



Noticias falsas en Internet: la estrategia para combatir la desinformación

Diciembre 2017

Facultad de Derecho

Centro de Estudios en Libertad
de Expresión y Acceso a la Información

UP
**Universidad
de Palermo**

Noticias falsas en Internet: la estrategia para combatir la desinformación

Este documento fue elaborado por Carlos Cortés y Luisa Isaza** para el CELE*

I. Introducción

El 4 de diciembre de 2016, un mes después de las elecciones presidenciales estadounidenses, Edgar Maddison Welch, un hombre de 28 años de Carolina del Norte, disparó tres veces contra la pizzería Comet Ping Pong en Washington D.C. Estaba decidido a entrar, investigar y rescatar a los niños que estaban siendo explotados por una supuesta red de tráfico sexual infantil dirigida por Hillary Clinton y otros miembros del Partido Demócrata.¹

Maddison Welch se había convencido de la existencia de la red a partir de las noticias que había leído en Internet. Luego de su arresto, y a pesar de que la policía desmintió la historia –conocida como *Pizzagate*–, el hombre se disculpó pero nunca admitió que la información que motivó su ataque era falsa.² La descabellada historia fue difundida por redes sociales y foros de discusión en Internet, junto con cientos de historias falsas relacionadas con los dos candidatos o con miembros de sus partidos.³

Semanas antes de las elecciones, millones de personas vieron en su *newsfeed* de Facebook una noticia en la que se informaba que, en una declaración sin precedentes, el papa Francisco había proclamado su respaldo a la candidatura de Donald Trump. Esta noticia falsa recibió 960.000 interacciones en la red social (comentarios, reacciones y compartidos), más que ninguna otra noticia real sobre las elecciones.⁴ De hecho, según un análisis de BuzzFeed publicado días después de las elecciones, las veinte principales noticias falsas disponibles a través de Facebook durante los tres meses anteriores a las elecciones generaron más interacciones que las veinte principales historias reales publicadas en la misma red social por los medios de comunicación más reconocidos (New York Times, Washington Post, Los Angeles Times, Wall Street Journal, FOX News, entre otros).

The Guardian y BuzzFeed revelaron que gran parte de estas historias estaban siendo producidas por un grupo de jóvenes macedonios, quienes a través de titulares sugestivos que producían clics –técnica engañosa conocida como *clickbait*– hicie-

* Carlos Cortés. Investigador consultor de la Iniciativa por la Libertad de Expresión (ILEI) del CELE. Abogado de la Universidad de Los Andes (Colombia), con maestría en gobernanza de medios de London School of Economics. Es consultor en libertad de expresión y regulación de internet.

** Luisa Isaza. Abogada de la Universidad Javeriana (Colombia). Asesora legal de la Coordinación de Defensa y Atención a Periodistas de la Fundación para la Libertad de Prensa (FLIP), en Colombia.

¹ Siddiqui, Faiz y Svrluga, Susan, “N.C. man told police he went to D.C. pizzeria with gun to investigate conspiracy theory”, *Washington Post*, 5 de diciembre de 2016, disponible en: <http://wapo.st/2gW0yFo>.

² Goldman, Adam, “The Comet Ping Pong Gunman Answers Our Reporter’s Questions”, *The New York Times*, 7 de diciembre de 2016, disponible en: <http://nyti.ms/2Er1xrM>.

³ Otras noticias falsas relacionadas con las elecciones y que fueron ampliamente difundidas sostenían, por ejemplo, que se había confirmado que Hillary Clinton habría vendido armas a ISIS, que Donald Trump estaba ofreciendo tiquetes de un solo trayecto a África y México para aquellos que quisieran abandonar el país y que un líder de ISIS estaba pidiendo a los musulmanes que votaran por Clinton. Silverman, Craig, “Here Are 50 Of The Biggest Fake News Hits On Facebook From 2016”, *Buzzfeed*, 30 de diciembre de 2016, disponible en: <http://bzfd.it/2Ge4ZXo>.

⁴ Silverman, Craig, “This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook”, *Buzzfeed*, 16 de noviembre de 2016, disponible en: <http://bzfd.it/2ssm265>.

ron miles de dólares en publicidad gracias al tráfico de sus páginas de Internet.⁵ En la ciudad de Veles, Macedonia, se crearon más de cien sitios web con este objetivo, diseñados para parecer portales auténticos de noticias. Otras fábricas de noticias falsas operaban directamente desde Estados Unidos.⁶ De acuerdo con sus propios creadores, gran parte del tráfico de estos sitios provenía de los clics originados en Facebook, donde tenían además cientos de miles de seguidores.

La información falsa se difundió hasta el mismo día de las elecciones en forma de bulos, cuentas de Twitter falsas, trinos de desinformación e incluso resultados en Google.⁷ Después de las votaciones, Mediaite reportó que el primer resultado en el buscador de Google a “cuenta final de votos 2016” era un sitio llamado 70News donde falsamente se decía que Donald Trump había ganado tanto el voto electoral como popular.⁸

Tras la sorpresiva victoria de Donald Trump, la discusión sobre las noticias falsas estalló.⁹ Algunas voces –incluyendo la de un autor de *fake news*– dijeron que la desinformación en redes sociales había influido directamente en el resultado de las elecciones.¹⁰ Y aunque no hay por ahora un estudio que mida ese impacto de manera clara, resulta indudable que Facebook, Twitter y Google sí jugaron un rol importante al ser la fuente principal de información de muchas personas.¹¹

Los servicios de estas compañías no sólo fueron explotados por astutos jóvenes para hacer ganancias por publicidad. De acuerdo con los hallazgos de la inteligencia norteamericana, el gobierno ruso usó estas plataformas para esparcir noticias falsas y propaganda, buscando influenciar el debate público durante el tiempo de campaña y beneficiar la candidatura de Trump en detrimento de la de Clinton.¹² Más recientemente, la Primera Ministra del Reino Unido Theresa May hizo acusaciones similares en contra del gobierno ruso.¹³

⁵ Tynan, Dan, “How Facebook powers money machines for obscure political ‘news’ sites”, *The Guardian*, 24 de agosto de 2016, disponible en: <http://bit.ly/2bhzDzv>; Silverman, Craig y Alexander, Lawrence “How Teens In The Balkans Are Duping Trump Supporters With Fake News”, *Buzzfeed*, 3 de noviembre de 2016, disponible en: <http://bzfd.it/2EubVDu>; Ohlheiser, Abby, “This is how Facebook’s fake-news writers make money”, *The Washington Post*, 18 de noviembre de 2016, disponible en: <http://wapo.st/2Bt9wGk>.

⁶ Silverman, Craig y Singer-Vine, Jeremy, “The True Story Behind The Biggest Fake News Hit Of The Election”, *Buzzfeed*, 16 de diciembre de 2016, disponible en: <http://bzfd.it/2BZ7kHs>.

⁷ Rogers, Kathy y Bromwich, Jonah, “The Hoaxes, Fake News and Misinformation We Saw on Election Day”, *The New York Times*, 8 de noviembre de 2016, disponible en: <http://nyti.ms/2o4NhOA>.

⁸ Abrams, Dan, “Now Even Google Search Aiding in Scourge of Fake, Inaccurate News About Election 2016”, *Mediate*, 13 de noviembre de 2016, disponible en: <http://bit.ly/2ssmbq9>.

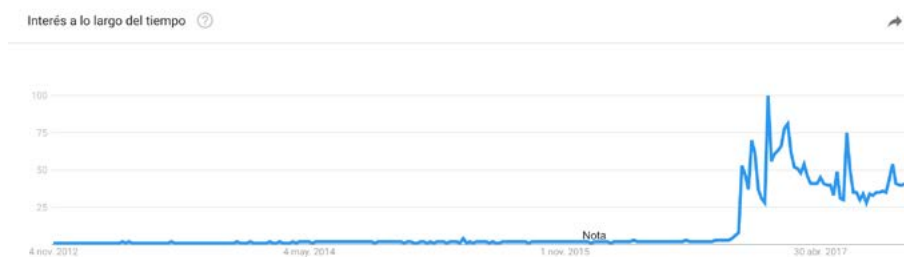
⁹ Una consulta en Google Trends revela que las búsquedas del término *fake news* se disparó en noviembre de 2016, luego de ser casi insignificantes durante los últimos años. El pico de las búsquedas se encuentra en la semana del 8 al 14 de enero de 2017, semana en la que se dio el conocido incidente en el cual Donald Trump se negó a contestar una pregunta de un periodista de CNN, diciéndole “*You are fake news!*”. Google Trends, <http://bit.ly/2Gfn3QZ>, último acceso: 4 de diciembre de 2017.

¹⁰ Read, Max “Donald Trump Won Because of Facebook”, *New York Magazine*, 9 de noviembre de 2016, disponible en: <http://nymag.com/selectall/2016/11/donald-trump-won-because-of-facebook.html>; Dewey, Caitlin, “Facebook fake-news writer: ‘I think Donald Trump is in the White House because of me’”, *The Washington Post*, 17 de noviembre de 2016, disponible en: <http://wapo.st/2f3NIIC>; Parkinson, Hannah, “Click and elect: how fake news helped Donald Trump win a real election”, *The Guardian*, 14 de noviembre de 2016, disponible en: <http://bit.ly/2fSyaDH>.

¹¹ De acuerdo con el último estudio sobre noticias digitales del Instituto Reuters para el Estudio del Periodismo de la Universidad de Oxford, una encuesta realizada en 36 países reveló que 54% de los usuarios de internet usan los medios sociales como fuente de noticias (Instituto Reuters para el Estudio del Periodismo, “Digital News Report”, <http://bit.ly/SBryvD>, último acceso: 14 de diciembre de 2017). En el caso de Estados Unidos, un estudio del Centro de Investigaciones Pew dice que el 67% de la población consume noticias a través de estas plataformas. Haciéndole honor a su nombre, el NewsFeed de Facebook es la principal fuente de noticias entre quienes respondieron que usaban los medios sociales como fuente de noticias (Centro de Investigaciones Pew, “News Use Across Social Media Platforms 2017”, <http://pewrsr.ch/2vMCQWO>, último acceso: 14 de diciembre de 2017).

¹² Oficina del Director Nacional de Inteligencia de la Comunidad de Inteligencia de los Estados Unidos, “Assessing Russian Activities and Intentions in Recent US Elections”, 6 de enero de 2017, disponible en: <http://bit.ly/2iRbS9b>.

¹³ Mason, Rowena, “Theresa May accuses Russia of interfering in elections and fake news”, *The Guardian*, 14 de noviembre de 2017, disponible en: <http://bit.ly/2mnQaML>.



Nota. Búsquedas en Google del término fake news durante los últimos cinco años (Google Trends).

Varias autoridades estadounidenses iniciaron investigaciones sobre la interferencia rusa en las elecciones, para lo cual han solicitado información a las empresas. Inicialmente, Facebook quiso restarle importancia al problema: dos días después de las elecciones, Mark Zuckerberg sostuvo públicamente que pensar que las noticias falsas tuvieron un impacto en la elección era una idea “bastante loca”.¹⁴ Pero posteriormente la empresa admitió ante el Senado de Estados Unidos que millones de usuarios habían visto publicidad producida por Rusia, publicada en Facebook e Instagram.¹⁵ Por su parte, Twitter declaró que encontró 2.752 cuentas controladas por rusos y que *bots* rusos tuitearon 1,4 millones de veces durante la elección. Google encontró en YouTube más de mil videos sobre el tema.

