

Internet y derechos humanos III

Aportes para
la discusión en
América Latina

Agustina Del Campo
Compiladora

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



Internet y derechos humanos III
Aportes para la discusión en América Latina

Internet y derechos humanos III

Aportes para la discusión
en América Latina

Agustina Del Campo

COMPILADORA

Facultad de Derecho

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



Internet y derechos humanos III : aportes para la discusión en América Latina / Carlos Cortés ... [et al.] ; compilado por Agustina Del Campo ; editor literario Carla Ortiz Rocha. - 1a ed. - Ciudad Autónoma de Buenos Aires : Universidad de Palermo - UP, 2020.

156 p. ; 16 x 23 cm. - (INTERNET Y DERECHOS HUMANOS)

Traducción de: María Soledad Vázquez.
ISBN 978-950-9887-49-7

1. Internet. 2. Derechos Humanos. 3. América Latina. I. Cortés, Carlos II. Del Campo, Agustina, comp. III. Ortiz Rocha, Carla, ed. Lit. IV. Vázquez, María Soledad, trad.
CDD 341

Compiladora:
Agustina Del Campo

Diseño general:
Departamento de Diseño
de la Universidad de Palermo

Corrección:
Carla Ortiz Rocha

Editado por la Universidad
de Palermo, mayo de 2019,
Buenos Aires, Argentina

© 2019 Fundación Universidad
de Palermo

ISBN 978-950-9887-49-7

Mayo de 2019

Hecho el depósito que marca la
ley 11.723

Cantidad de ejemplares: 350

Impresión: MPA
Luis Sáenz Peña 647 (1110)
Ciudad Autónoma de Buenos Aires

Impreso en la Argentina / Printed
in Argentina

Universidad de Palermo
Rector
Ing. Ricardo H. Popovsky

Facultad de Derecho
Decano
Dr. Leandro Vergara

Centro de Estudios en Libertad de Expresión
y Acceso a la Información (CELE)
Directora
Agustina Del Campo

Mario Bravo 1050
(C1175ABW) Ciudad de Buenos Aires
Argentina
Tel.: (54 11) 5199-4500 | cele@palermo.edu |
www.palermo.edu/cele



Licencia Creative Commons 4.0. Los artículos de este libro se distribuyen bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional. Pueden ser compartidos y adaptados mientras no se haga un uso comercial del material, bajo la condición de reconocer a los autores y autoras y mantener esta licencia para las obras derivadas.

Este libro fue realizado en el marco de un proyecto auspiciado por **Fundación Ford**.

Índice

7 Prólogo

iLEI

9 La implementación del “olvido digital”: el olvido de los detalles

Carlos Cortés y Luisa Isaza

31 Del cómic a los memes: viejas y nuevas problemáticas en torno al humor y la libertad de expresión

Agustina Del Campo y Paula Roko

63 Noticias falsas en Internet: la estrategia para combatir la desinformación

Carlos Cortés y Luisa Isaza

99 Servicios Over-the-Top: principios fundamentales para su tratamiento regulatorio en Argentina

Maia Levy Daniel

Anexos:

121 Moderación de contenidos y censura privada: estándares extraídos de la jurisprudencia del sistema interamericano de derechos humanos

133 Estudio sobre la Junta de Supervisión de Facebook: modificar las expectativas

- 147 Comentarios en respuesta a los cambios
propuestos por Twitter a las reglas sobre “contenido
deshumanizante”

En los últimos años los debates en torno a la regulación de internet han cobrado mayor complejidad y sobre todo visibilidad y relevancia en la agenda pública. En el primer volumen de esta serie se destacaba la incipiente intervención estatal en la regulación de internet y la necesidad de encarar dicha intervención desde una perspectiva de derechos humanos. El volumen II abordó algunos de los temas más relevantes de la agenda legislativa y problemas asociados a la gobernanza de internet de modo más general. Este volumen III ahonda en el estudio de algunos temas que gozan de cierta permanencia en la agenda pública, como el mal llamado “derecho al olvido”, e incorpora otros nuevos temas como la desinformación, la regulación de los servicios Over-The-Top (OTTs) o los desafíos del discurso humorístico en el ámbito online.

El primer artículo, *La implementación del “olvido digital”: el olvido de los detalles*, escrito por Carlos Cortés y Luisa Isaza, estudia la implementación judicial del “olvido digital”, destacando sus diversas acepciones y abordajes a nivel comparado en América Latina. “Existe el riesgo de que la implementación de una orden técnica genere efectos imprevistos, especialmente cuando esta no toma en cuenta la praxis del entorno digital y el contexto en que debe ejecutarse. Ese es el escenario que identificamos en nuestra región: el riesgo de no entender las implicaciones técnicas de una orden puede desembocar en decisiones desproporcionadas o en un incentivo de autocensura”, advierten los autores.

En el segundo artículo, *Del cómic a los memes: viejas y nuevas problemáticas en torno al humor y la libertad de expresión*, las autoras, Agustina Del Campo y Paula Roko, evalúan la regulación pública y privada que actualmente impacta sobre la producción y difusión del humor, particularmente el humor político. Las autoras concluyen que “la imposición de plazos cortos al estilo NetzDG, la resolución de controversias en manos de empresas y no de justicia estatal, y más recientemente los incentivos para la adopción de filtros (previos y posteriores a su carga) sin denuncia o notificación previa, están ocasionando fuertes incentivos para la automatización de procesos de detección, remoción y moderación de contenidos. En este marco, los contextos, los simbolismos y los modismos que caracterizan a las expresiones satíricas, la parodia y el humor en

general –pero particularmente el humor político– se ven indudablemente postergados, invisibilizados, asfixiados.”

Noticias falsas en internet: la estrategia para combatir la desinformación es el tercer artículo en esta publicación. Los autores Cortés e Isaza identifican, sistematizan y contrastan las principales políticas adoptadas por Google, Facebook y Twitter para combatir la desinformación en sus plataformas y concluyen con un llamado a mayor transparencia.

Y el cuarto artículo, *Servicios Over-the-Top: principios fundamentales para su tratamiento regulatorio en Argentina*, de Maia Levy Daniel, analiza las tensiones entre las posiciones de los diversos actores en el debate en torno a la regulación de OTTs, haciendo especial énfasis en la necesidad de enmarcar estas discusiones desde una perspectiva de derechos humanos.

Finalmente, este volumen incluye además tres contribuciones que desde el iLEI preparamos vinculadas con reglas de autorregulación y responsabilidad de las empresas por la protección y promoción de la libertad de expresión. La primera dirigida al Relator de Libertad de Expresión y Opinión de Naciones Unidas en la preparación de su informe sobre moderación de contenidos en internet (2018); la segunda en respuesta a una consulta abierta por Twitter en torno a la posibilidad de modificar sus normas comunitarias referidas a discurso deshumanizante; y la tercera en respuesta a la propuesta de Facebook de crear una Junta de Supervisión (Oversight Board) de expertos independientes tendiente a interpretar y hacer recomendaciones concretas respecto de sus normas comunitarias y su implementación.

