificaron 1216 medios noticiosos; de ellos 111 eran colombianos, quienes a su vez aportaron el 63 % de las noticias capturadas y el 42,3 % de las noticias únicas presentes en la edición. Se identificó a ElTiempo.com y Caracol Radio, grandes, tradicionales y populares medios, como aquellos que tuvieron las mayores cuotas de agregación, tanto en noticias capturadas (visibilidad) como en noticias únicas (proveedor). En paralelo, a partir de una revisión documental, se encontró que ElTiempo.com aplicó una política de no mencionar a Google como una forma de protesta por la agregación; en contraposición de un medio pequeño y nativo digital como TuSemanario.com que la celebraba, ya que aparecía junto a grandes medios noticiosos colombianos.

Palabras clave: Google News / Colombia / noticias / agregadores de noticias / industrias periodísticas

* Doctora en Comunicación y Periodismo. Profesora del Programa de Comunicación Social de la Universidad Tecnológica de Bolívar, Colombia (<https://orcid.org/0000-0001-8348-3631>).

JOURNALISM INDUSTRIES IN THE INTERNET ERA: THE CASE OF COLOMBIAN NEWS MEDIA OUTLETS IN GOOGLE NEWS COLOMBIA

Abstract. In the era of news media outlets on the Internet, the way in which the Colombian news media outlets appear in the edition of Google News Colombia concerning providers' identification, visibility and role was analyzed. Google News, owned by Google, is a news aggregator with 72 international editions that, in an automated way, without human editors, takes snippets of news from thousands of news media outlets, without paying for them, and organizes, hierarchizes and offers them to audiences through the Web and apps. After a web scraping carried out between January 1 and March 31, 2015, 1,222,320 news were captured from 1,216 news media outlets, out of which 111 were Colombian. In this setting, Colombian news media outlets contributed 63% of the captured news and 42.3% of the total news presented in this edition. ElTiempo.com and Caracol Radio—large, traditional and popular news media outlets—were identified as those with the highest aggregate shares, both in captured news (visibility) and unique news (provider). In parallel, as a result of a documentary review, it was found that ElTiempo.com adopted a policy of not mentioning Google as a form of protest against the aggregation, as opposed to a small and native digital medium like TuSemanario.com that was delighted to appear together with large Colombian news media.

Keywords: Google news / Colombia / news / news aggregators / news media industry

INDÚSTRIAS JORNALÍSTICAS NA ERA DA INTERNET: O CASO DA MÍDIA COLOMBIANA NO GOOGLE NEWS COLOMBIA

Resumo. Na era dos meios de comunicação na Internet, a forma como os meios de comunicação colombianos aparecer na edição do Google Notícias Colômbia nas áreas de identificação, visibilidade e papel de provedor foi analisada. Google News, de propriedade do Google, é um agregador de notícias com 72 edições internacionais que, de forma automatizada, sem editores humanos, retêm trechos de notícias de milhares de mídias, sem pagar por isso, organizam, hierarquizam e as oferecem ao público através da web e apps. Após uma *web scraping* realizada entre 1 de janeiro e 31 de março de 2015, foram capturados 1 222 320 notícias, das quais 1216 foram identificadas, sendo 111 colombianas, que por sua vez contribuíram com 63 % das notícias capturadas e 42,3 % das notícias exclusivas presentes na edição. Eles foram identificados ElTiempo.com e Radio Caracol, mídia grandes, tradicionais e populares, como aqueles que tiveram as maiores quotas de agregação, tanto notícias capturado (visibilidade) e apenas notícias (fornecedor). Paralelamente, a partir de uma revisão documental, constatou-se que o ElTiempo.com aplicou uma política de não mencionar o Google como uma forma de protesto contra a agregação; ao contrário de um meio digital pequeno e nativo, como o TuSemanario.com, que o celebrou desde que apareceu ao lado de grandes meios de comunicação colombianos.

Palavras-chave: Google News / Colombia / notícias / agregadores de notícias / indústrias de jornais

INTRODUCCIÓN

En las últimas dos décadas del siglo xx, los avances en el desarrollo de los protocolos informáticos TCP/IP¹ permitieron el nacimiento de internet, cuyo servicio más exitoso ha sido la WWW o simplemente la Web. Los promotores de estos avances son una gran variedad de compañías en el ámbito de la electrónica, la informática y las telecomunicaciones, cuyos desarrollos de *hardware* y *software* suscitaron lo que Pérez (2004, p. 44) llama la quinta revolución tecnológica², la era de la informática y las telecomunicaciones, y Estados Unidos se encuentra a la cabeza como protagonista.

Esta nueva era o quinta revolución que estamos viviendo ha traído consigo fuertes y profundos cambios sociales en todos los aspectos de la vida humana, incluyendo, por supuesto, a los medios masivos de comunicación³. “La consolidación de un nuevo medio de comunicación de masas siempre ha conllevado impactos, en mayor o menor escala y con variados resultados, que se han traducido en readaptaciones en las diferentes esferas o ámbitos de las sociedades” (Cobos, 2017b, p. 42).

Inicialmente la Web fue adoptada como una extensión por la prensa, la radio y la televisión, pero, posteriormente, modeló una identidad digital individual y se formaron unidades de producción propias. En paralelo, surgieron los medios nativos digitales, muchos de ellos producto de emprendimientos de comunicadores sociales y periodistas que salían de las filas de los medios tradicionales y se lanzaban a la aventura de crear y administrar su propio medio. A su vez, las compañías de tecnología, por ejemplo, Google⁴, empezaron a incursionar, entre otros aspectos, en la distribución de las noticias en internet, para lo cual crearon servicios como los agregadores de noticias, que se convirtieron en punto de encuentro entre audiencias y medios noticiosos, y cuyas implicaciones para los segundos son mucho más profundas, como se detallará más adelante. Este estudio se centra en este último aspecto: a partir de la identificación de los medios noticiosos colombianos indexados en Google News Colombia, se determina su visibilidad, su rol como proveedores y las relaciones que estos mantenían con el agregador.