Este escenario se volvió la tormenta perfecta para las empresas de Internet, que por un menú amplio de temas y desde diferentes latitudes enfrentan presiones de gobiernos y sociedad civil para que transparenten sus prácticas.¹⁶ La desinformación no es un problema aislado ni tampoco nuevo. En el trasfondo está el dilema de los contenidos perjudiciales –difamatorios, incendiarios, violatorios de la intimidad, entre otros– y la respuesta de las empresas frente a estos. La moderación de contenidos está, pues, en el centro de esta discusión y sus implicaciones atraviesan el ejercicio de la libertad de expresión en línea.

Como se verá más adelante, Facebook y Google vienen adoptando medidas para lidiar con el problema. Unas se enfocan en la prevención, buscan educar al ciudadano para que tome decisiones informadas sobre el contenido que consume. Otras tienen efecto directo en la información que se publica en las plataformas, ya sea a través de cambios invisibles del algoritmo o de advertencias visibles sobre la veracidad de la información. La mayoría de estas medidas son ensayos o tienen una cobertura geográfica parcial;¹⁷ otras están pensadas únicamente para momentos específicos, como las elecciones. En cualquier caso, no es fácil determinar el alcance y profundidad de estas acciones ya que se trata, en esencia, de una serie de anuncios cuya implementación no es del todo clara.

Este documento expone las medidas que han anunciado Facebook y Google para combatir la desinformación. Incluye también una alusión breve a YouTube y Twitter. Nos enfocamos en aquellas medidas que tienen efecto directo en la plataforma y en la información que reciben sus usuarios, y no en medidas preventivas y educativas que se vienen trabajando paralelamente. Igualmente, el documento intenta determinar cuál será la cobertura geográfica de esas medidas. Posteriormente, a manera de conclusión expone cuáles pueden ser los problemas detrás de las soluciones propuestas en cuatro puntos: i) la escala y el tiempo, ii) el impacto, iii) el rol de la sociedad civil, y iv) la transparencia.

¹⁴ Entrevista realizada por el periodista David Kirkpatrick, fundador de Techonomy Media, el 10 de noviembre de 2016, disponible en: <http://bit.ly/2o6NMrv>.

¹⁵ Shaban, Hamza, Timberg, Craig y Dvoskin, Elizabeth “Facebook, Google and Twitter testified on Capitol Hill. Here’s what they said”, *The Washington Post*, octubre 31 de 2017, disponible en: <http://wapo.st/2gZM9fx>.

¹⁶ En junio de 2017, el parlamento alemán expidió una ley que obliga a las empresas de Internet que tengan más de dos millones de usuarios a eliminar de sus plataformas el discurso de odio y otros contenidos ilegales, como las noticias falsas difamatorias en máximo 24 horas, so pena de multas de hasta 50 millones de euros (Miller, Joe, “Germany votes for 50m euro social media fines”, BBC, 30 de junio de 2017, disponible en: <http://bbc.in/2C0rTna>). Más recientemente, un vocero de Theresa May informó que el Reino Unido está evaluando el rol que cumplen Facebook y Google en la provisión de noticias y cuáles podrían ser sus responsabilidades (“Britain looking at Google, Facebook role in news: PM May’s spokesman”, *Reuters*, 10 de octubre de 2017, disponible en: <http://reut.rs/2swt75t>).

¹⁷ El experimento de Facebook “Explore”, desarrollado en seis países, separa en una pestaña las publicaciones de amigos y la publicidad y en otra el contenido público de otras cuentas de Facebook. Facebook ha anunciado que no tiene planes de hacer esta medida definitiva ni extensiva a todo el mundo. Facebook, “Clarifying Recent Tests”, <http://bit.ly/2zwW0NF>, último acceso: 14 de diciembre de 2017.

II. Desinformación y manipulación: una clasificación tentativa

En noviembre de 2017, el Diccionario Collins eligió *fake news* como la palabra del año.¹⁸ La expresión, cuyo uso aumentó 365% este año, será incluida en la próxima edición de ese diccionario como “información falsa, frecuentemente sensacionalista, diseminada bajo el disfraz de reportaje de noticias”. *Fake news*, sin embargo, no tiene una única connotación. Las audiencias no solo la usan para referirse a reportajes falsos, sino en general para expresar una inconformidad con la desinformación, especialmente en línea.¹⁹ Y, para no ir demasiado lejos, el Presidente Donald Trump la usa para descalificar cualquier información con la que no está de acuerdo.

Lo cierto es que, más allá del uso político del término, *fake news* se relaciona también con opiniones extremistas, propaganda y manipulación. Una opinión desinformada y alarmista de un político a través de Twitter puede ser para muchos una noticia falsa, tanto como una noticia que, con deliberada mala fe, informe falsamente sobre la muerte de un líder mundial. En ambos casos encontramos diferencias de contenido (una opinión versus una noticia), de formato (un tuit versus una página web) y posiblemente de motivaciones (el político quiere arengar a su base mientras el sitio web quiere clics).

A continuación, se propone una clasificación orientada a entender cómo se produce el contenido y cómo alcanza al lector. Esta propuesta no desarrolla categorías excluyentes entre sí. Por ejemplo, un contenido puede ser una noticia falsa y tener, a la vez, un enfoque propagandístico. Tampoco intenta ordenar el tema en su totalidad, y excluye, en particular, los contenidos de periodismo satírico y los errores de reportaje cometidos de buena fe. Estos últimos podrían hacer parte de un debate sobre *fake news*, pero no se relacionan con el propósito de este documento.

1. Noticias falsas

Se trata de contenidos deliberadamente falsos que se publican en sitios web cuya apariencia intenta ser formal y auténtica.²⁰ A veces el diseño del sitio y su URL suplantan un portal de noticias reconocido. El propósito claro es engañar al usuario. Generalmente estos contenidos se mueven en redes sociales a través de las cuentas propias de esos portales, ya sea de manera orgánica –mediante *likes*, *retweets* y compartidos de los usuarios– o con acciones promocionadas, es decir, pagando para que estos contenidos sean publicitados por las plataformas.

¹⁸ Flood, Alison, “Fake news is ‘very real’ word of the year for 2017”, *The Guardian*, 2 de noviembre de 2017, disponible en: <http://bit.ly/2iTWYk4>.

¹⁹ “Our findings suggest that, from an audience perspective, fake news is only in part about fabricated news reports narrowly defined, and much more about a wider discontent with the information landscape— including news media and politicians as well as platform companies. Tackling false news narrowly speaking is important, but it will not address the broader issue that people feel much of the information they come across, especially online, consists of poor journalism, political propaganda, and misleading forms of advertising and sponsored content.” Nielsen, Rasmus y Graves, Lucas, “‘News you don’t believe’: Audience perspectives on fake news”, Instituto Reuters para el Estudio del Periodismo, octubre de 2017, disponible en: <http://bit.ly/2o4Exb6>.

²⁰ Media Matters, “Understanding The Fake News Universe”, <http://bit.ly/2EDq8NH>.



Nota. Historia falsa publicada en el sitio abcnews.com.co, creado para imitar a ABC News, cuya dirección web es abcnews.go.com. Según esta falsa noticia, a algunos de los manifestantes contra Donald Trump les pagaron para protestar. A pesar de la falsedad de la noticia, días después de su elección el mismo Trump sugirió que estas personas eran “manifestantes profesionales”.²¹



Nota. Una foto de una votación en Arizona fue alterada para incluir la imagen de un arresto. Quien la publicó en Twitter sostuvo que un inmigrante indocumentado había sido arrestado por tratar de votar.²²

²¹ Esta noticia fue compartida en Twitter por Eric Trump, hijo de Donald Trump, por Corey Lewandowski y Kellyanne Conway, dos jefes de campaña de Trump. Jacobson, Louis, “No, someone wasn’t paid \$3,500 to protest Donald Trump; it’s fake news”, *Politifact*, 17 de noviembre de 2016, disponible en: <http://bit.ly/2flpuUR>; Stahl, Lesley, “President-elect Trump speaks to a divided country”, *CBS News*, 13 de noviembre de 2016, disponible en: <http://cbsn.ws/2swFkKz>.

²² Wardle, Claire, “6 types of misinformation circulated this election season”, *Columbia Journalism Review*, 18 de noviembre de 2016, disponible en: <http://bit.ly/2CkRdjh>.

Las noticias falsas en sentido estricto pueden tener motivaciones económicas, políticas o un poco de ambas. En el primer caso, se trata de operaciones comerciales que buscan generar tráfico a partir de contenidos falsos y, sobre todo, titulares sensacionalistas a los que la gente les da clic, pero cuya información relacionada no tiene sentido o relevancia alguna. En el segundo caso, las noticias intentan parecer auténticas no tanto para generar tráfico y ganancias sino para manipular el debate público a favor de determinados intereses políticos. Un ejemplo de esta categoría es la información falsa ya mencionada sobre la adhesión del Papa Francisco a la candidatura de Donald Trump en 2016.

Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement

TOPICS: Pope Francis Endorses Donald Trump



Nota. La noticia falsa sobre el apoyo del Papa a Donald Trump.

El interés político en las *fake news* no es necesariamente excluyente del económico. Si bien algunas noticias falsas se crean con una u otra motivación, en muchos casos ambas órbitas pueden confluir. En el caso de la desinformación alrededor de la campaña presidencial en Estados Unidos, a los jóvenes macedonios podía resultarles indiferente quién ganara la elección; pero no a los operativos rusos que también influyeron en ella. En este último caso, los beneficios económicos por cuenta del tráfico y la interacción se sumaban a la agenda política subyacente.²³

2. Propaganda

Para Jacques Ellul, la propaganda es un concepto elusivo que se desarrolla alrededor de la acción psicológica y la acción bélica, la reeducación, el lavado de cerebro y las relaciones públicas entre humanos. En ese sentido, para el sociólogo francés la propaganda es, sobre todo, una técnica para influir en las acciones de grupos o individuos.²⁴

Entre otras manifestaciones, la propaganda puede abarcar información falsa o información cierta presentada con un enfoque engañoso.²⁵ Por ejemplo, se presentan unos hechos pero se omiten otros; se saca la información de contexto; se manipula el contenido; se presentan teorías u opiniones como hechos; se le da credibilidad a información altamente discutible; se niega información cierta con el objetivo de crear confusión, o se proclama una verdad única en oposición al 'otro' –la estrategia de los movimientos nacionalistas–.²⁶

La propaganda hace parte de la política y las comunicaciones al menos desde comienzos del siglo pasado. No es, por lo tanto, un fenómeno digital. Sin embargo, el alcance de estos contenidos en el entorno en línea sí reviste especial

²³ "These Macedonians on Facebook didn't care if Trump won or lost the White House. They only wanted pocket money to pay for things--a car, watches, better cell phones, more drinks at the bar". Subramian, Samantha, "Welcome To Veles, Macedonia, Fake: News Factory To The World", Wired, marzo de 2017, disponible en: <http://bit.ly/2o7BOxQ>.

²⁴ Ellul, Jacques, *Propaganda: The Formation of Men's Attitudes*, New York, Vintage Books, 1965.

²⁵ "Propaganda is false or misleading information or ideas addressed to a mass audience by parties who thereby gain advantage. Propaganda is created and disseminated systematically and does not invite critical analysis or response". Huckin, Thomas, "Propaganda Defined", en: Henderson, Gae Lyn y Braun, M.J (ed.), *Propaganda and Rhetoric in Democracy: History, Theory, Analysis*, Carbondale, Southern Illinois University Press, 1a ed., 2016, pp. 118-136.