“Internet y derechos humanos” se ha convertido en una serie de publicaciones de la iniciativa por la Libertad de Expresión en Internet (iLEI) del CELE. Estas publicaciones intentan constituirse como herramientas útiles para apoyar a legisladores, sociedad civil, periodistas, instituciones gubernamentales y sector privado en la difícil labor que encaran al diseñar, implementar y monitorear las políticas que regulan Internet, particularmente desde una perspectiva de derechos humanos. Este tercer volumen fue realizado gracias al generoso apoyo de la Fundación Ford.

La implementación del “olvido digital”: el olvido de los detalles

Carlos Cortés* y Luisa Isaza**

I. Introducción

“En un mundo conectado, una vida puede ser arruinada en cuestión de minutos y una persona congelada en el tiempo”, afirma la académica norteamericana Meg Leta Jones.¹ Internet nos ofrece ejemplos a diario: un momento humillante que queda capturado en un celular y se multiplica en redes sociales como un virus; un escándalo noticioso que nadie recuerda pero que persigue a su protagonista cada vez que alguien busca su nombre en Google; un viejo tuit que se le enrostra al autor para avergonzarlo... Pa-reciera que el pasado en internet no existiera. Vivimos en un eterno presente.

Llevamos varios años hablando del “derecho al olvido” en internet. Entendido de manera general como la potestad de silenciar eventos pasados de una persona, este “derecho” se viene abriendo paso en nuestra región.² Su origen se remonta, sobre todo, a la protección de los datos personales de un individuo. El referente internacional que permitió instrumentalizar

* Carlos Cortés es investigador 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 la London School of Economics. Es consultor en libertad de expresión y regulación de internet.

** Luisa Isaza es 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.

*** Este artículo fue publicado por el CELE en diciembre de 2018.

¹ Leta Jones, Meg, *Ctrl + Z: The Right to Be Forgotten*, Nueva York, NYU Press, edición para Kindle, 2016, loc. 105. Traducción libre.

² Cfr. Pino, Giorgio, “The Right to Personal Identity in Italian Private Law: Constitutional Interpretation and Judge-Made Rights”, en: Mark Van Hoecke y François Ost (eds.). *The Harmonization of Private Law in Europe*, Oxford, Hart Publishing, 2000, pp. 225-237.

un “olvido digital” se dio en Europa: en 2014, el Tribunal de Justicia de la Corte Europea determinó que el motor de búsqueda de Google debía evitar que al buscar a una persona por su nombre apareciera una noticia antigua referida a una deuda pasada de esta (conocido como el “fallo Costeja”).

En esa decisión, la justicia europea instó a Google a desarrollar un sistema privado mediante el cual los ciudadanos europeos pueden solicitarle al motor de búsqueda que desindexe información inexacta, inadecuada, irrelevante o excesiva sobre ellos. En términos generales, desindexar implica que el motor de búsqueda desvincula de los resultados de búsqueda un enlace en particular.³ Entre mediados de 2014 y finales de 2017, Google recibió solicitudes de desindexación de 2,4 millones de direcciones de internet (URL), de las cuales 43% fueron ejecutadas. El 89% de los casos fueron impulsados por ciudadanos particulares.⁴

El fallo Costeja centró la responsabilidad de garantizar el “olvido digital” en el motor de búsqueda en oposición a quien había producido la información o quien la alojaba. Es decir: según esa decisión, es el intermediario de búsqueda el que debe impedir que la información “dañina” pueda ser asociada al nombre del afectado. La información sigue existiendo, pero encontrarla es mucho más complicado al no hacer parte de los resultados de búsqueda bajo el nombre de esa persona.⁵

Este precedente ha sido incorporado en algunas decisiones administrativas y judiciales en América Latina. Mientras que en el contexto europeo ya existe el sistema privado de desindexación de Google, en nuestra región se trata de casos específicos donde una autoridad da una orden equivalente. En cualquier caso, ni las autoridades administrativas ni judiciales en nuestra región han llegado al punto de ordenar la creación de un sistema privado similar al europeo.

En contraste, otras decisiones en la región centraron la responsabilidad del “olvido digital” en el autor o publicador de la información. En conse-

³ A pesar de que se usa también el término “desreferenciación”, este documento opta por usar la palabra “desindexación”, ya sea total o parcial. Parte del objetivo del documento es que quede claro qué significa esta acción en cada caso.

⁴ Smith, Michee, “Updating our ‘Right to Be Forgotten’ Transparency Report”, 26 de febrero de 2018, disponible en: <https://bit.ly/2SGwtLT>.

⁵ Actualmente se debate si las medidas de desindexación adoptadas por Google con fundamento en esta decisión deberían aplicarse únicamente en sus dominios europeos (como Google.fr, Google.es, etc.) o si, por el contrario, existe el derecho a solicitar una desindexación global. El Tribunal de Justicia de la Unión Europea se encuentra próximo a resolver una cuestión prejudicial planteada por el Consejo de Estado de Francia en este sentido luego de que Google demandara a la Comisión Nacional de Informática y de las Libertades de Francia (CNIL) por haberle impuesto una sanción por limitar las desindexaciones a los dominios europeos.

cuencia, le ordenaron a este que actualizara o eliminara alguna información en particular. En este escenario, el motor de búsqueda no hace parte de la controversia, y si el autor de la información la elimina, esta última simplemente desaparece de los resultados de la búsqueda.

En medio de esos dos tipos de órdenes –aquellas centradas en el motor de búsqueda y aquellas dirigidas al publicador– hay decisiones combinadas, intermedias y abiertamente contradictorias. Al intentar implementar un “olvido digital”, los jueces de América Latina están transitando un territorio desconocido. Entender las implicaciones técnicas y explorar la implementación y sus consecuencias es el objetivo de este documento.

Es importante delimitar el alcance de este análisis: el “derecho al olvido” digital abarca varias discusiones sobre la necesidad de devolverle al individuo el control sobre los datos y la información que de él o de ella circula en internet. En este espacio, queremos enfocarnos en el “olvido digital” que impacta el ejercicio de la libertad de expresión y, en particular, la actividad de los medios de comunicación y de quienes divulgan información y opiniones de interés general. Es decir, no se estudia el “olvido digital” en relación con la discusión sobre la información personal que capturan y conservan las empresas en bases de datos. Nos centramos en las órdenes que resuelven tensiones entre la libertad de expresión y los derechos individuales como el buen nombre, la privacidad y el libre desarrollo de la personalidad, por información que se publica de manera abierta en internet.

El documento se enfoca en decisiones sobre “olvido digital” en su acepción más amplia: las que buscan eliminar, oscurecer o dificultar el acceso a cierta información. En cambio, no abordamos otros componentes de esas órdenes que buscan equilibrar el debate público. Es el caso de las órdenes de actualización de información, de aquellas que otorgan un derecho de réplica o las que exigen la publicación de una advertencia (*flagging*). Valga la pena añadir, sin embargo, que estas últimas órdenes no son excluyentes con las de “olvido digital” (piénsese, por ejemplo, en una disposición que requiera eliminar una información y a la vez exija abrir un espacio de réplica al afectado).