1 TCP/IP: TCP (*Transfer Control Protocol*) e IP (*Internet Protocol*) son dos protocolos que trabajan en conjunto para definir cuidadosamente cómo se mueve la información en internet desde el remitente hacia el destinatario y asegurarse de que se cumpla el envío (IBM, s. f.).

2 La primera es la Revolución Industrial; la segunda, la era del vapor y el desarrollo del ferrocarril; la tercera, la era del acero, ingeniería pesada y electricidad; la cuarta, la era del automóvil, el petróleo y la producción en masa (Pérez, 2004, p. 35).

3 Se refiere a aquellos canales o instrumentos cuyos mensajes transmitidos son recibidos de forma simultánea por enormes auditorios, heterogéneos y anónimos, llamados masas.

4 Google es una de las más grandes compañías de tecnología estadounidenses y la principal subsidiaria de Alphabet, que se especializa en productos y servicios relacionados con internet (*software* y *hardware*).

MEDIOS NOTICIOSOS, AGREGADORES DE NOTICIAS Y LAS NOTICIAS

Independientemente de que sean tradicionales o nativos digitales, las industrias de medios se caracterizan, entre otros aspectos, por la producción y distribución de noticias (Steven, 2005, p. 30). Se entiende por estas una información de interés general para la audiencia objetivo, donde criterios como oportunidad, impacto, proximidad, controversia, prominencia, actualidad y extrañeza determinan si se convertirá en una de ellas (Potter, 2006, pp. 2-6). Desde una perspectiva económica, las noticias son el bien que elaboran y difunden las industrias informativas, transformándose en mercancía (Torres, 1985, p. 50).

Profundizando desde esa perspectiva, la noticia es un “bien” raro, entendiéndolo en el sentido de su uso limitado. Su producción no es gratuita; por el contrario, es costosa y para ello se requiere un esfuerzo: la puesta en marcha de cualquier tipo de medio, es decir, de una organización. Esta es también “útil”, al ser un acontecimiento o problema de actualidad que le interesa a un determinado público. Por tanto, la noticia no tiene valor de uso para su poseedor —la empresa periodística—, sino un valor de cambio para este y un valor de uso para otros (Torres, 1985, p. 51).

Así como su producción, su distribución también tiene un costo elevado. La rapidez y eficiencia de esta fase igualmente condiciona el carácter perecedero de las noticias. La distribución requiere una organización sofisticada y costosa para poder dar respuesta a una demanda masiva y ampliamente dispersa (Torres, 1985, pp. 63-64). En este sentido, la llegada de la Web, y la posterior aparición de las aplicaciones, también trajo consigo el surgimiento de nuevas formas de distribución del contenido noticioso, de la mano de las compañías tecnológicas, que incrementan exponencialmente su exposición a diferentes públicos y abaratan costos. Una de estas formas son los agregadores de noticias.

Los agregadores de noticias son un tipo de agregador de contenido que, como su nombre lo indica, se especializa solo en estas. En este artículo se definen

[...] como sitios web de libre acceso que, a través de algoritmos o con editores humanos, constantemente exploran, listan y agrupan por temática, con enlace activo a la fuente original, titulares y primeras líneas o el texto completo, además de fotografías y videos, de cualquier tipo de información noticiosa, provenga esta de sitios de noticias de libre acceso y con muro de pago, y sea que cuenten o no con licencias del propietario del contenido para hacerlo. Estos suelen pertenecer a empresas tecnológicas multinacionales y su funcionamiento no está exento de polémica. (Cobos, 2017b, p. 61)

Un ejemplo de ellos es Google News, un servicio informativo propiedad de Google, con 72 ediciones para diferentes países y en diversos idiomas. Cuenta con una dinámica de mejoramiento continuo, pero también ha suscitado frecuentes y mediáticas polémicas con empresas periodísticas por su actividad.

Google News Colombia es una de estas ediciones. Inicialmente los medios noticiosos colombianos se indexaban, junto a otros iberoamericanos, en la edición Google News España, lanzada en septiembre del 2003 (“Nace el servicio Google News en español”, 2003). Posteriormente, en enero del 2006, se presentó oficialmente la edición Colombia bajo el subdominio: <http://news.google.com.co>.

Google News, que empezó a operar en el 2002, se definía a sí mismo como “Un sitio de noticias generado por ordenador que recopila titulares de más de 50 000 fuentes informativas de todo el mundo, agrupa noticias similares y las presenta de acuerdo a los intereses de cada lector”; posteriormente lo modificó a “Amplia cobertura actualizada de noticias, agregada a partir de fuentes de noticias de todo el mundo” (Google, s. f.). Este agregador rastrea la web, agrega los titulares (antes también incluía las primeras líneas), las fotografías y videos de las noticias, y las categoriza y ordena dependiendo de su temática, fecha de publicación, medio noticioso, entre otros criterios. Si el usuario desea leer la noticia completa, debe hacer clic en el enlace que lo lleva a la página web del medio que la publicó. En este proceso no intervienen seres humanos, manifiesta Google, sino que es enteramente automatizado y ejecutado por el StoryRank, el algoritmo que opera este servicio.

La estructura general de este agregador de noticias está compuesta por ocho canales estándar que recuerdan las secciones de un periódico y facilitan al usuario la exploración temática. Los fragmentos de noticias son agregados a cada una de estas de acuerdo con los temas que aborden. Tales canales son “Noticias destacadas”, “Internacional”, “Nacional” (aparece el nombre del país), “Negocios”, “Ciencia y tecnología”, “Deportes”, “Entretenimiento” y “Salud”. Adicionalmente, si el usuario consulta la edición del país en donde reside, se visualiza la sección “Tus noticias locales”.