²⁶ Powell, John A. "Us vs them: the sinister techniques of 'Othering' – and how to avoid them", *The Guardian*, 8 de noviembre de 2017, disponible en: <http://bit.ly/2iGLAUX>.

importancia: a través de la herramienta de anuncios –especialmente en Facebook– la propaganda se ajusta a comunidades y grupos específicos a partir de gustos, inclinaciones políticas y círculos de amistad. En el caso de las elecciones norteamericanas, este nivel de granularidad alcanzó a un nivel alarmante:

Los *dark ads* o anuncios oscuros –como se les conoce debido a la imposibilidad de saber quiénes los ven– permitieron reafirmar convicciones políticas, atizar diferencias y, en general, polarizar al electorado. Un hombre blanco de un estado republicano veía un anuncio contra la inmigración o en defensa del uso de armas; un afroamericano veía un anuncio que recordaba la persecución racial del Ku Klux Klan; un católico veía a Hillary encarnando al diablo en una pelea contra Jesús.²⁷



Nota. Un anuncio televisivo de la campaña de Donald Trump decía que el entonces candidato detendría la inmigración ilegal en la “frontera sur” construyendo un muro que sería pagado por México, mientras mostraba un video de decenas de personas cruzando la frontera entre Marruecos y la ciudad española de Melilla, no entre México y Estados Unidos.²⁸

3. Teorías conspirativas

Las teorías conspirativas buscan explicar un suceso particular como el resultado de un plan cuidadosamente coordinado por un individuo o un grupo. Las motivaciones son generalmente secretas y maliciosas, y las acciones se ejecutan en detrimento del interés general.²⁹ Estas teorías pululan en canales de videos y páginas de Internet, y se presentan muchas veces como noticias a pesar de su escaso fundamento fáctico.

En Colombia, desde hace algunos años, circula una teoría según la cual el presidente Juan Manuel Santos fue reclutado para trabajar secretamente para el gobierno cubano.³⁰ En Argentina, luego de la desaparición del joven Santiago Maldonado, se desataron en Internet múltiples teorías conspirativas, incluyendo una que señalaba que la página web santiagomaldonado.com, creada para exigir la aparición con vida del activista, había sido puesta en funcionamiento antes de su desaparición.³¹

La teoría conspirativa del *Pizzagate*, mencionado en la introducción de este texto, fue examinada por varios medios de comunicación en Estados Unidos con el propósito de identificar su origen.³² En octubre de 2016, un usuario de

²⁷ Cortés, Carlos, “El algoritmo imposible, redes sociales y noticias falsas”, *Revista Arcadia*, diciembre de 2017, disponible en: <http://bit.ly/2CmcdpT>.

²⁸ Edds, Carolyn, “Donald Trump’s first TV ad shows migrants ‘at the southern border,’ but they’re actually in Morocco”, *Politifact*, 4 de junio de 2016, disponible en: <http://bit.ly/1mvNQI9>.

²⁹ Media Matters, “Understanding The Fake News Universe”, <http://bit.ly/2EsVDdZ>, último acceso: 14 de diciembre de 2017. Según Michael Barkun, las teorías de la conspiración se caracterizan por responder a tres principios, a saber: i) nada sucede por accidente, ii) nada es lo que parece y iii) todo está conectado. Barkun, Michael, *A Culture of Conspiracy: Apocalyptic Visions in Contemporary America*, Berkeley, University of California Press, 2003, pp. 3-4, disponible en: <http://bit.ly/2Btakek>.

³⁰ “Los Santos y su militancia castrocomunista”, *Periodismo Sin Fronteras*, 1 de julio de 2013, disponible en: <http://bit.ly/1hjVlzx>.

³¹ “El sitio de Santiago Maldonado fue creado antes de su desaparición”, *Data 24*, 18 de septiembre de 2017, disponible en: <http://bit.ly/2w3lOed>.

³² Al respecto, pueden verse los siguientes artículos: (i) Aisch, Gregor, Huang, Jon y Kang, Cecilia, “Dissecting the #PizzaGate Conspiracy Theories”, *The New York Times*, 10 de diciembre de 2016, disponible en: <http://nyti.ms/2jcCzlu>. (ii) Silverman, Craig, “How The Bizarre Conspiracy Theory Behind ‘Pizzagate’ Was Spread”, *Buzzfeed*, 5 de diciembre de 2016, disponible en: <http://bzfd.it/2CjSQOI>.

Twitter replicó desde su cuenta un mensaje de Facebook en el que una mujer aseguraba que una fuente anónima del Departamento de Policía de la Ciudad de Nueva York le había dicho que había evidencia de que Hillary y Bill Clinton estaban involucrados en una red de tráfico sexual infantil. El tuit alcanzó miles de retuits, y horas después un usuario de un foro de discusión en Internet publicó un mensaje en el cual decía que “fuentes internas” habían confirmado la existencia de la red de pedofilia, la cual sería expuesta en cuestión de horas.³³ Al día siguiente, el sitio de noticias falsas YourNewsWire.com publicó una historia basada en los comentarios de un usuario de 4chan –un tablero en línea donde abunda material de *troleo*– donde se decía que una fuente del FBI había confirmado las acusaciones.³⁴ La historia fue replicada y ampliada por otros sitios de noticias falsas y compartida en Facebook.

Entre todas las versiones publicadas, la historia alcanzó cientos de miles de interacciones en Facebook y Twitter, y empezó a volverse viral bajo la etiqueta #PizzaGate.³⁵ Se crearon noticias falsas con fotos manipuladas. En foros de discusión y comentarios en redes se hablaba de túneles subterráneos, cámaras de tortura, satanismos y canibalismo en los sótanos de varios restaurantes. Cuando los medios de comunicación refutaron la teoría fueron acusados por los creyentes de querer ocultar la verdad. Aún semanas después de que fuera arrestado el hombre que disparó contra la pizzería Comet Ping Pong –donde, por supuesto, no operaba ninguna red ilegal– algunas personas seguían insinuando que la historia era cierta.³⁶

4. Información falsa, rumores, cadenas, memes

Una categoría más amplia de desinformación, que recoge distintas formas de noticias falsas, propaganda y teorías conspirativas, son los contenidos mezclados y difundidos a través de servicios de mensajería instantánea, principalmente WhatsApp. En ese entorno la información se rota de mano en mano a manera de imagen, video o memes, sin autor identificado ni identificable. La legitimidad y autoridad del contenido la da, en últimas, aquel conocido o conocida que lo comparte.

En países como Colombia esta clase de contenidos tuvieron un alcance aparentemente alto el año pasado. Durante las campañas a favor y en contra de los acuerdos de paz de La Habana previas al plebiscito de octubre de 2016, se difundieron a través de WhatsApp contenidos falsos, imprecisos y descontextualizados sobre los acuerdos, orientados a capturar el voto en contra de los acuerdos. Como lo reconoció uno de los líderes del movimiento del no, el objetivo era generar rabia e indignación.³⁷

(iii) LaCapria, Kim, “Chuck E. Sleaze”, *Snopes*, 21 de noviembre de 2016, disponible en: <http://bit.ly/2xX7xta>.

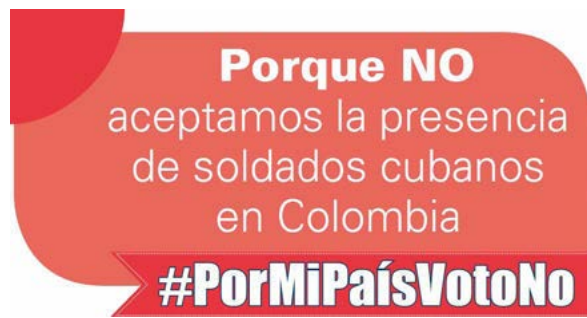
³³ El foro de discusión “Breaking: Its worse then classified emails. Political Pedophile Sex Ring exposed” está disponible en: <http://bit.ly/2ErmerY>, último acceso: 14 de diciembre de 2017.

³⁴ “FBI Insider: Clinton Emails Linked To Political Pedophile Sex Ring”, *YourNewsWire.com*, 31 de octubre de 2016, disponible en: <http://bit.ly/2fesMdl>.

³⁵ Silverman, Craig, “How The Bizarre Conspiracy Theory Behind ‘Pizzagate’ Was Spread”, *Buzzfeed*, 5 de diciembre de 2016, disponible en: <http://bzfd.it/2CjSQOI>.

³⁶ Allam, Hannah, “Conspiracy peddlers continue pushing debunked ‘pizzagate’ tale”, *Miami Herald*, 5 de diciembre de 2016, disponible en: <http://hrl.d.us/2CFYGKw>.

³⁷ Ramírez, Juliana, “El No ha sido la campaña más barata y más efectiva de la historia”, *Asuntos Legales*, 4 de octubre de 2016, disponible en: <http://bit.ly/2EHxlwc>.



Porque:

- Las Farc NO Entregarán Las Armas.
- Las Farc NO Devuelven Niños Secuestrados.
- Las Farc NO Se Arrepienten De Sus Crímenes.
- Las Farc NO Pagarán Un Día de Cárcel.
- Las Farc NO Abandonan El Narcotráfico.
- Las Farc NO Entregan Las Rutas de la Coca Ni \$.
- Las Farc NO Han Entregado A Secuestrados.
- Las Farc NO Han Dejado de Reclutar Menores.
- Las Farc NO Van A Indemnizar Víctimas.
- Las Farc NO Dejan de Hacer Terrorismo.
- Las Farc NO Se Someterán Al Voto Popular.
- NO Quiero Mantener a La Farc con + Impuestos.
- NO Quiero + Asesinatos De Militares y Civiles.
- NO Quiero a Colombia como Venezuela.
- La Decisión es sobre ACUERDOS no la Paz

#DespiertaColombia  #DespiertaMedellin

Nota. Mensajes con información falsa en redes sociales y Whatsapp durante la campaña previa al plebiscito de octubre de 2016.³⁸

Ya sea como noticias eminentemente falsas, propaganda, teorías conspirativas o rumores, las *fake news* no son un fenómeno aislado de la realidad social y política, ni mucho menos una externalidad de la tecnología. “Hoy, cuando hablamos de la relación de la gente con el Internet, tendemos a adoptar el lenguaje acríptico de la ciencia computacional. Las ‘fake news’ se describían como un ‘virus’ entre usuarios ‘expuestos’ a desinformación en línea”, explica la socióloga norteamericana Katherine Cross.³⁹ La desinformación se origina y retroalimenta en la incorporación humana, y es en esa relación donde podemos ubicar la dimensión del problema y las limitaciones de las soluciones propuestas.

III. Affordances y espacios

Así como el fenómeno de la desinformación responde a distintos contextos sociales, no se manifiesta de la misma forma en todos los entornos digitales: un espacio como WhatsApp posibilita el intercambio de contenidos acéfalos –sin autor o fuente aparente–; Facebook le da mayor relevancia a contenidos que se comparten masivamente, y en Twitter el usuario escoge las voces que quiere oír. Es en ese diálogo entre la plataforma y el usuario donde la producción y el intercambio de información adquiere connotaciones particulares. Las *fake news* no se manifiestan de la misma forma en todos estos espacios.

Los *affordances* u ofrecimientos son las propiedades que surgen a partir de la relación entre el objeto y la persona. El concepto de *affordances*, introducido por el psicólogo James Gibson, se refiere a las posibilidades de acción en determinado contexto, y se utiliza para hablar sobre cómo interactúan usuarios con objetos, ambientes o tecnologías para lograr determinados resultados.⁴⁰ Identificar los ofrecimientos de un objeto permite entender las distintas formas en que éste puede ser utilizado para diversos fines. Algunos de esos ofrecimientos están inscritos en el diseño del objeto; otros son “descubiertos” por el individuo. Una escoba sirve para barrer, pero también para golpear un objeto lejano, como una naranja en un árbol; una mesa sirve para poner objetos encima o sentarse sobre ella; un servicio de

³⁸ “¿Qué tan ciertos son algunos memes de la campaña del “NO” en el plebiscito?”, *Pacifista*, 30 de agosto de 2016, disponible en: <http://bit.ly/2bP16MN>.