Por otro lado, este documento no aborda críticamente el concepto del “derecho al olvido” ni la legitimidad del agente que lo implementa. Esto ha sido desarrollado en otras publicaciones del Centro de Estudios en Libertad de Expresión y Acceso a la Información (CELE) y en trabajos previos de los

autores.⁶ Consideramos que la idea del “olvido digital” enfrenta muchos cuestionamientos teóricos, y que la incorporación de este “derecho” en el contexto del Sistema Interamericano de Derechos Humanos debe seguirse discutiendo. En esta oportunidad, queremos centrarnos en la pregunta por la implementación.

El “olvido digital” implica una serie de acciones sobre información, datos y contenidos: eliminación, actualización, oscurecimiento e invisibilización. Estas acciones pueden operar en distintos niveles, involucrar a distintos actores y generar variaciones en el tipo de “olvido”. ¿Entienden los jueces, las autoridades administrativas y la sociedad civil interesada la puesta en práctica de estas órdenes? ¿Se está evaluando la proporcionalidad y posibilidad de esas órdenes? ¿Se está analizando el impacto que tienen y los incentivos que generan? ¿Permiten el entorno y las prácticas de las empresas privadas entender los problemas técnicos? ¿Hay una discusión transparente y honesta al respecto?

Para entender dónde se implementan las acciones de olvido, la primera parte de este documento describe los conceptos de memoria transactiva y de sistemas abiertos y cerrados. El propósito es explicar, por un lado, cómo entre más cerca esté el método de “olvido digital” al núcleo de almacenaje, más dominios serán afectados. Y, por el otro, cómo ese “olvido” usualmente involucra la acción de varios actores, los cuales no pueden ser impactados con una sola decisión o medida técnica.

Posteriormente, se aborda el “olvido digital” en la práctica desde los dos agentes involucrados: el motor de búsqueda y el autor de la información. En ambos casos, se hace énfasis en la naturaleza de las órdenes impartidas y su posible implementación. Finalmente, la última parte ofrece unas conclusiones enfocadas en los problemas de proporcionalidad y los incentivos.

Este documento intenta simplificar un asunto complejo en aras de acercar a la discusión a operadores judiciales e integrantes de la sociedad civil. En ese propósito, omite algunos detalles técnicos, pero también parte del supuesto de que la audiencia cuenta previamente con un contexto mínimo sobre el debate del “olvido digital”.

⁶ Véase, Cortés, Carlos, “Derecho al olvido: entre la protección de datos, la memoria y la vida personal en la era digital”, en: Bertoni, Eduardo (comp.). *Internet y derechos humanos. Aportes para la discusión en América Latina*. Buenos Aires, CELE-Universidad de Palermo, 2014; Ferrari, Verónica y Schnidrig, Daniela, “Responsabilidad de intermediarios y derecho al olvido. Aportes para la discusión legislativa en Argentina”, en: Bertoni, Eduardo (comp.). *Internet y derechos humanos II. Aportes para la discusión en América Latina*, Buenos Aires, CELE-Universidad de Palermo, 2016; Botero, Catalina, Camilleri, Michael J. y Cortés, Carlos, “Democracia en la era digital. Libertad de expresión y el derecho al olvido europeo”, en: *El Diálogo. Liderazgo para las Américas*, Informe del Programa de Estado de Derecho, noviembre, 2017.

II. Memoria transactiva y sistemas abiertos y cerrados

A lo largo de los siglos, los seres humanos hemos creado y perfeccionado técnicas y herramientas que apoyan nuestra memoria perecedera. El ejemplo más claro son los objetos externos que “memorizan” información por nosotros: una libreta de teléfonos, la alarma de un reloj o un *post-it* con un recordatorio. En el entorno digital, estos objetos han sido reemplazados por aplicaciones y servicios: los teléfonos almacenados en el celular, la alarma de este o la lista de tareas.

La memoria opera también —o sobre todo— de manera conjunta. En nuestro grupo de amigos o colegas de trabajo ocurre constantemente un proceso descentralizado de almacenamiento y consulta: “Recuerdo que el cumpleaños de Pablo es el 15 de julio porque ese día nació mi sobrina, pero no recuerdo qué día cumple años Felipe”, comenta alguien. “Es el 19 de noviembre, como la canción”, responde otro. En una conversación tan simple como esta se realiza un proceso de memoria transactiva, entendido como la “combinación de mentes individuales y la comunicación entre ellas”.⁷

La memoria transactiva se refiere a cómo se codifica, almacena y recupera información de manera colectiva. De acuerdo con los psicólogos que acuñaron el término, la memoria transactiva se define a partir de dos componentes: (i) un almacenamiento organizado de conocimiento que se alberga enteramente en los sistemas de memoria individual de los miembros del grupo, y (ii) un conjunto de procesos transactivos relevantes para el conocimiento que ocurren entre los miembros.

Al ser un espacio en el que las personas pueden consultar o aportar información, internet es un sistema de memoria transactiva.⁸ Esto resulta relevante para entender el “olvido digital”, cuáles son sus efectos y quién es el agente encargado de adoptar las medidas necesarias para “olvidar”. No estamos enfrentados a la memoria externa de un solo agente sino de múltiples de ellos, cada uno de los cuales puede tener intereses distintos en

⁷ Cfr. Wegner, Daniel M., Giuliano, Toni y Hertel, Paula T., “Cognitive Interdependence in Close Relationships”, en: Ickes, William J. (ed.). *Compatible and Incompatible Relationships*, Nueva York, Springer-Verlag, 1985, p. 256. Traducción libre. Disponible en: <https://bit.ly/2sf9qMQ>, último acceso: 28 de diciembre de 2018.

⁸ La idea de internet como memoria transactiva se popularizó a partir de un famoso estudio publicado en la revista *Science* en 2011 que reveló que las personas tienden a olvidar la información que esperan poder consultar nuevamente usando la red. Sparrow, Betsy, Liu, Jenny y Wegner, Daniel M., “Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips”, en: *Science*, Vol. 333, N° 6.043, 2011, disponible en: <https://bit.ly/2xTCB8i>, último acceso: 28 de diciembre de 2018.

la eliminación o retención de la información en cuestión. “El problema de cumplir con la necesidad de un individuo de ‘ser olvidado’ de un almacenamiento externo de memoria en red es que no se trata del almacenamiento externo del individuo, sino de uno transactivo”.⁹

Las memorias transactivas pueden residir en sistemas de acceso abiertos o cerrados. Según explica Vijfvinkel, en un sistema abierto la información puede ser copiada por cualquiera, sin que exista la posibilidad de hacer seguimiento a las copias.¹⁰ Es el caso, por ejemplo, de un foro abierto de discusión en el que cualquiera puede leer la información publicada por sus participantes. Mientras tanto, en un sistema cerrado la diseminación de la información se limita a espacios controlados, y todos sus operadores están limitados por las restricciones del espacio y por las obligaciones convencionales o legales de no difusión.¹¹ Piénsese en el sistema de consulta y edición de historiales médicos de los pacientes de un hospital.

Vijfvinkel explica que tanto los sistemas de acceso abiertos como los cerrados tienen distintos niveles de operación: primero está la capa del medio físico donde se almacenan los datos —el “dominio del sistema”— y la capa donde se pueden consultar los datos —el “dominio de la organización”—. En el sistema de historiales médicos, el dominio del sistema sería, por ejemplo, el servidor donde aquellos se almacenan, y en el dominio de la organización se moverían los miembros del personal médico y administrativo autorizados para la consulta. En un sistema cerrado como este, donde no hay un espacio para consulta pública, acá terminan los niveles de acceso.