La práctica de agregación que hace Google News, conocida como *deep linking* y realizada a través de la automatización sin pedir permiso, por lo tanto, sin pagar licencias por los fragmentos que agrega, ha sido fuertemente criticada por la industria de noticias en diferentes países del mundo: quejas, demandas, leyes, cierre de ediciones, acuerdos, entre otros. Google News ha sido referido como un “parásito”, “plagiario” y “vampiro digital” que succiona la “sangre” de los periódicos (Lee y Chyi, 2015, p. 4) y que se beneficia de la distribución, pero no asume costos por su producción (Cobos, 2017b, p. 107). Para los medios noticiosos, la permanencia en Google News es voluntaria; sin embargo, una desindexación del agregador puede afectar fuertemente su tráfico e incluso su posicionamiento en Google Search. Ejemplo de esto fue la experiencia del conglomerado mediático alemán Axel Springer AG. En el 2014, decidió eliminar voluntariamente cuatro de sus publicaciones (*Die Welt*, *Computer Bild*, *Sport Bild* y *Auto Bild*) por dos semanas; la caída del tráfico desde Google News fue cerca de un 80 % y estimaron fuertes pérdidas de ingresos por publicidad por cada marca; en vista de esto, regresaron (Axel Springer AG, 2014; Moreno, 2014). O, como manifiesta Pinto Balsemão (2014), CEO del Grupo Impresa:

Los operadores de sitios web tienen la posibilidad de autoexcluirse de las búsquedas desde 1994, gracias al Protocolo de Exclusión de Robots, por lo que no hay ningún cambio en esta área. Con este protocolo nos autoexcluimos de toda y cualquier búsqueda; o estamos dentro o estamos afuera, no hay término medio. Mas, reconozcámoslo, la autoexclusión, según lo propuesto por Google, solo puede ser útil si existen alternativas de autoinclusión en otros sistemas. Dado el estado *de facto* de Google como "puerta de acceso a internet", los proveedores de contenido dependen de este para proporcionar los enlaces a dicho contenido, pero no para la presentación de contenido en sí.

En su defensa, Google News destaca que su servicio no monetiza, ya que no muestra publicidad, y que ejerce un papel de *infomediario*, un intermediario entre las noticias del medio y los usuarios que las buscan. Asimismo, manifiesta que no viola leyes de derechos de autor, porque no agrega el contenido íntegro y porque otorga el respectivo crédito a la fuente de donde se han tomado los fragmentos (Pintos, 2010).

La situación descrita es tan solo un ejemplo de los retos y desafíos que afrontan las industrias de los medios de comunicación, en este caso, la noticiosa, ante los avances derivados de los desarrollos de las compañías tecnológicas, particularmente las que tienen a internet como su base de operaciones. Estas se han erigido como los actores detrás de un *quinto poder*⁵.

[El quinto poder está] basado en permanentes desarrollos tecnológicos de *hardware* y *software* que, en conjunto, hacen posible una amplia gama de servicios digitales interconectados en tiempo real y casi siempre a cero costo en términos de dinero; que empodera al ciudadano con instrumentos para hacer escuchar su voz a cambio de permitir el rastreo a través de algoritmos —sepa y quiera este o no— de su información personal, comportamientos y consumos digitales con diversos objetivos. Un poder que, al propiciar una relación de dependencia con otros estamentos de la sociedad, influencia, determina, encausa y controla; un poder que desafía, transforma y reevalúa, entre otros, al *cuarto poder* y a sus grandes y también pequeños actores, y les impone sus condiciones. (Cobos, 2017b, p. 51)

Teniendo en cuenta el marco anterior, para esta investigación se establecieron como objetivos, primero, identificar cuáles eran los medios noticiosos colombianos indexados en la edición de Google News Colombia (teniendo en cuenta que Google no da a conocer esta información); segundo, cómo era su tasa de agregación de noticias (para identificar visibilidad dentro del servicio a partir de las noticias capturadas, y sus roles como proveedores a partir de las noticias únicas); y tercero, indagar cómo eran las relaciones que mantenían estos medios indexados con el agregador.

5 Los tres primeros poderes los ostenta el Estado (Legislativo, Ejecutivo y Judicial) y el *cuarto poder* son los medios de comunicación de masas, identificación que se le atribuye a Edmund Burke en 1787 (Cobos, 2017b, p. 47).

METODOLOGÍA

Esta investigación empleó una metodología mixta, es decir, tanto técnicas cuantitativas como cualitativas de acuerdo a los objetivos estipulados. El enfoque, a su vez, fue exploratorio y descriptivo, dada la escasa literatura que existe en Colombia respecto al tema en cuestión. Los estudios de alcance exploratorio “se realizan cuando el objetivo es examinar un tema o problema de investigación poco estudiado, del cual se tienen muchas dudas o no se ha abordado antes” (Hernández, Fernández y Baptista, 2010, p. 79) y los de alcance descriptivo buscan “especificar propiedades, características y rasgos importantes de cualquier fenómeno que se analice” (Hernández, Fernández y Baptista, 2010, p. 80).

Con relación a la parte cuantitativa, guiada por los objetivos de identificación, visibilidad y rol de proveedor, se utilizaron los llamados métodos digitales o *digital methods*. Esto se refiere al uso de programas propios de la informática y la computación para la captura, visualización y análisis de datos de fenómenos que se suceden en internet o de estudiar objetos propios de este medio, con el objetivo de ser examinados desde la perspectiva de las ciencias sociales, en este caso, de la comunicación. Cabe mencionar que la automatización del proceso no descarta que en algún momento se requiera la intervención humana, particularmente en la revisión de los datos crudos.

La aplicación de *digital methods* se hizo a través del *web scraping*, una técnica informática que permite que los datos se obtengan a partir del “raspado” automatizado que se hace de un documento HTML (una página web), por ejemplo, y se almacenen en un documento estructurado, un archivo XLS (de Microsoft Excel), por ejemplo. Para ello se contó con el apoyo de un desarrollador de *software*, quien de acuerdo con las instrucciones dadas por la investigadora diseñó un *scraper bot*⁶ en PHP⁷ para la ejecución de esta tarea. El *scraper bot*⁸ visitó el sitio web de Google News Colombia cada hora entre el 1 de enero y el 31 de marzo del 2015, fechas que obedecieron a un muestreo no probabilístico por conveniencia; capturaba todos los fragmentos de noticias que encontraba por cada canal (ejemplos: “Entretenimiento”, “Deportes”, “Ciencia y tecnología”, etcétera), y los almacenaba en una base de datos en MySQL⁹ que, posteriormente, para su análisis, se exportó a Microsoft Excel¹⁰. Las variables capturadas fueron canal, titular, nombre del medio, enlace de la noticia, fecha y hora de captura.