³⁹ Cros, Katherine, “The Art of the Real: Disinformation vs. democracy”, *The Baffler*, junio de 2017, disponible en: <http://bit.ly/2Er1KLR>.

⁴⁰ Gibson, James, *The Ecological Approach to Visual Perception*, Hillsdale, Cornell University, 1986. Evans, Pearce y otros definen los ofrecimientos como la estructura relacional multifacética entre un objeto o tecnología y el usuario, que permite o restringe potenciales resultados conductuales en un contexto en particular. Para entender el concepto proponen el siguiente ejemplo: la cámara integrada de un smartphone es una función del teléfono, un ofrecimiento es la capacidad del teléfono de grabar (por ejemplo, imágenes o videos de una persona) y un posible resultado es la documentación de una violación de derechos humanos con esa cámara. Explican que las funciones o herramientas son estáticas, mientras que los ofrecimientos son dinámicos. Evans, Sandra K. y otros, “Explicating Affordances: A Conceptual Framework for Understanding Affordances in Communication Research”, en: *Journal of Computer-Mediated Communication*, Vol. 22, No. 1, Blackwell Publishing Ltd, 2017, disponible en: <http://bit.ly/2EEFYfM>.

mensajería sirve para difundir mensajes propios o de extraños. El intelectual francés Bruno Latour define el diseño en sí mismo como un proceso de inscripción de modos de uso: “Los productos resultante llevan consigo los ‘guiones’ que inhiben o precluyen ciertas acciones mientras invitan o exigen otras”.⁴¹

Cuando esos objetos, ambientes o tecnologías permiten acciones sociales, nos encontramos frente a un ofrecimiento social.⁴² En el contexto de las tecnologías sociales, distintos trabajos se han centrado en estudiar, por ejemplo, cómo las redes sociales son usadas por las personas para organizar su vida privada; por los gobiernos para tener un contacto directo con sus ciudadanos o vigilarlos; por empresas para estimular el trabajo en equipo; por instituciones educativas para promover fines pedagógicos, o por organizaciones políticas para motivar la participación ciudadana.⁴³ En los *affordances* del servicio influyen también las políticas de uso y los algoritmos, que son los que permiten esas interacciones y el acceso a contenidos o usuarios (a través de los recomendados del *newsfeed*, por ejemplo). Es a partir de esa incorporación del producto por parte del usuario y de las comunidades donde puede identificarse cómo se diseña, difunde y consume la desinformación. En muchos casos, se trata simplemente del uso ordinario del servicio: una noticia falsa es una unidad informativa como cualquier otra. En otros casos, es una consecuencia no deseada pero que el producto ofrece: usar un titular engañoso para generar clics y viralizar una mentira. De una u otra forma, se trata de usos que no son ajenos al producto.

Tomemos el ya mencionado ejemplo del *Pizzagate*. La génesis de la historia estuvo en las redes sociales: un rumor compartido en Facebook (“fuente anónima de la Policía de Nueva York dijo que...”) y replicado en Twitter. Este contenido falso surge orgánicamente, o sea, a través de los mecanismos normales que tiene el servicio para compartir información. La configuración de ambas plataformas, que premia la interacción de distintas formas, permite que este contenido se consuma masivamente. La veracidad de éste es irrelevante. De forma paralela, un sitio web que produce noticias falsas incorpora ese rumor a manera de reportaje y, además de publicarlo de manera orgánica, lo promociona. Es decir, le paga a una plataforma como Facebook para que exhiba esa información a manera de anuncio y obtenga mayor interacción y visibilidad. En este punto, el producto ofrece una herramienta para que esa publicidad llegue a un público definido en detalle. Es lógico suponer que el uso inscrito –deseado– del servicio sea promocionar productos veraces, pero su diseño ofrece –*affords*– la posibilidad de promocionar información fraudulenta.

Pasemos a Twitter. El diseño de la red permite y promueve la creación de conversaciones abiertas y descentralizadas. Este ofrecimiento ha permitido gestionar protestas sociales, movimientos de derechos e incontables acciones conjuntas. Pero la incorporación del producto permite que esas acciones conjuntas sean concertadas, con lo cual se abre la posibilidad de que un grupo de cuentas aparentemente auténticas se dediquen a promover contenido falso.⁴⁴ Y en ese entorno abierto y de participación, muchos usuarios considerarán que ese volumen de conversación es auténtico y que, por lo tanto, la información que transmite es veraz.

Terminemos con el caso de Whatsapp. Al ser un servicio de mensajería, su diseño está pensado para preservar la privacidad de las comunicaciones. Esta configuración cerrada imposibilita, en principio, el monitoreo de los contenidos,

⁴¹ Dow Schüll, Natasha, *Addiction by Design: Machine Gambling in Las Vegas*, Princeton, Princeton University Press, 2012 (Ed. Kindle). Traducción informal.

⁴² “Along the same lines as Gibson’s, our hypothesis is that the richest and most elaborate environmental affordances are provided by other animals and other people”. Kaufmann, Laurence y Clément, Fabrice, “How Culture Comes to Mind: From Social Affordances to Cultural Analogies”, en: *Intellectiva*, No. 46-47, 2007, disponible en: <http://bit.ly/2sv5lqD>.

⁴³ Ver, Kursat Ozenc, Fatih y Farnham, Shelly “Life “modes” in social media” en: *CHI ‘11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Vancouver, 7 al 12 de mayo de 2012; Mergel, Inés, “Implementing Social Media in the Public Sector”, octubre de 2013, disponible en: <http://bit.ly/2o6K0hT>; Zeiller, Michael y Schauer, Bettina “Adoption, motivation and success factors of social media for team collaboration in SMEs”, en: *i-KNOW ‘11 Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, Graz, 7 al 9 de septiembre de 2011; Wang, Qiyun, Lit Woo, Huay y Lang Quek, Choon, “Exploring the Affordances of Facebook for Teaching and Learning”, en: *International Review of Contemporary Learning Research*, No. 23-31, 2012, disponible en: <http://bit.ly/2Gd7ol3>; Blegind Jensen, Tina y Dyrby, Signe “Exploring Affordances Of Facebook As A Social Media Platform In Political Campaigning”, en: *Proceedings of the 21st European Conference on Information Systems*, ECIS 2013 Completed Research, 2013, disponible en: <http://bit.ly/2EIGgxm>.

⁴⁴ Un estudio de la Universidad de Edimburgo identificó cientos de cuentas falsas operadas desde Rusia con el objetivo de influenciar el referendo sobre la permanencia del Reino Unido en la Unión Europea (*Brexit*). Booth, Robert y otros, “Russia used hundreds of fake accounts to tweet about Brexit, data shows”, *The Guardian*, 14 de noviembre de 2017, disponible en: <http://bit.ly/2iUPF8b>.

ya sea por la propia plataforma o por terceros (de ahí que se le conozca como un entorno social oscuro o *dark social*).⁴⁵ A esto hay que sumar que los únicos que pueden compartir esos contenidos son los contactos que hacen parte de una conversación –individual o colectiva–.⁴⁶ Al ser los individuos quienes comparten videos, imágenes, audios o textos, la información no incluye necesariamente una fuente, pero aun así se comparte en un contexto de intimidad y con un interlocutor conocido, todo lo cual le da una legitimidad. El objeto tiene el uso inscrito de conectar personas conocidas en un entorno cerrado, lo cual no excluye que ese canal se use para compartir cualquier tipo de información.

Tener claras estas diferencias es importante por dos razones. Por una parte, permite entender las soluciones propuestas e identificar sus limitaciones inherentes. Por la otra, sitúa el problema de las *fake news* en la órbita de la incorporación social de una tecnología. El uso de la tecnología –y por ende, los cuestionamientos que surjan por el uso de ella– es un proceso mediado por las personas y no una ecuación aislada: “Las tecnologías no se insertan en la vida cotidiana, provocando una revolución o una ruptura radical como suele decirse; por el contrario, esta inserción suele conllevar una evolución gradual, un negociación entre las prácticas heredadas y el deseo de transformación de las sociedades”.⁴⁷

IV. La solución del problema

Mucho antes de que surgiera un debate alrededor de la desinformación, las redes sociales ya enfrentaban el reto general de moderar contenidos en línea. La acción de arbitrar el flujo de información es quizá la muestra más grande del poder que tienen estas plataformas en su condición de intermediarios:⁴⁸ “No son los que producen el contenido, pero toman importantes decisiones sobre ese contenido: qué van a distribuir y a quién, cómo van a conectar a los usuarios y agenciar sus interacciones, y qué van a rechazar”.⁴⁹ Mediante la moderación de contenidos, estos actores intentan aplicar sus normas comunitarias para fidelizar a los usuarios con el servicio y mantener una operación libre de interferencias indeseadas.

Los discursos de odio y el terrorismo, la discriminación contra minorías, el acoso contra mujeres y el contenido tóxico, en general, obligaron a estas empresas a buscar un complicado balance entre la libre circulación de contenidos y la restricción oportuna. En ese proceso, y de manera creciente, enfrentan cuestionamientos frente a la transparencia y rendición de cuentas de estos procesos –*accountability*–. El rol de los algoritmos en las decisiones sobre contenidos, la suspensión de cuentas y los mecanismo de apelación, entre otros, hacen parte de una agenda de exigencias que tanto gobiernos como sociedad civil les piden que atiendan a estas empresas.⁵⁰

Es en ese contexto donde empresas como Facebook, Google y Twitter intentan responder a la problemática de las *fake news*. Y aunque sería un tema para desarrollar en otra oportunidad, resulta relevante ubicar esas respuestas dentro de las políticas de moderación de contenidos que ya tienen esas plataformas –y no como un asunto aparte–. Por ejemplo: mientras Facebook tiene una política de nombres reales, Twitter no prohíbe los seudónimos o las cuentas de parodia. Ese punto de partida demarca desarrollos distintos frente a la moderación de la desinformación.

⁴⁵ No se descartan los esquemas de vigilancia de las comunicaciones, pero el punto central es que el producto no está diseñado para generar y medir interacción y consumo de contenidos como Facebook o Twitter.

⁴⁶ Que una persona desconocida obtenga el teléfono de uno y comparta algo por Whatsapp tendrá posiblemente un efecto alienante. El receptor rechazará el mensaje o no le dará relevancia a la información.

⁴⁷ Gómez, Rocío y otros (comp.), *Facebook como obra mundana. Poetizar la vida y recrear vínculos personales*, Universidad del Valle, Programa Editorial, 2016, pp. 66.

⁴⁸ Véase, Cortés, Carlos “Las llaves del ama de llaves: la estrategia de los intermediarios en Internet y el impacto en el entorno digital”, en: Bertoni, Eduardo (comp.), *Internet y Derechos Humanos. Aportes para la discusión en América Latina*. Buenos Aires, CELE, Universidad de Palermo, 2014.

⁴⁹ Gillespie, T. Citado en: Myers West, S. ‘Raging Against the Machine: Network Gatekeeping and Collective Action on Social Media Platforms’, *Media and Communication* (ISSN: 2183–2439) 2017, Volume 5, Issue 3, p. 28 (traducción informal).