En el sistema de acceso abierto, en cambio, hay capas adicionales donde aumentan las posibilidades de consulta e intercambio de información. Y, entre más abierto sea el acceso al sistema, más transactivo será. Facebook, por ejemplo, requiere de una autenticación del usuario para acceder sus contenidos, esta sería una memoria semitransactiva. Wikipedia, que —en contraste— no requiere esta autenticación ni para consultar los contenidos ni para producirlos, sería una memoria transactiva.

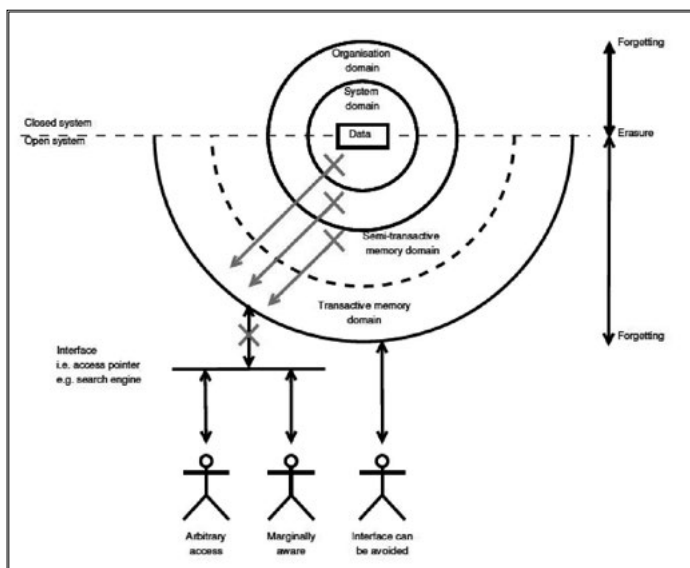
Otro ejemplo de un sistema de acceso abierto de memoria transactiva,

⁹ Korenhof, Paulan, Ausloos, Jef, Szekely, Ivan y otros, “Timing the Right to Be Forgotten: A Study into ‘Time’ as a Factor in Deciding About Retention or Erasure of Data”, en: Gurwirth, Serge, Leenes, Ronald y de Hert, Paul (eds.). *Reforming European Data Protection Law*, Vol. 20, Dordrecht, Springer, 2015.

¹⁰ Vijfvinkel, M.M., “Technology and the Right to be Forgotten”, tesis presentada para obtener el grado de Máster en Ciencias de la Computación, Universidad de Radboud, Nimega, Países Bajos, 2016.

¹¹ Druschel, Peter, Backers, Michael y Tirtea, Rodica, “The Right to Be Forgotten. Between Expectations and Practice”, European Network and Information Security (ENISA), 2012.

relevante además para este documento, es el que se forma entre la página web de un medio de comunicación y el motor de búsqueda que ayuda al usuario a ubicar un artículo periodístico a partir del término que introduce. Aunque el conocimiento se alberga en un lugar –el servidor del medio de comunicación–, el motor de búsqueda genera el proceso para que el conocimiento esté disponible para que los demás usuarios puedan “recordar”.



“Entre más cerca esté el método de ‘olvido’ al lugar de almacenaje de los datos, más dominios serán afectados”, afirma Vijfvinkel.

El sistema de dominios permite entender el tipo de medidas implementadas por el “olvido digital”: “Entre más cerca esté el método de ‘olvido’ al lugar de almacenaje de los datos, más dominios serán afectados”, explica Vijfvinkel.¹² Los efectos de las medidas variarán y serán desiguales entre los usuarios de la información dependiendo de en qué capa se apliquen. Por ejemplo, una orden de eliminación de información en el dominio del sistema –en el archivo digital de un medio de comunicación, por ejemplo– afectará a todos los demás dominios y a todos los usuarios. En cambio, una medida aplicada al dominio de la memoria transactiva –una orden de desindexación al motor de búsqueda– no afectará la información: esta todavía estará disponible para aquellos usuarios que se mueven en el dominio semitransactivo

¹² Vijfvinkel, *supra* nota 10. Traducción informal.

(son suscriptores del medio de comunicación o conocen el enlace de acceso, por ejemplo), o en el dominio de la organización (periodistas del medio de comunicación con acceso al servidor). Sin embargo, unos y otros tendrán que saber cómo buscarla.

Otro aspecto importante del sistema de dominios es entender quién controla la información y quién está realmente en posición para eliminarla del sistema o para controlar la forma cómo se intercambia en un sistema transactivo. Siguiendo con el ejemplo anterior, el motor de búsqueda que indexa las páginas de internet para los resultados de búsqueda no puede modificar o eliminar la información en el dominio del sistema. Por el contrario, si el medio de comunicación que publicó el artículo lo elimina de sus servidores —es decir, del dominio del sistema— esto afectará el resultado del motor de búsqueda. Sin embargo, si esa noticia estuvo disponible en un sistema de acceso abierto, el medio de comunicación no podrá controlar todas las acciones posteriores que se construyan en la memoria transactiva: un artículo publicado en un medio puede compartirse en capturas de pantalla o copiarse y publicarse en el blog de un tercero, y muy posiblemente estos contenidos serán a su vez indexados por el motor de búsqueda.¹³

III. El “olvido digital” en la práctica

Implementar el “derecho al olvido” digital requiere meterse en el cuarto de máquinas de un medio de comunicación o de un intermediario. Implica, en otras palabras, indicarle a estos agentes los términos en que su servicio debe relacionarse con un contenido. Como planteamos en la introducción, muchas de estas decisiones se toman sin tener en cuenta las posibilidades y las consecuencias técnicas. A continuación, tomaremos ejemplos de decisiones judiciales o administrativas para clasificar y analizar esas órdenes desde la responsabilidad que se les asigna a los dos actores relevantes: el motor de búsqueda y el autor o publicador de la información. En este último caso, como dijimos, queremos hacer énfasis en los medios de comunicación, pero haremos alusión a algunas decisiones que se centran en otros agentes y que son relevantes para el presente análisis.

¹³ El medio sí tendría la posibilidad de controlar las acciones posteriores si funciona como un sistema de acceso cerrado, por ejemplo, en aquellos casos en los que haya implementado medidas de gestión de derechos digitales o programas anticopia (también conocidos como DRM por su sigla en inglés para *digital rights management*), a través de los cuales logra controlar el acceso y el uso de sus publicaciones.