6 Un *scraper bot* es el programa informático diseñado en PHP que ejecutó la tarea de *web scraping*.

7 PHP es un lenguaje de programación orientado al desarrollo de aplicaciones web dinámicas con acceso a información almacenada en una base de datos, como MySQL.

8 Para conocer más sobre estos programas se recomienda visitar la Digital Methods Initiative, donde aparece un modelo de Google News Scraper (<https://wiki.digitalmethods.net/Dmi/ToolGoogleNewsScraper>).

9 MySQL es un sistema de gestión de bases de datos.

10 Microsoft Excel es un *software* de hojas de cálculo empleado para análisis estadísticos de datos.

En cuanto a la parte cualitativa, guiada por la indagación de antecedentes de las relaciones entre los medios noticiosos colombianos y este agregador de noticias, se procedió a una revisión documental en línea en búsqueda de literatura, particularmente noticias, en cualquier fecha de publicación, que diera testimonio sobre la posición de los medios noticiosos colombianos con respecto a su presencia en Google News, permitiendo identificar si había tensiones, neutralidad o encumbramientos. Esta etapa se cumplió en las mismas fechas que la parte cuantitativa y se hizo de forma manual a través de búsquedas en Google Search utilizando diversos criterios.

Como principal limitación cabe mencionar el cierre de la edición de Google News España en diciembre del 2014, que eliminó, al menos en su gran mayoría, a los medios noticiosos españoles de otras ediciones en español, como ocurrió en Google News Colombia. Esto pudo implicar un cambio en el ecosistema noticioso, aunque es imposible saber de qué magnitud.

Finalmente, los datos crudos de esta investigación se encuentran disponibles para su libre uso en el Digital Depòsit de Documents (DDD) de la Universitat Autònoma de Barcelona. En la bibliografía se halla el enlace respectivo.

LOS MEDIOS NOTICIOSOS COLOMBIANOS EN GOOGLE NEWS COLOMBIA

El *scraper bot* capturó en total 1 222 320 noticias en la edición de Google News Colombia entre el 1 de enero y el 31 de marzo del 2015. Con los datos organizados en hojas de cálculo en Microsoft Excel, se procedió a hacer una revisión o curaduría humana que permitió detectar anomalías de origen, es decir, presentes en la fuente (fueran causadas por el algoritmo o por inserción humana de datos de identificación). Estas fueron las siguientes: el mismo medio usando más de un nombre, medios diferentes usando un mismo nombre, diferentes ediciones de un medio bajo un mismo nombre, nombres desactualizados, nombre de medios con enlaces que no les corresponden, inexactitudes en etiquetación de fuente y errores de ortografía.

Una vez subsanadas manualmente tales anomalías, se llevó a cabo el análisis implementando funcionalidades como filtros, ordenamientos, quitar duplicados y tablas dinámicas. Se identificaron en total 1216 medios noticiosos procedentes de 43 países distintos (y cinco de procedencia indeterminada). En la tabla 1 se aprecia que en primer lugar se ubicó México, con una cuota de procedencia del 27,96 %; seguido por Argentina, con 20,06 %, y Colombia en un lejano tercer lugar, con 9,12 %. En otras palabras, a pesar de que la edición manifiesta ser de Colombia, predominó fuertemente la indexación de medios noticiosos mexicanos y argentinos.

Pero, desde la perspectiva de noticias capturadas, variable que buscó establecer qué tan visible era el medio noticioso en la edición (que implicó el conteo total de noticias,

incluyendo duplicadas), se encontró que los medios colombianos marcaron el más alto porcentaje; es decir, el 63 % de las 1 222 320 noticias capturadas eran de medios de este país, mientras que los medios mexicanos tuvieron 8,87 % y los argentinos 5,16 %. Desde la perspectiva de noticias únicas, variable que buscó determinar el rol de proveedor del medio noticioso en la edición (que se obtuvo al contarlas por una sola vez al eliminarse las duplicadas), se encontró que nuevamente los medios colombianos fueron los mayores proveedores con un 42,30 % de las 175 502 noticias únicas, ubicándose los mexicanos con un 13,60 % y los argentinos con un 8,23 %.

Tabla 1
País de procedencia de los medios en Google News Colombia

Orden	País de procedencia	Medios	Noticias capturadas	Noticias únicas
1	México	340	108 446	23 884
2	Argentina	244	63 175	14 461
3	Colombia	111	770 663	74 240
4	Estados Unidos	80	34 136	7 205
5	Perú	60	54 842	12 609
6	Chile	54	9 388	2 571
7	Venezuela	54	52 943	11 741
8	Cuba	35	18 237	3 405
9	España	30	5 762	1 693
10	Bolivia	24	11 532	2 587
11	Rep. Dominicana	21	1 510	272
12	Uruguay	20	6 029	1 536
13	Ecuador	19	12 326	2 563
14	Panamá	13	3 689	789
15	Paraguay	13	6 005	1 409
16	Costa Rica	11	8 328	1 629
17	El Salvador	9	2 930	724
18	Guatemala	8	2 444	701
19	Reino Unido	8	6 754	1 283
20	Honduras	7	12 076	3 139
21	Italia	6	1 656	416
22	Francia	5	2 728	707
23	Puerto Rico	5	1 210	239
24	China	4	8 215	1 530
25	Nicaragua	4	3 427	922
26	India	3	5	3
27	Suiza	3	2 745	759

(continúa)

(continuación)