⁵⁰ Ver, entre otras iniciativas, www.onlinecensorship.org y <http://responsible-tech.org/>. El Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión de Naciones Unidas, por ejemplo, hará un reporte en junio de 2018 sobre la regulación de contenido en la era digital.

1. Facebook

A mediados de 2017 el diario inglés *The Guardian*, a través del proyecto Los archivos Facebook (*The Facebook Files*), ofreció un panorama de las prácticas de moderación de contenidos de la empresa.⁵¹ Las *fake news* en particular hacen parte de lo que denominan ‘operaciones informativas’, a las que Facebook entiende como “las acciones desarrolladas por actores organizados para distorsionar el sentimiento político nacional o extranjero”.⁵² Las operaciones informativas se dividen en noticias falsas, desinformación y cuentas falsas. Nótese como la empresa reconoce el carácter conjunto y organizado de estas acciones, entendiendo que se trata de un uso adaptado y particular del producto.

Bajo ese entendido, Facebook ha anunciado soluciones para enfrentar las noticias falsas enfocadas en dos áreas. Por una parte, ha anunciado medidas para promover la alfabetización en noticias (*news literacy*) de sus usuarios, para ayudarlos a tomar decisiones informadas sobre las noticias y fuentes en las que se puede confiar. Para ello, ha desarrollado dos proyectos: el *Facebook Journalism Project* y *New Integrity Initiative*.⁵³ Estas medidas preventivas y educativas no se abordan en este documento. Por otra parte, anunció la adopción de medidas técnicas que impactan en los contenidos que los usuarios reciben en la plataforma. Según lo publicado por la empresa, éstas son las características de dichas medidas:

1.1. Reporte, verificación de hechos y advertencias

En diciembre de 2016 Facebook anunció que estaba ensayando distintos mecanismos para facilitar el reporte de posibles noticias falsas por parte de sus usuarios.⁵⁴ Los reportes de la comunidad y “otras señales” –Facebook no aclara cuáles– serían considerados para seleccionar historias que enviarían a organizaciones independientes para hacer verificación de hechos (*fact-checking*). Facebook no aclara qué cantidad de reportes son suficientes para generar una alerta.

Si las organizaciones disputan el contenido luego del proceso de verificación, Facebook mostrará advertencias (*flagging*) señalando que el artículo ha sido cuestionado. Al lado de la bandera se incluirá un enlace para consultar información más detallada. Si a pesar de estos mensajes un usuario quiere compartir el contenido, Facebook podrá mostrar un nuevo mensaje en el que advierte que se va a compartir contenido que ha sido cuestionado. Además, según Facebook, estas historias podrán tener menos prevalencia en el *newsfeed*, y no podrán ser promovidas o convertidas en anuncios.

Mark Zuckerberg ha dicho que Facebook no quiere ser árbitro de la verdad.⁵⁵ En esa medida, la tarea de verificación de los hechos se la encargó a organizaciones externas firmantes del Código de Principios Internacionales de Verificación de Hechos de Poynter.⁵⁶ Inicialmente, la empresa está trabajando con ABC News, FactCheck.org, Associated Press, Snopes y Politifact.⁵⁷ Entre finales de 2016 y 2017, estas medidas se han anunciado como ensayos en países como Estados Unidos, Alemania, Francia y Países Bajos, y no están disponibles de manera permanente.⁵⁸

⁵¹ Véase, “The Facebook Files”, *The Guardian*, disponible en: <http://bit.ly/2r6h5yb>.

⁵² Facebook, “Information Operations and Facebook”, 27 de abril de 2017, disponible en: <http://bit.ly/2oOOS9s>.

⁵³ A través de esta iniciativa, Facebook ofrece herramientas de formación para periodistas. “Facebook Journalism Project”, <http://bit.ly/2ioDPAO>, último acceso: 14 de diciembre de 2017; La *News Integrity Initiative* es la gran apuesta de Facebook en el tema de alfabetización en noticias. A través de este proyecto se creó un consorcio de líderes de la industria tecnológica, instituciones académicas y otras organizaciones, que tendrá como objetivo ayudar a la gente a tomar decisiones informadas sobre las noticias que leen y comparten en Internet. “News Integrity Initiative”, <http://bit.ly/2D5udGi>, último acceso: 14 de diciembre de 2017.

⁵⁴ “Enfrentando engaños y noticias falsas”, <http://bit.ly/2HgPW0k>, último acceso: 14 de diciembre de 2017.

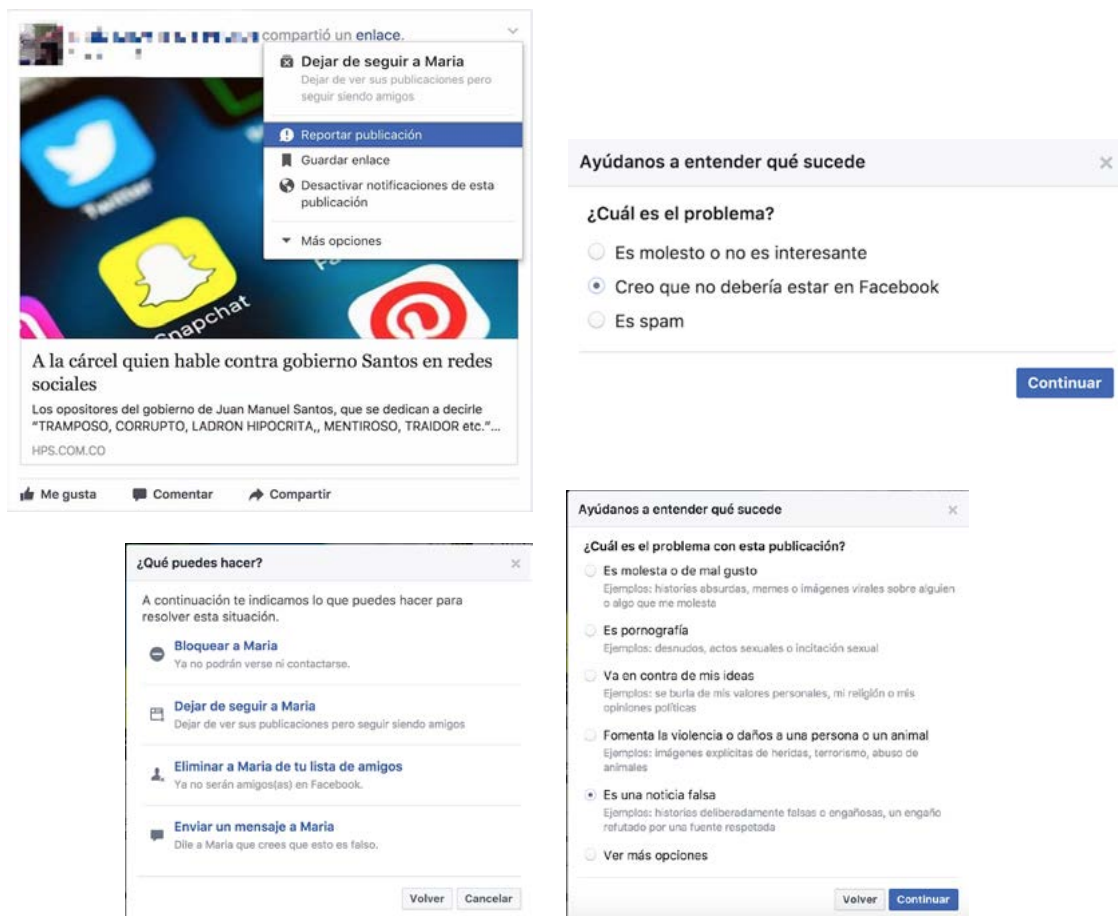
⁵⁵ <http://bit.ly/2fOsvha>

⁵⁶ International Fact-Checking Network, “International Fact-Checking Network fact-checkers’ code of principles”, disponible en: <http://bit.ly/2BwtakU>.

⁵⁷ Mullin, Benjamin, “Facebook rolls out plan to fight fake news”, *Poynter*, 15 de diciembre de 2016, disponible en: <http://bit.ly/2h5bPm5>.

⁵⁸ Ver: Facebook, “News Feed FYI: Addressing Hoaxes and Fake News”, 15 de diciembre de 2016, disponible en: [http://bit.ly/2gFFVw](http://bit.ly/2gFFVw;); Facebook, “Umgang mit Falschmeldungen”, 15 de enero de 2017, disponible en: [http://bit.ly/2gFvye](http://bit.ly/2gFvye;); Facebook, “Facebook annonce le lancement de son outil de fact-checking en France”, 6 de febrero de 2017, disponible en: <http://bit.ly/2swl4Ws>; Facebook, “Addressing Fake News Globally”, 15 de enero de 2017, disponible en: <http://bit.ly/2jyZ1FV>.

Al hacer su anuncio, Facebook publicó algunas imágenes del esquema de reporte. Una prueba hecha desde una conexión con IP en Estados Unidos confirmó ese mecanismo. Sin embargo, el mismo ejercicio realizado en agosto de 2017 desde una IP en Colombia mostró que si bien los usuarios podían señalar la opción “Es una noticia falsa”, no era posible enviar un reporte a Facebook. Las únicas opciones eran bloquear, dejar de seguir, eliminar o enviar un mensaje a quien había compartido la información cuestionada. Un nuevo ejercicio realizado en diciembre de 2017 muestra que la opción de reporte está deshabilitada en ambos países, al menos con el usuario de Facebook que usamos.



Nota: capturas de pantalla del esquema de reporte de noticias falsas.

1.2. Cambios en el newsfeed y medidas para contrarrestar la falsa amplificación

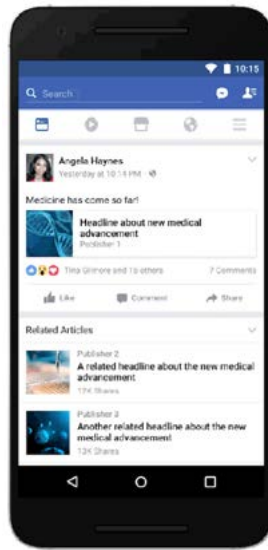
Como se señaló antes, las noticias que los verificadores marquen como falsas podrán tener menos relevancia en el *newsfeed*. Adicionalmente, hay otras señales que pueden llevar a que Facebook les dé menor relevancia a estos contenidos: por ejemplo, si al leer un artículo la gente está menos inclinada a compartirlo, puede ser una señal de que la historia es engañosa.⁵⁹ O cuando los usuarios regularmente hacen muchas publicaciones por día con contenidos externos que pueda considerarse de baja calidad —engañoso o sensacionalista— la publicación con un enlace externo puede ser “castigada”.⁶⁰ Por otra parte, se han anunciado medidas adicionales en el *newsfeed*:

– De acuerdo con un anuncio de abril de 2017, Facebook está ensayando que los artículos relacionados con una pu-

⁵⁹ Facebook, “Enfrentando engaños y noticias falsas”, 15 de diciembre de 2016, disponible en: <http://bit.ly/2HgPW0k>.