1. Desde el buscador

1.1. Desindexación (total y parcial)

En el fallo Costeja, el Tribunal de Justicia de la Unión Europea consideró que los motores de búsqueda (y Google en particular) realizan tratamiento de datos personales al indexar los sitios web de terceros. En consecuencia, son responsables de dar cumplimiento a la regulación europea en esa materia.¹⁴ Y aunque el tribunal reconoció que los autores de contenidos externos –como los medios de comunicación– pueden evitar que su publicación sea indexada por los buscadores (a través de medidas que abordaremos más adelante), concluyó que esa opción no excluye la responsabilidad de los buscadores, especialmente cuando aquellos autores no responden frente al contenido cuestionado.¹⁵

La sentencia considera entonces que es Google quien tiene que adoptar la medida de “olvido digital”. La instrucción técnica que dio el tribunal no fue muy minuciosa: el buscador está obligado a evaluar y, cuando corresponda, eliminar de los resultados de búsqueda –obtenidos a partir del nombre de una persona– los ‘links’ a páginas web que contengan la información cuestionada.¹⁶ Esto significa que, según el fallo, Google no debe desindexar del todo el “link” cuestionado de los resultados de búsqueda. Se trata de una desindexación parcial: Google restringe los resultados de una búsqueda

¹⁴ Tribunal de Justicia (Gran Sala), “Google Spain, S.L., Google Inc. vs. Agencia Española de Protección de Datos (AEPD), Mario Costeja González”, sentencia del 13 de mayo de 2014, disponible en: <https://bit.ly/2sug3Nu>, último acceso: 28 de diciembre de 2018.

¹⁵ El tribunal justificó la responsabilidad del buscador diciendo que: (i) el autor de una página web puede beneficiarse de la excepción que protege el tratamiento de datos para fines periodísticos, (ii) la inclusión en los resultados de búsqueda a partir del nombre de una persona facilita el acceso a la información por parte de cualquier usuario de internet, creando una injerencia mayor en la intimidad que la de la publicación que hace el autor de la página web, y (iii) dado que la información puede ser copiada fácilmente por sitios que no estén sujetos al Derecho de la Unión Europea, no podría lograrse la protección efectiva de los titulares de datos si estos tuvieran que recurrir a tales sitios web.

¹⁶ De acuerdo con el tribunal, la decisión debe tomarse en relación con “links” a páginas web con información verídica, publicada legítimamente por terceros, cuando dicha información es “inadecuada, no es pertinente, o ya no lo es, o es excesiva en relación con los fines del tratamiento en cuestión realizado por el motor de búsqueda”. Y dice el fallo: “El gestor de un motor de búsqueda está obligado a eliminar de la lista de resultados obtenida tras una búsqueda efectuada a partir del nombre de una persona vínculos a páginas web, publicadas por terceros y que contienen información relativa a esta persona, también en el supuesto de que este nombre o esta información no se borren previa o simultáneamente de estas páginas web, y, en su caso, aunque la publicación en dichas páginas sea en sí misma lícita”.

asociada a un nombre en particular, pero mantiene los enlaces si se buscan otros términos relacionados.¹⁷

Ilustremos el punto con un ejemplo: Juan Pérez es un ciudadano que estuvo involucrado hace varios años en el delito de evasión de impuestos, noticia que fue reportada por un medio de comunicación local y que aparece indexada como primer resultado de búsqueda de Google al digitar su nombre. Siguiendo el criterio establecido por el fallo Costeja, la información sobre Pérez es excesiva e inadecuada, de manera que Google debe desindexar parcialmente la noticia: cuando alguien digite “Juan Pérez” no debe aparecer ese resultado en el motor de búsqueda. Pero si alguien realiza una búsqueda y usa términos distintos (“evasión de impuestos” y el año o ciudad donde ocurrió el hecho), el resultado sigue estando disponible.

Basándose en el precedente Costeja, otros tribunales en nuestra región adoptaron criterios análogos o similares. En 2014, por ejemplo, la autoridad de protección de datos de México emitió una orden de desindexación parcial por información de medios de comunicación.¹⁸ La entidad administrativa le ordenó a Google México que se abstuviera de “tratar los datos personales del Titular, consistentes en su nombre y apellidos, de tal manera que, al ser tecleados en el motor de búsqueda del Responsable, no aparezcan los ‘links’ o URL –indexación– que dicho Titular refirió en su solicitud”.¹⁹ En 2016, la decisión fue anulada por un recurso de amparo.²⁰

Las órdenes del tribunal europeo y de la autoridad mexicana están dirigidas al buscador y no al autor del sitio web que publicó la información, lo cual nos sitúa en el dominio de la memoria transactiva. La medida de “olvido digital” no afecta los dominios internos: la información seguirá disponible para quien tenga el “link” directo a la nota o para quien haga una búsqueda en la misma página web del medio de comunicación, pero no estará disponible para quien limite su búsqueda en Google al nombre de la persona afectada.

¹⁷ El tribunal europeo obliga a que aquellas desindexaciones parciales se hagan en todos los dominios europeos de Google Search (Google.fr o en Google.es) e incluso en búsquedas desde dominios no europeos que se hagan dentro del territorio de la persona que solicitó la medida de “olvido digital”.

¹⁸ La autoridad es el Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI).

¹⁹ INAI, expediente: PPD.0094/14, responsable: Google México, S. de R.L. de C.V., 2014, disponible en: <https://bit.ly/1jr0U6E>, último acceso: 28 de diciembre de 2018.

²⁰ “¡Ganamos! Tribunal anula resolución del INAI sobre el falso ‘derecho al olvido’”, Red en Defensa de los Derechos Digitales, 24 de agosto de 2016, disponible en: <https://bit.ly/2VCS0qW>, último acceso: 28 de diciembre de 2018.

1.2. Degradación en el ranking de búsqueda

Un conocido “meme” de internet dice que el mejor lugar para ocultar un cadáver es en la segunda página de los resultados de búsqueda de Google. ¿La razón? Nadie lo buscará ahí. Según un estudio de 2014, el 95% del tráfico de un motor de búsqueda se concentra en la primera página de los resultados.²¹

Bajo esa lógica, el Comisionado de Privacidad de Canadá propuso recientemente que se obligara a los buscadores a bajar ciertos “links” en el ranking de los resultados de búsqueda como forma de proteger la reputación de las personas en internet.²² Sin embargo, a la fecha esta solución no ha sido ordenada por ninguna autoridad.



The best place to hide a dead body
is page 2 of Google search results.

“El mejor lugar para esconder un cuerpo es la página dos de los resultados de Google”.

Esta solución tendría un impacto en el dominio de la memoria transactiva, aunque con un efecto distinto al de la desindexación parcial o total, ya que el “link” se mantendría, incluso si se buscara el nombre de una persona. La pregunta sería si esta forma de “ocultar” el “link” operaría para cualquier búsqueda o si, al estilo de la decisión europea, se bajaría su ranking solo cuando se busque el nombre de la persona interesada. Por ahora, se trata de una propuesta sin mayor desarrollo.

²¹ Jacobson, Madeline, “How Far Down the Search Engine Results Page Will Most People Go?”, en: *Leverage Marketing*, 14 de agosto de 2017. Disponible en: <https://bit.ly/2scfc1F>, último acceso: 28 de diciembre de 2018.

²² Office of the Privacy Commissioner of Canada (OPC), “Draft OPC Position on Online Reputation”, 26 de enero de 2018, disponible en: <https://bit.ly/2CU0rXg>, último acceso: 28 de diciembre de 2018.