28	Australia	2	3	2
29	Canadá	2	56	11
30	Pakistán	2	2	2
31	Singapur	2	2	2
32	Alemania	1	511	125
33	Bélgica	1	16	2
34	Brasil	1	22	2
35	Corea del Sur	1	63	22
36	Grecia	1	9	3
37	Irán	1	604	173
38	Israel	1	40	12
39	Japón	1	50	2
40	Nueva Zelanda	1	3	1
41	República Checa	1	12	4
42	Rusia	1	410	115
43	Vietnam	1	121	29
	Procedencia indeterminada	5	9 195	1 978
	Total	1 216	1 222 320	175 502

Elaboración propia

Se destaca que en la edición Colombia los medios colombianos se ubicaron terceros en país de procedencia (9,12 %), aunque primeros en noticias capturadas (63 %) y noticias únicas (42,30 %). El primer hecho puede deberse a que México y Argentina son países con mayor extensión territorial que Colombia; por tanto, cuanto más grande sea el país, mayor cantidad de medios noticiosos tiene. Lo segundo refuerza el carácter local de la edición Colombia, pues el StoryRank da predominancia de agregación a noticias de medios de este país por encima de cualquier otro.

En la tabla 2 se aprecia una vista detallada de la agregación de noticias de los medios noticiosos colombianos en cuanto al volumen de noticias capturadas, que identifica su visibilidad dentro del servicio. La tabla 3, por su parte, ofrece una vista detallada del volumen de noticias únicas, que identifica su rol de proveedor. A partir del porcentaje de noticias capturadas y del porcentaje de noticias únicas correspondiente a cada fuente noticiosa indexada, se obtuvo la razón de cambio porcentual en cada caso y, con esto, se determina la clasificación de los medios noticiosos en diferentes grupos.

Tabla 2
Visibilidad de medios noticiosos colombianos en Google News Colombia

Rango de noticias capturadas	% de noticias capturadas	Número de medios	Razón de cambio porcentual	Grupo	Medios en el rango
85 001 - 90 000	7,34 %	1	7,34	Muy alta visibilidad	ElTiempo.com
75 001 - 80 000	12,70 %	2	6,35	Alta visibilidad	Caracol Radio (6,44 %), ElEspectador.com (6,25 %)
55 001 - 60 000	4,75 %	1	4,75	Intermedia visibilidad	El Colombiano, Radio Santa Fe, Vanguardia Liberal
50 001 - 55 000	4,10 %	1	4,10		
40 001 - 45 000	3,29 %	1	3,29		
25 001 - 30 000	11,68 %	5	2,33	Baja visibilidad	Pulzo, El Universal - Colombia, El Heraldo (Colombia), Terra Colombia, Semana.com, RCN Radio (comunicado de prensa, blog), W Radio, El Universal, LaRepública.com.co, El País - Cali, Colombia, Portafolio.co, Blu Radio, No colombianos: Terra Perú, EntornoInteligente, Prensa Latina
20 001 - 25 000	1,67 %	1	1,67		
15 001 - 20 000	12,47 %	9	1,38		
10 001 - 15 000	4,68 %	5	0,93	Muy baja visibilidad	Otros (colombianos y no colombianos)
5 001 - 10 000	11,91 %	22	0,54		
* 0 - 5 000	25,35 %	1 168	0,02		
1 222 320	100 %	1 216			

* El valor 0 corresponde a elEconomista.es

Elaboración propia

Tabla 3

Rol de proveedor de medios noticiosos colombianos en Google News Colombia

Rango de noticias únicas	% de noticias únicas	Número de medios	Razón de cambio porcentual	Grupo	Medios en el rango
9 001 - 10 000	5,29 %	1	5,29	Grandes proveedores	Caracol Radio, ElTiempo.com
8 001 - 9 000	4,88 %	1	4,88		
6 001 - 7 000	3,65 %	1	3,65	Medianos proveedores	ElEspectador.com, El Colombiano, Terra Colombia
5 001 - 6 000	6,23 %	2	3,11	Pequeños proveedores	Vanguardia Liberal, No colombianos: El Universal, Terra Perú, EntornoInteligente
4 001 - 5 000	2,41 %	1	2,41		
3 001 - 4 000	5,88 %	3	1,96		
2 001 - 3 000	11,96 %	8	1,49	Micro-proveedores	Otros (colombianos y no colombianos)
1 001 - 2 000	19,29 %	23	0,83		
* 0 - 1 000	40,38 %	1 176	0,03		
175 502	100 %	1 216			

* El valor 0 corresponde a elEconomista.es

Elaboración propia

Al cruzar la información de las tablas 2 y 3, se descubrió, en primera instancia, que los 1216 medios noticiosos encontrados en la edición de Google News Colombia se concentraban masivamente en el rango de muy baja visibilidad y en el rol de microproveedores. En el detalle de los 111 medios noticiosos colombianos hallados en la edición, solo 7 gozaban de intermedia a muy alta visibilidad y 5 se ubicaban como medianos a grandes proveedores.

En esa misma dirección, en una vista detallada, en muy alta visibilidad (entre 85 001 a 90 000 noticias capturadas, el 7,34 %) y como gran proveedor (entre 8001 a 9000 noticias únicas, el 4,88 %) se ubicó el diario ElTiempo.com, la versión digital del diario generalista *El Tiempo*, fundado en 1911, de circulación nacional y propiedad del conglomerado de medios Casa Editorial El Tiempo (empresa de la Organización Luis Carlos Sarmiento Angulo), que lanzó su versión en línea en 1996.

En segundo lugar, se destacó la posición de Caracol Radio, en alta visibilidad (entre 75 001 a 80 000 noticias capturadas, el 6,44 %) y como gran proveedor (entre 9001 a 10 000 noticias únicas, el 5,29 %). Caracol Radio es un sistema de estaciones de radio de cubrimiento nacional, fundado en 1948 y presente en internet desde 1998-2000; es propiedad de la empresa multinacional de medios Grupo PRISA.