⁶⁰ Facebook, “News Feed FYI: Showing More Informative Links in News Feed”, 30 de junio de 2017, disponible en: <http://bit.ly/2uzG8HM>.

blicación aparezcan antes de que el usuario lea el contenido deseado, de forma tal que tenga acceso rápido a otras perspectivas sobre el mismo tema, incluyendo artículos que hayan pasado el filtro de la verificación de hechos.⁶¹



Nota: Artículos relacionados en Facebook

- En una actualización de agosto de 2017, Facebook anunció que empezará a usar aprendizaje automatizado (*machine learning*) para detectar más contenido engañoso. Facebook podrá mostrar las historias de verificación de hechos debajo de los artículos originales.
- Desde julio de 2017, Facebook eliminó la opción de personalizar la imagen, título o descripción de previsualización de los enlaces que se publiquen en los perfiles o grupos de Facebook. Estas previsualizaciones dependerán únicamente de los metadatos del sitio web que se enlaza.⁶² Esto impide modificar esta información para atraer clics a través de titulares o imágenes engañosas.
- Desde 2013, Facebook anunció la adopción de un nuevo algoritmo para detectar y dar más relevancia al contenido de “alta calidad”. En su momento se consideró, por ejemplo, si el contenido era relevante o si las fuentes eran confiables.⁶³ Sobre el particular, en 2016 Facebook estableció una política para evitar que los anunciantes con páginas de contenidos de baja calidad puedan hacer publicidad en la plataforma.⁶⁴

Facebook también ha anunciado la adopción de medidas para contrarrestar la “falsa amplificación”, a la que entiende como la actividad coordinada que busca manipular la discusión política.⁶⁵ Hacen parte de esta práctica irregular la creación de cuentas falsas (muchas veces ejecutada a gran escala); la distribución coordinada de contenido o mensajes repetidos; los *likes* o reacciones coordinadas; las creación de grupos con el propósito de distribuir noticias o titulares sensacionalistas, y la creación de memes, de videos o de fotos manipuladas; entre otros.

Para este fin, la compañía anunció el desarrollo de herramientas tecnológicas que le permitan identificar los falsos *likes* y falsos comentarios que provengan de cuentas falsas, *malware* o de las llamadas “granjas de clics” (grupos de cuentas que tiene una actividad aparentemente orgánica que se crean para generar una interacción falsa).⁶⁶

⁶¹ Facebook, “News Feed FYI: New Test With Related Articles”, 25 de abril de 2017, disponible en: <http://bit.ly/2q2zxHe>.

⁶² Facebook, “API Change Log: Modifying Link Previews”, disponible en: <http://bit.ly/2u1ndq6>.

⁶³ Facebook, “News Feed FYI: Showing More High Quality Content”, disponible en: <http://bit.ly/2BZYU2U>.

⁶⁴ Ver: “Contenido molesto o de baja calidad”, disponible en: <http://bit.ly/2q9hkFk>, último acceso: 14 de diciembre de 2017.

⁶⁵ Facebook, “Information Operations and Facebook”, 27 de abril de 2017, disponible en: <http://bit.ly/2oOOS9s>.

⁶⁶ Ver: Facebook, “Breaking new Ground in the Fight Against Fake Likes”, 17 de abril de 2015, disponible en: <http://bit.ly/1yCsKDM>; Facebook, “Disrupting a major spam operation”, 14 de abril de 2017, disponible en: <http://bit.ly/2oQJZQX>.

2. Google

Según informó el New York Times en octubre pasado, anuncios de noticias falsas usaron el sistema de publicidad de Google y aparecieron incluso en portales de *fact-checking* como Snopes y Politifact. Un portal falso de Vogue anunciaba, por ejemplo, que la primera dama de Estados Unidos Melania Trump se había rehusado a vivir en la Casa Blanca.⁶⁷ Este, sin embargo, no era el único frente de desinformación para Google. En resultados de su motor de búsqueda han aparecido portales de noticias falsas, y en YouTube se recomiendan con frecuencia videos de teorías de conspiración.⁶⁸

Algunas medidas de la empresa para enfrentar la desinformación están enfocadas en la alfabetización digital sobre el consumo de noticias, como el programa *Internet Citizens* de YouTube, que da talleres a jóvenes entre 13 y 18 años. Por otra parte, Google financia cientos de proyectos en Europa para producir periodismo original y orientar a los ciudadanos sobre contenido confiable.⁶⁹ Más allá de este enfoque, la respuesta en el servicio se ha enfocado en dos frentes: los servicios de búsqueda (Google Search y Google Noticias) y los servicios de publicidad (Google AdWords, Google AdSense y YouTube).

1.1. Cambios en los servicios de búsqueda: Google Search y Google Noticias

Google ha anunciado varios cambios que podrían incidir en la experiencia de los usuarios del servicio de búsqueda: verificación de hechos para los resultados de búsqueda, cambios en el algoritmo de búsqueda y recepción de retroalimentación por parte de los usuarios. Además, Google ha anunciado que hará más transparente la forma de funcionamiento de las búsquedas.

- Verificación de hechos

Desde 2009, Google implementó un sistema de etiquetas para marcar algunos de los resultados que se muestran en las búsquedas de Google News.⁷⁰ Por ejemplo, las etiquetas Contenido de opinión, Blog o Sátira sirven para ayudar a los usuarios a identificar el tipo de contenido con el que se encontrarán al ingresar a un artículo. En octubre de 2016, Google anunció que implementaría la etiqueta Verificación de hechos, con la cual pretende marcar artículos que hayan pasado por un proceso de *fact-checking*.

Por ejemplo, durante meses el presidente Trump ha dicho que su gobierno hará la reducción de impuestos más grande de la historia. Una búsqueda de este tema en Google News (de noviembre de 2017) muestra entre los resultados un artículo de FactCheck.org marcado con la etiqueta Verificación de hechos, donde se explica qué tan veraz y plausible es esa afirmación.

⁶⁷ Wakabayashi, Daisuke y Qiu, Linda, "Google Serves Fake News Ads in an Unlikely Place: Fact-Checking Sites", *The New York Times*, 17 de octubre de 2017, disponible en: <http://nyti.ms/2hO00Ca>.

⁶⁸ Abrams, Dan, "Now Even Google Search Aiding in Scourge of Fake, Inaccurate News About Election 2016", *Mediaite*, 13 de noviembre de 2016, disponible en: <http://bit.ly/2C0uaic>.

⁶⁹ Al respecto, puede verse el proyecto Cross Check y la Digital News Initiative.

⁷⁰ Google, "¿Qué significa cada etiqueta (p. ej., "blog")?", <http://bit.ly/2F2tiYW>, último acceso: 14 de diciembre de 2017.

Lobbying Frenzy Begins on Tax Bill
New York Times · 2h ago

RELATED COVERAGE
Largest Tax Cut in History?
 Fact Check · FactCheck.org · Nov 1, 2017

MORE ABOUT
 Republican Party
 Donald Trump
 Tax reform

GOP predicts few defections on tax vote
The Hill · 6h ago

GOP Tax Plan Puts Billions in Muni Market Savings at Risk
Featured · Governing · 7h ago

Republicans will be raising taxes on some Americans. And there will be a backlash.
Opinion · Washington Post · 9h ago

GOP Rep. Kevin Brady On GOP Tax Plan: Tax Reform, Spending Constraints Needed | Morning Joe | MSNBC
MSNBC

Nota. Bajo el subtítulo “Cobertura relacionada”, aparece la historia de FactCheck.org. Tomado de Google News.

Originalmente, esta función estaba disponible en Estados Unidos y Reino Unido. Sin embargo, en abril de 2017 Google anunció que la etiqueta estaría disponible en todo el mundo y que se extendería a su sistema general de búsqueda Google Search en todos los idiomas.⁷¹ Así, cuando un usuario realice una búsqueda en Google y los resultados arrojen contenidos que han sido verificados, Google mostrará esta información señalando qué se dijo, quién lo dijo, quién verificó la información y el resultado de esa verificación.⁷² El resultado de la verificación no sólo podrá ser verdadero o falso, sino también en su mayoría cierto o parcialmente cierto. Para indexar esta información, Google parte de criterios como la confiabilidad de la fuente, los hechos verificados, las fuentes y las citas, y las conclusiones a la que se llegó por medio de esa revisión.⁷³

Los editores que deseen que sus verificaciones de hechos se muestren en los resultados de búsqueda, pueden señalárselo a Google de dos maneras: i) usando la etiqueta ClaimReview⁷⁴ en el código de la publicación, a partir de las directrices de Google para verificadores de datos; ii) usando el *widget* Share The Facts, que hace las veces de un sello de calidad para las organizaciones participantes y puede incrustarse –como si fuera un tuit– en una página web.⁷⁵ Vemos entonces cómo Google se limita a resaltar las verificaciones hechas por terceros a partir de ciertos criterios, incluso si éstas han alcanzado conclusiones diferentes.⁷⁶ En contraste, Facebook incorpora el trabajo de verificación de algunas organizaciones en sus propias decisiones, bien sea para decidir la prominencia de la historia o hacer una advertencia en particular.

⁷¹ Google, “Fact Check now available in Google Search and News around the world”, 7 de abril de 2017, disponible en: <http://bit.ly/2tZyRSf>.

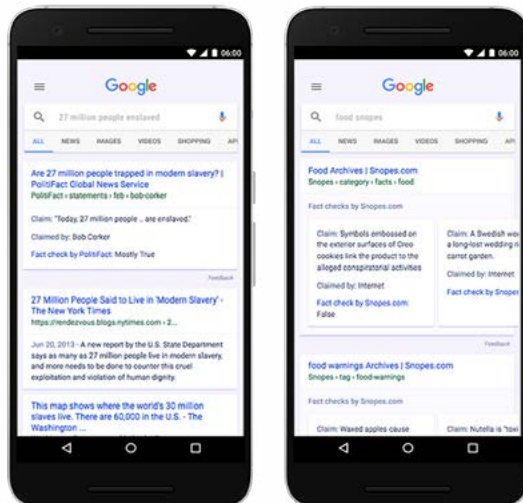
⁷² Distintos verificadores pueden llegar a distintas conclusiones. Varias conclusiones pueden ser presentadas en los resultados.

⁷³ Ver verificaciones de datos en los resultados de búsqueda, <http://bit.ly/2CjpcJ5>, último acceso: 14 de diciembre de 2017.

⁷⁴ Ver: Schema.org, “ClaimReview”, disponible en: <http://bit.ly/2EpwIDP>, último acceso: 14 de diciembre de 2017; Google, “Verificaciones de datos”, disponible en: <http://bit.ly/2CmGPrd>, último acceso: 14 de diciembre de 2017.

⁷⁵ Share the Facts, “Share the Facts”, disponible en: <http://bit.ly/2srvQxi>, último acceso: 14 de diciembre de 2017. A la fecha, trabaja junto con doce organizaciones, ninguna de ellas en Latinoamérica: PolitiFact, The Washington Post, FactCheck.org, Gossip Cop, Pagella Política, AGI, La Stampa, The Ferret, Climate Feedback, Demagog, Newsweek y VERA Files Fact Check (<http://www.sharethefacts.org/about/>).

⁷⁶ Google, “Fact Check now available in Google Search and News around the world”, 7 de abril de 2017, disponible en: <http://bit.ly/2tZyRSf>.



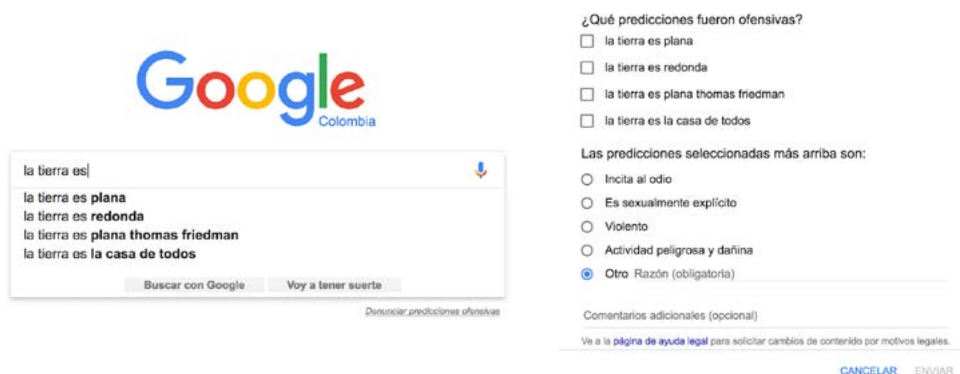
Nota. Ejemplos de cómo aparece la información verificada en las búsquedas en Google.