2. Desde el autor o publicador

2.1. Eliminación

En enero de 2016, la Tercera Sala de la Corte Suprema de Justicia de Chile le ordenó a un medio de comunicación “eliminar el registro informático de la noticia que afecta negativamente al recurrente, dentro del plazo de tres días”.²³ La información había sido publicada diez años antes, y se refería a la participación de un ciudadano de apellido poco común en un delito. Cumplida la condena, el ciudadano solicitó la eliminación de la información con el argumento de que la noticia ya no era relevante.

Para la Corte, si la ley penal señala la duración de una pena y permite su eliminación de registros públicos una vez esta sea cumplida, es coherente que los medios de comunicación hagan lo mismo y permitan así la reinserción social del individuo.²⁴ Mantener ese registro digital no reporta ningún beneficio para la libertad de expresión –argumenta la Corte– y en todo caso la información “puede ser consultada por métodos análogos mediante el ejercicio investigativo profesional de quien esté interesado en ello”. En todo caso, la Corte aclara que no se trata de eliminar la información de todos los registros digitales, sino de que el acceso a esta se circunscriba a las fuentes oficiales y a los términos que sobre el particular señale la ley.

Acá el agente encargado de implementar la orden es el medio de comunicación que originó la información. Y eso fue lo que hizo el medio chileno: eliminó la noticia sobre la condena penal del ciudadano.²⁵ La eliminación de la información es la medida más radical de “olvido digital”. Si se aplica en el lugar donde se alojan los datos y la información –en el dominio del sistema–, repercute sobre todas las demás capas y afecta a todos los usuarios. En cambio, si la eliminación se aplica únicamente al acceso público, la información seguirá disponible para quienes se mueven en el dominio de la organización (es decir, se mantiene la información en un sistema cerrado).

Sin dejar de lado el carácter radical de la medida, es importante tener en cuenta que al tratarse de información que hizo parte de un sistema abierto,

²³ Corte Suprema de Justicia de Chile, decisión del 21 de enero de 2016, disponible en: <https://bit.ly/2MbDW6m>, último acceso: 28 de diciembre de 2018.

²⁴ La decisión tuvo un voto en contra de la ministra María Eugenia Sandoval, quien resaltó que la noticia trataba un tema de interés público.

²⁵ Desde una perspectiva técnica, la supresión absoluta de información de un sistema no es tan sencilla como parece. Se han propuesto distintos métodos para asegurar la irrevocabilidad de la eliminación, que van desde la destrucción del medio físico y la sobrescritura hasta la encriptación de la información con la posterior eliminación de las claves.

la eliminación en el núcleo no abarca otras ubicaciones descendentes donde aquella hubiera sido replicada, como un blog o en una red social. Estas órbitas exceden el control que el publicador o el autor tuvo inicialmente sobre la información. Una orden judicial o estatal no puede aspirar entonces a que el originador de la información responda técnicamente por esas publicaciones, cuyos controladores son distintos. Igualmente, si el medio tomó la información de otra publicación –es decir, si el dominio del sistema original es otro– el acto de eliminación solo afectará su órbita de republicación.

Un precedente colombiano complementario nos permite entender mejor las limitaciones inherentes a la orden de eliminación. El caso se origina en una deuda no pagada. En 2014, “Esther” decidió exponer públicamente en Facebook a “Lucía” por un préstamo no pagado.²⁶ En una práctica común de humillación en línea –conocida como *online shaming*– “Esther” publicó una foto de “Lucía” en la que denunciaba la deuda y la negativa de esta a responder mensajes y llamadas. Después de que la acreedora se negara a retirar la publicación, “Lucía” le interpuso una acción de tutela (acción de amparo). A pesar de que “Lucía” no negó la existencia de la deuda, la Corte Constitucional confirmó la solicitud de protección. Para este tribunal, la publicación de esta información en Facebook puso a “Lucía” en estado de indefensión y vulneró sus derechos a la intimidad, el buen nombre, la imagen y la honra. En consecuencia, a “Esther” le correspondió eliminar la imagen, el mensaje y publicar una disculpa pública.

La eliminación que llevó a cabo “Esther” afectó el dominio del sistema de Facebook: una vez borrados, son contenidos que no pueden ser recuperados o extraídos por un tercero. No obstante, las capturas de pantalla y la información relacionada con la deuda que se haya intercambiado en otros espacios –una práctica muy común en redes sociales– no se verán afectadas por la eliminación en la fuente y, por lo tanto, no serán impactadas por la orden judicial.

2.2. Seudonomización

En 2014, una ciudadana colombiana interpuso un recurso de amparo contra la Corte Suprema de Justicia por publicar en su página web información sobre ella relacionada con una condena penal por los delitos de concusión, falsedad en documento público y fraude procesal. La reclamante ya había cumplido su pena, pero este antecedente era fácilmente accesible mediante una búsqueda de su nombre en Google.

²⁶ La Corte modificó los nombres de las protagonistas para preservar su privacidad.

La Corte Constitucional acogió las pretensiones de la ciudadana. Argumentó que, si bien las sentencias ejecutoriadas se rigen por el principio de publicidad, deben someterse al marco general de la administración de datos. Mantener ciertos datos disponibles para consulta en internet –afirmó el tribunal– vulnera los principios de finalidad del tratamiento, de acceso y de circulación restringida de datos. En esa medida, se estaba creando un sistema paralelo de consulta de antecedentes judiciales por fuera de los controles que existen en la materia. La Corte Constitucional le ordenó entonces a la Corte Suprema de Justicia que reemplazara el nombre de esa persona por una sucesión de letras o números que impidieran su identificación en las versiones públicas de la sentencia que se encuentran en internet.²⁷

Esta decisión del tribunal colombiano es una orden de seudonomización, que consiste en ocultar el nombre real de una persona, al cambiarlo por otro o al hacer públicas únicamente sus iniciales. Una orden de este tipo debe ser ejecutada por el autor del sitio web, y operaría en el dominio del sistema si se aplica a todas las versiones del documento. Por el contrario, se limitaría al dominio de la organización si se mantiene una versión íntegra del documento para los miembros de esta, pero se hace pública una versión seudonomizada.

La seudonomización ha existido en los medios de comunicación como una forma de protección a fuentes y a sujetos mencionados en artículos periodísticos. Esta medida autorregulatoria, desplegada de manera voluntaria por el medio a partir de su capacidad operativa, podría prevenir un eventual reclamo de “olvido digital”. Al proteger una identidad desde el comienzo, el medio decide evitar la exposición de un individuo. Según cómo se decida implementar, esta medida operaría en el dominio del sistema o de la organización.

Si bien esta orden no iba dirigida a un medio de comunicación, es relevante por el tipo de precedente que establece para la actividad periodística. Por un lado, la búsqueda de información de interés público sobre un ciudadano se vuelve más compleja. De hecho, luego de esta sentencia y ante solicitudes similares, la Corte Suprema generalizó la regla: cuando se compruebe que una condena se cumplió o prescribió, se deben suprimir los nombres de las personas de las bases de datos de acceso abierto.