Por otra parte, en lo que respecta a los hallazgos de la revisión documental, en el 2008, Francisco Forero, entonces representante de Google en Colombia, se manifestó acerca de si los periódicos debían temer la presencia de Google:

Yo no diría temerle. En el caso puntual de los periódicos, por el contrario, tenemos alianzas muy interesantes. Por ejemplo, con la Casa Editorial El Tiempo hay un acuerdo para compartir los ingresos provenientes de publicidad en internet mediante un mecanismo llamado AdSense [...]. Pero, en concreto, en el caso de los periódicos y los grandes medios, lo que tenemos es varias formas de asociarnos. Nosotros no somos un editor de información, sino un intermediario de esta. Lo que hacemos es poner esa información, organizada, a disposición de cientos de usuarios [...]. Google es muy respetuoso de los derechos intelectuales y derechos de propiedad de ese contenido. Los dueños de la información y otros jugadores de la industria han entendido que es un beneficio que buscadores como Google enlacen sus sitios. Lo que vamos es a generar más tráfico hacia esos lugares. Un ejemplo: en Portafolio.com.co sale una noticia, y alguien está buscando en Google noticias sobre la economía colombiana y el primer resultado de la búsqueda es Portafolio.com.co. Entonces, la persona que hace clic en ese resultado va directamente a Portafolio.com.co. Yo no creo que a Portafolio.com.co le moleste mucho que Google le lleve tráfico hacia su sitio. De eso se trata Google News. Creo que ya hay un entendimiento importante por parte de los productores de contenido sobre el hecho de que eso es un beneficio. ("Google reconoce que vino a Colombia hasta por el mercado publicitario de las pymes", 2008)

Aun así, se encontró evidencia de que hubo medios noticiosos del país que exigieron a Google una compensación monetaria por la agregación de noticias en su servicio, aunque tal pretensión no prosperó. En el 2014, con el cierre de Google News España, Luis Fernando Santos (2014a, 2014b), expresidente de El Tiempo, manifestó que durante un período de tiempo, sin referenciar fechas, este periódico aplicó la política de no mencionar a Google en sus noticias como una forma de protesta por la agregación a cero costo: "Ningún medio [colombiano] mencionó que Google ya les había respondido a los medios impresos en Colombia que no pagaba por la agregación, y que quienes no quisieran ser agregados, le informaran para excluirlos. Siguen todos en Google News".

En contraposición a la postura de este gran, tradicional y popular medio noticioso, se expresa la de un pequeño y poco conocido medio, TuSemanario.com, un diario nativo digital con sede en la ciudad de Neiva. Alejandro Cabrera (2012), su director, declaró:

Desde hace un par de meses, Google News (la plataforma de noticias de Google) incluyó al periódico dentro de su grupo de fuentes de noticias de Colombia, para que millones de navegantes en el mundo puedan encontrar noticias de interés de Neiva, del Huila y de Colombia. Así pues, TuSemanario.com se encuentra ya en Google News a la par con la *Revista Semana*, el diario *El Tiempo*, Caracol Radio y otros grandes pesos pesados de la información.

En la revisión documental se halló, asimismo, la confrontación que se suscitó entre algunos medios noticiosos colombianos y el agregador y curador de noticias local Pulzo.com, justamente por esta labor de agregación de noticias. Aunque el funcionamiento de Pulzo.com difiere del de Google News, se considera relevante mencionar el suceso. En el 2014, la Asociación Colombiana de Editores de Diarios y Medios Informativos (Andiarios), de la que forma parte *El Tiempo*, notificó a Pulzo.com para que no incluya en su servicio los contenidos generados por sus socios sin autorización y sin consultarles (“Andiarios pide a Pulzo no utilizar contenido de sus medios sin autorización”, 2014):

Los contenidos constituyen la esencia de los medios de comunicación y su producción es el fruto del inmenso esfuerzo que realizan nuestros reporteros, fotógrafos, diseñadores y editores, así como de las grandes inversiones en tiempo y dinero que realizan las empresas editoras. Solo quienes dirigimos las empresas que tienen por objeto la generación y edición de contenidos tenemos la potestad de decidir dónde y cuándo se publican los mismos.

En respuesta, Guillermo Franco (“Medios tradicionales colombianos piden a Pulzo.com ‘abstenerse de incluir sus contenidos’”, 2014), entonces director de Pulzo.com, manifestó “extrañeza por una solicitud que va en contravía no solo de la realidad tecnológica, a la que Colombia no es ajena, sino de las buenas prácticas de agregación y curaduría de contenidos [...] que ha caracterizado a Pulzo.com”. Además, contabilizando más de dos mil fuentes usadas cotidianamente, “Pulzo.com, literalmente, ‘agrega valor’ de muchas formas, entre otras: complementando, sintetizando, confrontando versiones, comentando, contextualizando, explorando nuevos ángulos, haciendo versiones en español de diferentes lenguas, etcétera”. Cerraba mencionando que le llamaba la atención que los medios pertenecientes a Andiarios habían sido agregados durante años por grandes jugadores internacionales como Google o Yahoo!, sin haber ameritado una reacción como la plasmada en la notificación.

Ese mismo año, Nora Sanín, entonces directora ejecutiva de Andiarios, afirmó en torno a los derechos de los periodistas y de los periódicos frente a Google:

Esta es una polémica que está muy viva en Europa, Brasil, Colombia. Estamos luchando por el respeto de los contenidos generados por nuestros periódicos. El periodismo de calidad necesita y exige el apoyo del Estado. Necesita y exige el respeto por los derechos de autor, por la competencia leal. Consideramos que los agregadores de noticias, como Google, están obligados a, primero, solicitar autorización a quienes generan contenidos para su utilización y, segundo, a remunerar, a pagar una remuneración justa por su uso. (Citada por Ruiz, 2014)

En el 2015, Sanín manifestó:

Los editores enviaron una carta a dicho agregador [Pulzo.com] reclamándole por la inclusión de sus contenidos sin pagar y se está iniciando demanda por infracción

a la propiedad intelectual. Para el caso de Google, se envió una carta requiriéndole que en cada caso se tratara con los editores. (Citada por Sarmiento, 2015)

Mencionó que se veían tres opciones: (1) demandar, (2) retirarse de Google o (3) legislar para que Google pagara por los contenidos. Señaló que los editores colombianos habían llegado al consenso de que Google debía pagar por los contenidos, pero no había uno respecto a retirarse de Google, ya que se requerían más elementos de juicio, y cerró mencionando que en Colombia no se contemplaba ninguna acción legislativa.