- Mejoras en el ranking de búsqueda

Aunque Google afirma que apenas el 0,25% de las búsquedas arrojan contenido ofensivo o engañoso, en abril de 2017 anunció cambios en Google Search para que los resultados reflejen contenido más fidedigno.⁷⁷ Esto incluye actualizar las directrices que usan los equipos que evalúan la calidad de los resultados de las búsquedas.⁷⁸

- Retroalimentación de los usuarios

En abril de 2017 Google anunció que mejoraría los mecanismos de reporte de los usuarios en relación con los términos sugeridos en la función de autocompletar y los contenidos destacados en los resultados de búsqueda (*featured snippets*). Según Google, los nuevos mecanismos para obtener retroalimentación incluyen categorías preestablecidas de contenidos ofensivos que faciliten el reporte. Estos mecanismos de reporte no sólo funcionan para señalar noticias falsas, sino también otros contenidos problemáticos (sexualmente explícitos, violentos, peligrosos, etc.).⁷⁹



Nota. Los usuarios pueden enviar comentarios sobre la función de autocompletar haciendo clic en la opción ‘Denunciar predicciones ofensivas’, que despliega varias opciones de reporte.

77 Google, “Our latest quality improvements for Search”, 25 de abril de 2017, disponible en: <http://bit.ly/2sya4ro>.

78 Google, “General Guidelines”, 27 de julio de 2017, disponible en: <http://bit.ly/1ZdLFyy>.

79 Google, “Our latest quality improvements for Search”, 25 de abril de 2017, disponible en: <http://bit.ly/2sya4ro>. Antes del anuncio, era posible para un usuario reportar los términos sugeridos por la función de autocompletar si violaban la Política de Autocompletar de Google, sin embargo, era necesario ir hasta la página de ayuda, encontrar y llenar un formulario (<http://bit.ly/2C0rj95>).



¿Cuál es tu opinión?

- Es útil
- No me gusta esto
- El contenido es racista, ofensivo o incita al odio
- El contenido es vulgar o sexualmente explícito
- El contenido es dañino, peligroso o violento
- El contenido es engañoso o incorrecto

¿Tienes comentarios o sugerencias?

Opcional

Los datos que proporcionas ayudan a mejorar la Búsqueda de Google. [Más información](#)

Si se trata de una cuestión jurídica, [crea una solicitud de eliminación legal](#).

CANCELAR ENVIAR

Nota. Los reportes sobre resultados destacados en forma de snippets pueden hacerse dando clic en la opción 'Comentarios' y luego señalando 'El contenido es engañoso o incorrecto'.

De acuerdo con Google, esta retroalimentación sirve para evaluar si los cambios introducidos en su sistema de búsqueda son exitosos y si deberían aplicarse a los demás usuarios. Además, la empresa dio más información sobre cómo funciona Google Search y divulgó unas breves políticas de contenido para su servicio de autocompletar.⁸⁰

1.2. Cambios en los servicios de publicidad: Google Adwords, Google Adsense y YouTube

En marzo de 2017, una investigación de The Times reveló que anuncios en YouTube de L'Oréal, The Guardian, Nissan, la BBC, entre muchos otros, aparecían mientras se reproducían videos con mensajes extremistas, antisemitas u homofóbicos.⁸¹ De acuerdo con la investigación, quienes publican videos en YouTube generalmente reciben hasta \$7,6 dólares por cada mil reproducciones, y algunos videos tenían millones de vistas. Luego de la revelación, grandes anunciantes como PepsiCo, Starbucks, Walmart, AT&T, Verizon, Volkswagen, Johnson y Johnson, decidieron retirar sus anuncios, en un suceso que fue conocido en los medios de comunicación como el YouTube Ad Boycott.⁸²

En respuesta, Google publicó un comunicado en el que reconocía que, a pesar de los esfuerzos, sus herramientas tecnológicas no siempre detectaban los malos anuncios o los anunciantes que violan sus políticas. Unos días después, la empresa anunció, entre varias medidas, el fortalecimiento de sus políticas sobre contenido de odio u ofensivo, la creación de mecanismos para que los anunciantes controlen dónde aparece su publicidad, y el desarrollo de nuevas herramientas para revisar contenidos cuestionables.⁸³

El problema de los anuncios con contenido indeseable es igual al de las noticias falsas: los anunciantes no quieren que sus productos se asocien a desinformación. Antes de que estallara esa polémica, Google ya estaba trabajando para identificar "anuncios malos": por ejemplo, productos que prometen falsamente ayudar a bajar de peso, productos ilegales o falsificados o anuncios que llevan código malicioso o virus.⁸⁴ Siguiendo esa línea, en noviembre

⁸⁰ Google, "Políticas de autocompletar", <http://bit.ly/2EqUeRb>, último acceso: 14 de diciembre de 2017; Cómo ayudar a mejorar los resultados de búsqueda de Google", <http://bit.ly/2o9oqci>, último acceso: 14 de diciembre de 2017; Google, "How Search works", <http://bit.ly/2rudrP8>, último acceso: 14 de diciembre de 2017.

⁸¹ Mostrous, Alexi, "YouTube hate preachers share screens with household names", *The Times*, 17 de marzo de 2017, disponible en: <http://bit.ly/2HfmUON>.

⁸² Solon, Olivia, "Google's bad week: YouTube loses millions as advertising row reaches US", *The Guardian*, 25 de marzo de 2017, disponible en: <http://bit.ly/2n41Ccw>.

⁸³ Google, "Improving our brand safety controls", 17 de marzo de 2017, disponible en: <http://bit.ly/2uPIDdc>; Google, "Expanded safeguards for advertisers", 21 de marzo de 2017, disponible en: <http://bit.ly/2EJTCJT>; YouTube, "YouTube Partner Program overview", <http://bit.ly/1bQpskt>, último acceso: 14 de diciembre de 2017.

⁸⁴ Google, "How we fought bad ads, sites and scammers in 2016", 25 de junio de 2017, disponible en: <http://bit.ly/2jpTTWP>; Google, "How we fought bad ads, sites and scammers in 2016", 25 de junio de 2017, disponible en: <http://bit.ly/2jpTTWP>.

de 2016 Google introdujo cambios a su política de Google AdSense –sus red de anuncios que aparece en distintos sitios de Internet–, y prohibió la publicación de anuncios en sitios web que tergiversen u oculten información sobre el editor, sus contenidos o el propósito principal del sitio.⁸⁵ Esta política, sin embargo, no se refiere a noticias falsas sino a falsa representación.⁸⁶

En el último informe sobre la ejecución de sus políticas para combatir los anuncios y sitios prohibidos, Google reportó que después de introducir los cambios referidos se tomaron acciones contra 340 sitios por falsa representación y otras ofensas.⁸⁷ Doscientos de estos sitios fueron expulsados definitivamente de la red publicitaria de la empresa. Sin embargo, Google no aclaró cuántos de ellos corresponden a casos de noticias falsas. Lo más cercano al tema aparece en un aparte del informe donde reporta el bloqueo de más de 1.300 cuentas de *tabloid cloak*ers, portales que publican anuncios pretendiendo ser titulares de noticias pero que dirigen a sitios con publicidad.

3. Twitter

A pesar de que tiene el 16% de los usuarios de Facebook, Twitter también ha hecho parte de la polémica alrededor de la desinformación. Según la propia empresa explicó en las audiencias con el Congreso norteamericano, más de 2.700 cuentas asociadas con agentes rusos movieron 1,4 millones de tuits entre septiembre y noviembre de 2016.⁸⁸

Este no es un problema del todo nuevo en Twitter, que es cuestionado constantemente por las cuentas falsas y *bots* en su plataforma. Según la empresa, este tipo de cuentas no representa más del 5% de sus usuarios, pero fuentes externas aseguran que son muchas más. Por ejemplo, de acuerdo con un estudio realizado por Alessandro Bessi y Emilio Ferrara, los *bots* fueron responsables de una quinta parte de los tuits relacionados con las elecciones presidenciales norteamericanas.⁸⁹ No obstante, es importante recordar que Twitter permite las cuentas automatizadas, y muchas de ellas cumplen un servicio relevante en términos informativos (canales noticiosos, servicios gubernamentales) y no se consideran *spam*.

Twitter ha reconocido que es imposible determinar la veracidad de los tuits publicados y, al igual que Facebook, ha sostenido que no quiere ser árbitro de la verdad.⁹⁰ Que Twitter entre a evaluar una información no solo sería indeseable, sino imposible de implementar en la práctica: por la plataforma pasan al menos mil millones de tuits al día. En esta medida, el foco de la respuesta de la empresa es detectar y remover las cuentas que de manera automatizada o manual difunden contenido malicioso (*spam*, falsedades o ataques, entre otros). Algunas de las acciones concretas anunciadas por la empresa, son las siguientes:

- Reducir la visibilidad de tuits y de posibles cuentas de spam mientras se investiga si efectivamente ocurrió una violación.
- Suspensión de cuentas una vez que se ha detectado actividad prohibida.
- Medidas para detectar aplicaciones que abusan de la API pública de Twitter.⁹¹

Twitter considera que cualquier sistema de detección automática conlleva un riesgo alto de falsos positivos; es decir, cuentas que para el algoritmo pueden estar violando las políticas pero que en realidad tienen un comportamiento

⁸⁵ Google, "Prohibited content", <http://bit.ly/2nZvWYe>, último acceso: 14 de diciembre de 2017.

⁸⁶ De hecho, en enero de 2017, la organización Media Matters for America reportó que Google eliminó esta expresión, que antes se incluía como ejemplo de contenido prohibido. "Google Quietly Removes "Fake News" Language From Its Advertising Policy", *Media Matters*, 12 de enero de 2017, disponible en: <http://bit.ly/2BZZY6U>.

⁸⁷ Google, "How we fought bad ads, sites and scammers in 2016", 25 de enero de 2017, disponible en: <http://bit.ly/2kclNnZ>.

⁸⁸ Solon, Olivia y Siddiqui, Sabrina. "Russia-backed Facebook posts 'reached 126m Americans' during US election", 31 de octubre de 2017, disponible en: <http://bit.ly/2hoiMRe>.

⁸⁹ Bessi, Alessandro y Ferrara, Emilio. "Social bots distort the 2016 US presidential election online discussion", 7 de noviembre de 2016, disponible en: <http://bit.ly/2oZNst8>.

⁹⁰ Twitter, "Our Approach to Bots & Misinformation", 14 de junio de 2017, disponible en: <http://bit.ly/2HhmFCC>.