Según la decisión, “las sentencias condenatorias que expida la Sala o los autos en los que haga referencia a ellas (...) se ofrecerán íntegras a la comunidad en su servidor de acceso público”, esto es con los nombres completos de los procesados. Sin embargo, “cuando se compruebe que judicialmente

²⁷ Corte Constitucional de Colombia, sentencia T-020 del 27 de enero de 2014. Magistrado ponente, Luis Guillermo Guerrero, disponible en: <https://bit.ly/2AAZ3as>, último acceso: 28 de diciembre de 2018.

se declaró cumplida o prescrita la pena, se suprimirán de las bases de datos de acceso abierto los nombres de las personas condenadas, salvo en los eventos en que la ley obligue a conservar pública esa información en todo tiempo”.²⁸ Esta versión, reitera la Corte, no afectará los archivos internos de la entidad, lo cuales podrán consultarse físicamente en desarrollo de las normas de acceso a la información. Siguiendo la clasificación planteada en el documento, se trata de una decisión que afecta al sistema de la memoria-transactiva, pero no el dominio del sistema ni de la organización.

2.3. Desindexación (total y parcial)

En agosto de 2000, el diario colombiano *El Tiempo* publicó la nota “Empresa de trata de blancas”, donde informaba, entre otras, que la ciudadana “Gloria” —como fue identificada en el expediente— había sido vinculada a una investigación penal por el delito de tráfico de personas. En 2013, “Gloria” inició una acción de amparo contra *El Tiempo* y Google, toda vez que el proceso en su contra había finalizado por prescripción y la noticia seguía publicada —fácilmente accesible a ella a través de Google—. En 2015, la Corte Constitucional de Colombia abordó el caso.²⁹

El alto tribunal descartó dos soluciones antes de resolver el caso. Por un lado, liberó al motor de búsqueda de cualquier responsabilidad, con lo cual planteó un precedente en contravía al fallo Costeja. Para la Corte, obligar a Google a que desindexe resultados de búsqueda lo convertiría en un censor de contenidos. Esto podría comprometer la libertad de expresión e información y afectar la arquitectura de internet. Por otro lado, y también bajo la óptica de la libertad de expresión, la Corte consideró que una orden de eliminación de contenidos no se ajustaba a los estándares nacionales e internacionales.

Estaba claro entonces que la Corte iba a centrar la responsabilidad en el autor del contenido. Sin embargo, afirmó que una actualización de la noticia (con la introducción de la aclaración de que la acción penal había prescrito a favor de “Gloria”, por ejemplo) era insuficiente para evitar la estigmatización de la afectada. Por lo tanto, ordenó al medio que, además de hacer la actualización, debía tomar las medidas para que el motor de

²⁸ Corte Suprema de Justicia, sentencia del 19 de agosto de 2015. Magistrada ponente Patricia Salazar Cuéllar. Ficha de resumen disponible en: <https://bit.ly/2Rydl9Q>, último acceso: 28 de diciembre de 2018.

²⁹ Corte Constitucional de Colombia, sentencia T-277 del 12 de mayo de 2015. Magistrada ponente María Victoria Calle, disponible en: <https://bit.ly/1iQCR1b>, último acceso: 28 de diciembre de 2018.

búsqueda no indexara ese contenido cuando se hiciera una búsqueda del nombre de “Gloria”.

Se trata de una orden de desindexación parcial igual a la del fallo Costeja, pero en cabeza de un agente distinto. En este caso, el tribunal requiere que el medio de comunicación le dé la instrucción al motor de búsqueda de que no indexe una página web específica de su sitio cuando se busque un nombre. El efecto debería ser el mismo del precedente europeo: la noticia debe ocultarse en los resultados de Google si se busca el nombre del reclamante (“Gloria Pérez”, por ejemplo), pero debe mostrarse para otros términos relacionados con la información (“tráfico de personas en Colombia”). Sin embargo, que el responsable sea el medio de comunicación y no el motor de búsqueda tiene implicaciones técnicas que no fueron previstas por la Corte.

La orden dictada por la Corte es técnicamente imprecisa. El fallo dispone que el medio de comunicación, “por medio de la herramienta técnica ‘robots.txt’, ‘metatags’ u otra similar, neutralice la posibilidad de libre acceso a la noticia ‘Empresa de trata de blancas’ a partir de la mera digitación del nombre de la accionante en los buscadores de internet”.³⁰ El problema es que el protocolo “robots.txt”, el uso de etiquetas u otra herramienta similar, no permiten un desindexación parcial que excluya ciertas palabras del contenido de un “link”. El uso de estos protocolos sí permite darle instrucciones de indexación al motor de búsqueda, pero en un sentido distinto.

2.3.1. “Robots.txt” y “metatags”

El protocolo de exclusión de robots —conocido como el protocolo “robots.txt”— es la forma en la que una página de internet les informa a los rastreadores (*crawlers*) de los motores de búsqueda qué páginas de su sitio web no desea que sean indexadas. Que no sean indexadas implica que, en principio,

³⁰ En este punto, la Corte acogió el argumento del motor de búsqueda: “Al respecto, la Sala toma nota de la respuesta de Google Colombia Ltda., en la cual se señala que por medio del uso de herramientas como técnicas como ‘robots.txt’ y ‘metatags’ es posible para los titulares y administradores de un sitio web impedir que contenidos específicos sean mostrados como resultados al realizar una consulta por medio de un buscador de internet. En relación con este asunto, la vinculada manifiesta que ‘utilizando la herramienta robots.txt lo que se logra es que un determinado contenido no sea rastreado por el buscador. Sin embargo, pese al uso de dicha herramienta, el buscador sigue reconociendo que el buscador existe y, por ende, puede llegar a mostrar el título de la nota o URL en los resultados de búsqueda, aunque no se podría acceder a la misma por no estar indexado el contenido’. De igual forma, en relación con el uso de metatags, expresa que ‘(...) lo que se logra es que un determinado URL, pese a ser indexado, no sea mostrado como resultado de búsqueda’”.

no aparecerán en los resultados de búsqueda del motor bajo ningún término de búsqueda.³¹ Es decir que la consecuencia es una desindexación total.

El protocolo “robots.txt” surgió como una respuesta técnica para aquellos administradores de sitios web que querían que sus contenidos tuvieran cierto grado de privacidad sin ser necesariamente secretos. De la misma manera en que dos personas pueden tener una conversación en un sitio público que eventualmente puede ser escuchada, una página que no quiere ser indexada desea permanecer relativamente oculta, pero no enteramente privada.³² El académico norteamericano Jonathan Zittrain explica que “robots.txt” no es un estándar oficial: no hay garantía de que los rastreadores actuales o futuros lo respeten.³³ No obstante, en la lógica autorregulatoria y consensuada de internet, los buscadores más populares acatan las solicitudes de no indexación incluidas en “robots.txt”.³⁴

El “robots.txt” funciona a través de un archivo de texto que se inserta en la raíz del sitio web y en el que se incluyen las direcciones web (URL) que desean mantenerse desindexadas. La exclusión de la URL es total: no es posible, y esto es importante reiterarlo, definir a partes de una página que deben ser indexadas y a otras que no. Si suponemos que el motor de búsqueda es un gran directorio telefónico, el “robots.txt” solo permite solicitar que un apartamento de un edificio en particular no aparezca enlistado, pero no permite escoger en qué casos sí debe aparecer el apartamento y en cuáles no.