CONCLUSIONES

Tal como se parafraseó páginas atrás, desde una perspectiva económica, la noticia es el “bien” que producen las industrias periodísticas y para ellas tiene un valor de cambio (Torres, 1985, p. 51). Ante el arribo de las compañías de tecnología en el ámbito de la distribución digital de las noticias, en este caso, Google y su servicio Google News, se genera una tensa relación entre ambas industrias, donde solo las primeras asumen los costos de producción y las segundas solo los beneficios que obtienen de la distribución.

En esta investigación se encontró que, de los 1216 medios noticiosos indexados en Google News Colombia, solo 111 (9,12 %) eran de origen colombiano, pero, al cuantificar la tasa de agregación de noticias, se demostró que estos tuvieron el más alto porcentaje en visibilidad (63 % de las noticias capturadas), así como el más alto porcentaje como proveedores (42,30 % de las noticias únicas); de esta manera, se logró que la edición cumpliera con su carácter local.

Al efectuar una vista detallada del total de los medios noticiosos indexados en esta edición, se observó que se ubicaban masivamente en el rango de muy baja visibilidad y como microproveedores. El 96,05 % tenía menos de 5000 noticias capturadas y el 96,71 %, menos de 1000 noticias únicas. En el otro extremo de la pirámide, solo el 0,08 % se ubicó en muy alta visibilidad y el 0,16 % como grandes proveedores. A partir de esto se infiere que aunque hubo un heterogéneo ecosistema noticioso (medios tradicionales migrantes, medios nativos digitales; grandes y populares, pequeños y poco conocidos; de prensa, televisivos, radiofónicos, agencias de noticias, entre otros) y que a primera vista podría entenderse como diversidad, en realidad, la agregación de noticias fue muy desigual y que el algoritmo StoryRank tuvo una fuerte y marcada preferencia de agregación por aquellos medios que son grandes productores de noticias, es decir, aquellos que cuentan con condiciones económicas y de infraestructura que les permiten producir con frecuencia, en gran volumen y con poca probabilidad de cesar tal actividad. Esto así, garantiza a Google News Colombia tener un inventario fresco de noticias y en permanente actualización.

Los medios que se ubicaron en la mencionada punta de la pirámide fueron el diario ElTiempo.com y la emisora Caracol Radio, ambos medios noticiosos colombianos tradicionales migrantes (sus ediciones fuera de internet tienen cubrimiento nacional y sus ediciones en línea superan los 20 años), grandes (el primero pertenece a la Organización Luis Carlos Sarmiento Angulo y la segunda al Grupo PRISA) y muy populares (sus sitios web tienen un alto tráfico en el país, especialmente el primero).

ElTiempo.com evidencia en la edición de Google News Colombia la tirantez de las relaciones existentes entre los medios noticiosos y las compañías tecnológicas por la distribución digital de noticias y la ausencia de pago en dinero. Ese diario manifestó en algún momento su desaprobación a la agregación hecha por este servicio y, en paralelo, también fue el medio noticioso con la mayor cantidad de noticias agregadas, tanto en noticias capturadas como en noticias únicas, es decir, en las primeras, tuvo una alta probabilidad de recibir tráfico referido desde Google News Colombia, y en las segundas, fue el más grande proveedor de noticias para la edición. En contraposición, TuSemana.com (identificado como *TU SEMANARIO* en el agregador) celebraba su indexación junto a los grandes medios noticiosos del país, a la vez que se ubicaba en la franja de muy baja visibilidad con el 0,11 % de noticias capturadas y como microproveedor con el 0,07 %, es decir, probablemente con muy bajo tráfico referido desde el servicio. Estos hallazgos dan pie para inferir que, a mayor tráfico referido desde Google News, los medios noticiosos parecen incrementar su interés hacia este y buscan obtener compensaciones adicionales por el uso de su contenido; a menor tráfico referido, se conforman con los escasos clics que llegan y carece de relevancia si este los indexa o no.

Pero los titulares de noticias de diversas fuentes agrupadas en un solo lugar convierten a Google News en un concentrador de audiencias, y su labor de intermediación (y de paso el envío de tráfico), a la que hacían referencia Forero ("Google reconoce que vino a Colombia hasta por el mercado publicitario de las pymes", 2008) y Pintos (2010), es algo de carácter secundario. Desde esta perspectiva, Google News es para Google una herramienta que rentabiliza de forma indirecta, no genera ingresos, pero sí le permite identificar los consumos temáticos de sus usuarios (independientemente de si estos hacen clic o no sobre la noticia), información que se suma a la que captura a través de sus otros servicios y que finalmente monetiza, entre otros, a través de sus conocidos servicios de publicidad digital.

A modo de reflexión final, cabe recordar lo que mencionaba McCombs (1996):

Las noticias influyen en muchas facetas de nuestra vida cotidiana. Nuestro modo de vestir para ir al trabajo, el camino que elegimos a veces para llegar a él, los planes del próximo fin de semana, nuestros sentimientos generales de bienestar o de inseguridad, el enfoque de nuestra atención hacia el mundo más allá de la experiencia inmediata y nuestras preocupaciones sobre los temas del día [...]. De vez en

cuando un dictamen de las noticias trastorna nuestro comportamiento global de forma instantánea y por completo. (p. 13)

Por tanto, la presencia de los grandes, populares y tradicionales medios noticiosos, que tienen la mayor cantidad de noticias agregadas y que de paso son también los principales proveedores, sienta un sesgo, causado por razones técnicas, hacia la información que están consumiendo los usuarios, y esto tiene claras implicaciones sociales o, en palabras de Segev (2008), “la popularidad de Google News y su expansión global lo hacen un influyente canal que puede tener importantes implicaciones en la forma como las personas perciben el mundo” (p. 1).