⁹¹ API es una interfaz de programación de aplicaciones que permite que un software externo use varias funcionalidades de un servicio, en este caso, Twitter.

legítimo. Por ejemplo, una cuenta de un activista que tuitea constantemente y que puede confundirse con un *bot* o con una persona que hace *spam* deliberadamente. Por otro lado, la desinformación en la plataforma se hace en desarrollo de acciones conjuntas entre grupos. Se trata de operaciones coordinadas que se adaptan a las reglas y restricciones de la plataforma para conseguir su cometido. “Es mucho más complicado detectar coordinación no-automatizada”, afirma la empresa. “Los riesgos de silenciar inadvertidamente actividad legítima son mucho más altos”.⁹²

V. Conclusión: el problema de la solución

El objetivo principal de este documento era exponer las medidas que han contemplado algunos intermediarios – Facebook y Google principalmente– para combatir la desinformación. Después de explicar el problema y situarlo en el espacio de cada plataforma, describimos las soluciones propuestas, la mayoría de las cuales están en fase preliminar. En esta parte final del texto planteamos algunos de los problemas que conllevan esas soluciones y hacemos algunas recomendaciones. Para llevar a cabo este análisis, proponemos cuatro puntos: i) la escala y el tiempo de la solución; ii) el impacto; iii) el rol de la sociedad civil, y iv) la transparencia. Este análisis parte de un supuesto básico que planteamos al comienzo y reiteramos ahora: no es posible ni deseable que la solución a las *fake news* sea automática. Se trata de un fenómeno de incorporación tecnológica, que es por definición social.

Este análisis se enfoca en el contenido orgánico y no en el promocionado. Las noticias falsas que se promocionan comercialmente suponen unos retos y exigencias distintas para las plataformas. Sería razonable exigirles un mayor nivel de monitoreo y control en ese caso, toda vez que reciben un lucro directo por anunciar información con fines comerciales. El contenido orgánico es aquel que difunden los usuarios entre sus redes de contactos, y en contraste con los anuncios comerciales abarca por igual información espontánea de usuarios, manipulación y acciones coordinadas. Es allí donde encontramos los mayores retos para la libertad de expresión y, por tanto, donde nos queremos centrar.

1. Escala y tiempo

Para analizar estos dos factores, es necesario retomar el concepto de *affordances*. Como se explicaba anteriormente, la respuesta a la desinformación debe partir necesariamente de la configuración del espacio donde se produce el problema. Esto evidencia un asunto estructural: algunos elementos que posibilitan la desinformación solo podrían eliminarse en el diseño mismo del servicio. Se trata de limitaciones inherentes a la arquitectura del espacio. Por ejemplo, supóngase que Facebook introdujera un sistema de revisión previa de todos los contenidos: si fuera posible, eliminaría por igual la desinformación y mucho contenido legítimo. Pero como Facebook no va a modificar ese aspecto sustancial de su estructura y espacio, el “ofrecimiento” de esta plataforma demarca los alcances de la solución.

Tanto Facebook como Google, YouTube y Twitter, “ofrecen” un espacio donde se puede divulgar contenido sin revisión previa. Esto significa que, en la gran mayoría de casos, las estrategias para enfrentar la desinformación serán alguna forma de control posterior: etiquetas, artículos relacionados, menor visibilidad, etc. Ese tipo de control, como vimos, no es susceptible de automatización total, y en esa medida hay un reto en términos de escala y tiempo.

La escala se refiere tanto a los equipos de la compañía como al trabajo con actores externos ¿Puede la empresa replicar las iniciativas contra la desinformación para todos sus usuarios (más de 2.000 millones en el caso de Facebook)? ¿Tiene suficientes equipos humanos internos para revisar todos los contenidos cuestionables? ¿Hay suficientes *fact-checkers* para toda la información disponible? La respuesta invariablemente será negativa. Si la estrategia para enfrentar la desinformación exige un esfuerzo humano –como es deseable, además–, no será posible atender toda la demanda del problema.

⁹² Op. Cit., Twitter, “Our Approach to Bots & Misinformation”. Traducción informal.

En septiembre pasado, Facebook llevó a cabo varias acciones preventivas para evitar la desinformación durante las elecciones alemanas: suspensión de decenas de cuentas, trabajo directo con autoridades locales y campañas educativas.⁹³ Con el antecedente de las elecciones norteamericanas, y esta vez en medio del tenso contexto regulatorio europeo, la empresa no podía quedarse cruzada de brazos.

Difícilmente este tipo de esfuerzos podrían desplegarse en mercados con menor relevancia, como son la gran mayoría de los países latinoamericanos. La escala de la solución, entonces, está enfocada principalmente en países clave, y aunque algunas de las respuestas pueden extenderse a otros mercados, están pensadas y localizadas para esos escenarios prioritarios (lenguaje, abordaje del problema, soluciones, contexto). En el caso de Colombia, como vimos, los reportes de noticias falsas al parecer no han estado disponibles en ningún momento. Los esquemas de *fact-checking*, por otro lado, podrían hacerse en países de la región (Share the Facts, reiteramos, no incluye por ahora organizaciones latinoamericanas), pero requieren de un trabajo mancomunado local que no será fácilmente escalable a toda América Latina.

La limitación en escala se relaciona con la variable de tiempo. El riesgo de la desinformación en el debate público es aún más alto en época de elecciones. Una noticia falsa o un rumor propagado como verdad puede incidir en el resultado de un certamen electoral. En esa medida, es deseable que la respuesta para enfrentar ese problema sea oportuna. Una acción *post-mortem* es relevante para entender el fenómeno, pero no para enfrentar sus efectos inmediatos.

En momentos clave de una campaña, las respuestas descritas en este texto pueden ser tardías. Mientras que construir una mentira es algo rápido, verificar un hecho requiere de tiempo. Resulta fundamental reconocer esta limitación, no tanto como argumento para justificar medidas más restrictivas para el debate público, sino para entender la realidad que enfrentamos y los alcances de las soluciones propuestas.

2. Impacto

La escala y el tiempo de la solución influyen en el impacto de ésta. Por un lado, un conjunto de acciones parciales y aisladas, muchas de ellas tardías, difícilmente servirá para combatir la desinformación en el debate público. Por el otro, estas acciones pueden tener un efecto contrario al que se busca.

Sobre los posibles efectos indeseados, la evidencia es aún precaria. No obstante, algunos estudios señalan que las herramientas de verificación y las advertencias pueden tener un impacto negativo. Según una investigación de un grupo de académicos –aún en etapa de revisión–, marcar historias falsas como tales no cambia necesariamente la percepción del usuario frente a éstas. Y, más grave aún, si un usuario empieza a ver historias donde hay advertencias de posible falsedad, puede concluir que todas las que no tienen advertencia son verdades, lo cual, por supuesto, es una generalización equivocada.⁹⁴

En cuanto al primer punto, este documento no pretende descalificar acciones como el *fact-checking*, las advertencias o los artículos de contexto. Se trata de respuestas que buscan ponderar el problema de la desinformación con la garantía de la libertad de expresión de los usuarios. Descartar medidas restrictivas o abiertamente arbitrarias dificulta la solución del problema, pero evita la creación de otros peores. Bajo esa perspectiva, también es relevante analizar las decisiones de invisibilizar o remover contenidos bajo clasificaciones potencialmente arbitrarias, como la “baja calidad” o no tratarse de una fuente “confiable”. Estos esquemas basados en reputaciones corren el riesgo de favorecer principalmente medios masivos y comerciales, en detrimento de voces que no pueden acceder a ese tipo de alianzas o *sell*, pero cuyo contenido es legítimo y relevante.

⁹³ Facebook, “Update on German Elections”. 27 de septiembre de 2017, disponible en: <http://bit.ly/2Buursp>. Ver también: Locklear, Mallory. “Facebook axed ‘tens of thousands’ of accounts before German election”. Engadget, 27 de septiembre de 2017 disponible en: <http://engt.co/2k1Etdw>.

⁹⁴ Ver, Pennycook, Gordon, et. al. “Prior exposure increases perceived accuracy of fake news”. Versión de 26 de agosto de 2017; Pennycook, Gordon, et. al, “Assessing the effect of “disputed” warnings and source salience on perceptions of fake news accuracy”. Versión de 15 de septiembre de 2017. Estos textos no han tenido revisión de pares.

Las medidas para enfrentar la desinformación se ubican en el contexto de los términos de servicio de la plataforma y, en particular, de la moderación de contenidos. Por lo tanto, siguiendo lo planteado por la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos, “en el diseño y conformación de sus términos de servicio y reglas de comunidad, las empresas no deben limitar o restringir la libertad de expresión de manera desproporcionada o innecesaria”.⁹⁵

3. El rol de la sociedad civil

La mayoría de estrategias para enfrentar la desinformación incluyen la participación de la sociedad civil. De distinta manera, las empresas trabajan con organizaciones para verificar información, hacer advertencias o proveer mayor contexto. Esta participación es fundamental, y en la medida en que estas iniciativas se localicen regionalmente deben abarcar organizaciones que tienen legitimidad y conocimiento para ponderar información sobre el debate público. El usuario, sin embargo, es un convidado de piedra en esta discusión. Salvo las acciones de retroalimentación de Google, no está claro de qué manera la opinión de los usuarios es tomada en cuenta para enfrentar la desinformación en estos servicios.

Por último, frente a este punto es preciso tener en cuenta que la sociedad civil es también un actor en la producción de la desinformación: los partidos políticos, las agencias de comunicación y distintos grupos de interés hacen parte del problema que hoy en día se espera que las plataformas resuelvan directamente. Las acciones concertadas, por ejemplo, evidencian una actuación organizada y sistemática para explotar los servicios en beneficio de un fin individual y en detrimento del interés general.

Si bien esto no exime de responsabilidad a los intermediarios, sí pone de relieve la necesidad de buscar respuestas, además de las preventivas y educativas, en distintos sectores económicos, políticos y sociales. Dicho de otra forma, la solución a este problema no se limita a los actores tradicionales de la gobernanza de internet.

4. Transparencia

La elaboración de este documento tuvo una dificultad constante: entender si la gran cantidad de decisiones y acciones cubiertas en medios de comunicación y anunciadas por las empresas eran efectivamente implementadas y en qué medida. De hecho, mientras escribimos estas palabras se informa sobre medidas que reversarían cambios anunciados hace poco y referenciados en este texto.⁹⁶ En definitiva, la conclusión frente a este punto es que muchas de las medidas se quedan en anuncios, y las que se implementan alcanzan apenas un grado parcial y temporal de aplicación. De cualquier forma, no es posible determinarlo con precisión, lo cual evidencia un grave problema de transparencia.

Es entendible y deseable que las empresas experimenten con posibles soluciones. La desinformación debe enfrentarse con acciones creativas e innovadoras, y las estrategias no pueden ser inamovibles. Sin embargo, la falta de claridad para los usuarios y el público en general oscurece el entendimiento del problema e impide que la sociedad civil pueda hacer retroalimentaciones juiciosas.

La tecnología produce retornos informativos en la medida en que se masifica su adopción. Es a través de ese uso —y del efecto de red— que se obtienen datos que sirven para mejorarla y adaptarla.⁹⁷ Ese retorno social debe masificarse también, o al menos tener algún grado de apertura. El conocimiento sobre cómo funciona la desinformación y las medidas

⁹⁵ Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos, Estándares para una Internet Libre, Abierta e Incluyente OEA/Ser.L/V/II CIDH/RELE/INF.17/17 15 de marzo 2017, Original: Español, Párr. 98

⁹⁶ Cohen, David. “Facebook Is Opting for Context Over Labeling in Its Battle Against Fake News”. *Adweek*, 21 de diciembre de 2017. Disponible en: <http://bit.ly/2Ero77P>.

⁹⁷ Ver, MacKenzie, Donal. *Knowing Machines. Essays on Technical Change*. MIT Press, 1998.

para enfrentarla está controlado por las empresas, que se ven poco dispuestas a compartirlo con la sociedad civil. No hay claridad en el diagnóstico ni en las respuestas adoptadas, todo lo cual se suma a los cuestionamientos ya existentes contra los intermediarios por su falta de transparencia en la forma como prestan su servicio. La transparencia, en esta medida, no es una acción unilateral posterior, sino un proceso conjunto que hace parte de la solución.