REFERENCIAS

- Andiarios pide a Pulzo no utilizar contenido de sus medios sin autorización. (30 de mayo del 2014). *El Espectador*. Recuperado de <https://www.elespectador.com/entretenimiento/medios/andiarios-pide-pulzo-no-utilizar-contenido-de-sus-medio-articulo-495613>
- Axel Springer AG. (5 de noviembre del 2014). Axel Springer schließt Datendokumentation ab: Gravierender Schaden durch verschlechterte Suchanzeigen bei Google. Recuperado de <https://www.axelspringer.com/de/presseinformationen/axel-springer-schliesst-datendokumentation-ab-gravierender-schaden-durch-verschlechterte-suchanzeigen-bei-google>
- Cabrera, A. (23 de agosto del 2012). Ahora TuSemanario.com en Google News. *TuSemanario.com*. Recuperado de http://www.tusemanario.com/noticia/ahora-tusemanario-com-en-google-news_5000
- Cobos, T. (2017a). *Dataset: noticias de Google News edición Colombia, 2015*. Recuperado de https://ddd.uab.cat/pub/dadrec/2017/182292/Noticias_Google_News_Colombia_Ene-Mar2015_v1_0.xlsx
- Cobos, T. (2017b). *Medios de comunicación iberoamericanos y agregadores de noticias: análisis a las ediciones de Google News Brasil, Colombia, España, México y Portugal* (tesis doctoral). Universitat Autònoma de Barcelona, España. Recuperado de <https://ddd.uab.cat/record/188096>
- Google. (s. f.). Acerca de Google Noticias. Recuperado de https://www.google.com/intl/es/about_google_news.html
- Google News Colombia. (s. f.). Recuperado de <http://news.google.com.co>
- Google reconoce que vino a Colombia hasta por el mercado publicitario de las pymes. (27 de mayo del 2008). *Portafolio*. Recuperado de http://www.portafolio.co/detalle_archivo/CMS-4211011

- Hernández, R., Fernández, C., y Baptista, P. (2010). *Metodología de la investigación* (5.ª ed.). México D. F.: McGraw-Hill.
- IBM. (s. f.). *Protocolos TCP/IP*. Recuperado de https://www.ibm.com/support/knowledgecenter/es/ssw_aix_72/network/tcpip_protocols.html
- Lee, A., y Chyi, H. (2015). The rise of online news aggregators: consumption and competition. *International Journal on Media Management*, 17(1), 3-24.
- McCombs, M. (1996). Influencia de las noticias sobre nuestra imagen del mundo. En J. Bryant y D. Zillmann (Comps.), *Los efectos de los medios de comunicación. Investigaciones y teorías* (pp. 13-34). Barcelona: Paidós.
- Medios tradicionales colombianos piden a Pulzo.com "abstenerse de incluir sus contenidos". (30 de mayo del 2014). *Pulzo*. Recuperado de <http://www.pulzo.com/medios/medios-tradicionales-colombianos-piden-pulzocom-abstenerse-de-incluir-sus-contenidos/146736>
- Moreno, M. (16 de diciembre de 2014). El ejemplo de Axel Springer: ¿cómo afectará a los medios españoles la salida de Google News? *TreceBits*. Recuperado de <http://www.trecebits.com/2014/12/16/el-ejemplo-de-axel-springer-como-afectara-a-los-mediosespanoles-la-salida-de-google-news>
- Nace el servicio Google News en español. (30 de septiembre de 2003). *El Mundo*. Recuperado de <http://www.elmundo.es/navegante/2003/09/30/empresas/1064923609.html>
- Pérez, C. (2004). *Revoluciones tecnológicas y capital financiero: la dinámica de las grandes burbujas financieras y las épocas de bonanza*. México D. F.: Siglo Veintiuno Editores.
- Pinto Balsemão, F. (18 de junio de 2014). As práticas comerciais injustas do Google. Público. Recuperado de <http://www.publico.pt/economia/noticia/as-praticas-comerciais-injustas-do-google-1659055>
- Pintos, L. (29 de abril de 2010). Google News es parte de la solución, no del problema [mensaje en un blog]. Recuperado de <http://233grados.lainformacion.com/blog/2010/04/google-news.html>
- Potter, D. (2006). *Manual de periodismo independiente*. Recuperado de https://photos.state.gov/libraries/amgov/30145/publications-spanish/handbook_journalism-sp.pdf
- Ruiz, H. (21 de septiembre del 2014). Quieren poner de rodillas a la prensa. *ABC Color*. Recuperado de <http://www.abc.com.py/edicionimpresa/politica/quieren-poner-de-rodillas-a-la-prensa-1288073.html>
- Santos, L. F. (1 de junio de 2014a). Era de esperarse reacción de medios tradicionales con Pulzo.com. *Pulzo*. Recuperado de <http://www.pulzo.com/opinion/era-de-esperarse-reaccion-de-medios-tradicionales-con-pulzocom/147866>

- Santos, L. F. (15 de diciembre del 2014b). Salen medios españoles de Google News, que ya había rechazado pagar a diarios colombianos. *Pulzo*. Recuperado de <http://www.pulzo.com/opinion/salen-medios-espanoles-de-google-news-que-ya-habia-rechazado-pagar-diarios-colombianos/258741>
- Sarmiento, A. (2015). *Relatorio del encuentro "El papel de la SIP ante los desafíos de Google y otros agregadores de contenido digitales"*. Lima: Sociedad Interamericana de Prensa (SIP).
- Segev, E. (2008). The imagined international community: dominant American priorities and agendas in Google News. *Global Media Journal*, 7(13). Recuperado de <http://www.globalmediajournal.com/open-access/the-imagined-international-community-dominant-american-priorities-and-agendas-in-google-news.php?aid=35252>
- Steven, P. (2005). *Dominatrix: la influencia de los medios de comunicación*. Barcelona: Intermón Oxfam.
- Torres, J. (1985). *Economía de la comunicación de masas*. Madrid: Grupo Cultural Zero.