³¹ Decimos “en principio” porque el motor de búsqueda sí puede llegar a incluir en sus resultados páginas que el “robots.txt” pedía no incluir. Esto sucede cuando muchas otras páginas de terceros –que sí fueron indexadas– incluyen ese enlace en su contenido, con lo cual el algoritmo considera que es un resultado relevante. En este caso, el enlace a la página en cuestión aparecerá en la lista de resultados, pero sin la descripción que suele acompañarlo. En esta medida, no aparecerá como consecuencia directa del contenido que se encuentre en esa página, sino indirectamente por haber sido referenciada externamente por cuenta de este contenido. Técnicamente, no se trata de “indexar” sino de “enlistar”. De Valk, Joost, “Preventing your Site from Being Indexed, the Right Way”, junio de 2017, disponible en: <https://bit.ly/2AyaCiG>, último acceso: 28 de diciembre de 2018.

³² De acuerdo con Zittrain, la lección de “robots.txt” es que la creación de un estándar básico y sencillo puede avanzar significativamente en resolver o anticipar un problema con importantes dimensiones éticas y legales. Zittrain, Jonathan, “Privacy 2.0”, en: *University of Chicago Legal Forum*, Vol. 2008, artículo 3, 2008, p. 104.

³³ *Ibíd.*

³⁴ En la página de ayuda del buscador de Google puede consultarse “Información sobre los archivos robots.txt” en la sección “Cómo bloquear URLs con robots.txt”, disponible en: <https://support.google.com/webmasters/answer/6062608>, último acceso: 28 de diciembre de 2018. En Yahoo, puede consultarse “Elimina tu sitio web del índice de búsqueda”, disponible en: <https://es.ayuda.yahoo.com/kb/SLN2214.html?guccounter=1>, último acceso: 28 de diciembre de 2018. En Bing, “How to Create a Robots.txt File”, disponible en: <https://binged.it/2FcaCcc>, último acceso: 28 de diciembre de 2018.

```

User-agent: *
Allow: /ade/public/
Allow: /svc/news/v3/all/pahb.rss
Disallow: /ade/
Disallow: /ade/bin/
Disallow: /archives/
Disallow: /auth/
Disallow: /cnet/
Disallow: /college/
Disallow: /external/
Disallow: /financialtimes/
Disallow: /idg/
Disallow: /indexes/
Disallow: /library/
Disallow: /nytimes-partners/
Disallow: /packages/fish/multimedia/TEMPLATES/
Disallow: /pages/college/
Disallow: /paidcontent/
Disallow: /partners/
Disallow: /register/
Disallow: /thetreet/
Disallow: /evo
Disallow: /video/embedded/*
Disallow: /web-services/
Disallow: /gst/travel/travsearch*
Disallow: /1996/06/17/nyregion/guest-at-diplomat-s-party-accused-of-rape.html
Disallow: /*.asp.html$

User-agent: googlebot
Allow: /*.asp.html$

User-agent: bingbot
Allow: /*.asp.html$

User-Agent: omgilibot
Disallow: /

User-Agent: omgill
Disallow: /

Sitemap: https://www.nytimes.com/sitemaps/www.nytimes.com/sitemap.xml.gz
Sitemap: https://www.nytimes.com/sitemaps/sitemap_news/sitemap.xml.gz
Sitemap: https://www.nytimes.com/sitemaps/sitemap_video/sitemap.xml.gz
Sitemap: https://www.nytimes.com/sitemaps/www.nytimes.com_realstate/sitemap.xml.gz
Sitemap: https://www.nytimes.com/sitemaps/www.nytimes.com/2016_election_sitemap.xml.gz

```

Página de “robots.txt” de nytimes.com. La orden “disallow” da la instrucción de no indexación de una URL en particular.

Por su parte, las “metatags” son etiquetas (en código de programación conocido como HTML) que se utilizan para consignar información importante sobre un sitio web: la descripción de la página, las palabras clave y el nombre del autor, entre otros. Esta información se incluye en la cabecera del código de la página y no en el cuerpo de esta, por lo cual se trata de información invisible para el usuario. El uso de la “metatag” y de los “robots” con valores como no “noindex” o “nofollow” permite dar instrucciones mediante el uso de etiquetas o categorías, por ejemplo, para evitar que una o varias URL sean indexadas o rastreadas por los motores de búsqueda.³⁵

El “robots.txt” y las “metatags” no permiten dar cumplimiento a la orden de la corte colombiana. Tampoco parecen existir otras herramientas “similares” –como sugiere el tribunal– para lograr una desindexación parcial. En círculos especializados se mencionan etiquetas que permiten instruir al motor de búsqueda para que no indexe algunas partes de una página web.³⁶ Estas

³⁵ Google Webmaster Central Blog, “Using the robots meta tag”, 2007, disponible en: <https://bit.ly/2AvMRHV>, último acceso: 28 de diciembre de 2018. Wikipedia, “Noindex”, disponible en: <https://bit.ly/2CX37Uh>, último acceso: 28 de diciembre de 2018.

³⁶ Wikipedia, “Noindexing Part of a Page”, disponible en: <https://bit.ly/2LVC9z9>, último acceso: 28 de diciembre de 2018.

etiquetas, sin embargo, no han sido reconocidas hasta ahora por Google ni están presentes en la discusión técnica sobre las soluciones. Y, en cualquier caso, no se trata de herramientas que permitan dictarle al motor de búsqueda bajo qué palabras puede enlistar o indexar el contenido.³⁷

Teniendo este contexto presente, ¿cuál fue la solución en el caso colombiano? Enfrentado a una orden que no podía cumplir técnicamente, *El Tiempo* optó simplemente por seudonomizar el nombre de la reclamante: “La Fiscalía acaba de capturar a 16 personas sindicadas de cometer el delito de trata de blancas y concierto para delinquir, entre ellos la señora Gloria”, dice un apartado de la nueva versión del texto.³⁸

IV. Los efectos no deseados del desconocimiento técnico

Uno de los cuestionamientos teóricos que enfrenta el “olvido digital” es el de la proporcionalidad de su protección. En relación con el derecho a la información, hablamos de que una medida es proporcional cuando esta “no implica un costo para la libertad de expresión mayor al beneficio alcanzado”.³⁹ El reparo, entonces, es que una orden de desindexación encaminada a proteger la privacidad o la reputación de una persona pueda terminar afectando de manera desproporcionada el acceso a información de interés general.

Existe el riesgo de que la implementación de una orden técnica genere efectos imprevistos, especialmente cuando esta no toma en cuenta la praxis del entorno digital y el contexto en que debe ejecutarse. Ese es el escenario que identificamos en nuestra región: el riesgo de no entender las implicaciones técnicas de una orden puede desembocar en decisiones desproporcionadas o en un incentivo de autocensura. Enfrentados a deci-

³⁷ No podemos decir de forma tajante que no existe una solución técnica posible. Lo cierto es que en el estado de la discusión y con las herramientas y conocimiento a nuestro alcance no encontramos una respuesta, lo cual resulta diciente frente a la dificultad que encontraría un medio de comunicación para dar cumplimiento a una orden de esta naturaleza.

³⁸ “Empresa de trata de blancas”, *El Tiempo*, 2015, disponible en: <https://bit.ly/2AB6H4S>, último acceso: 28 de diciembre de 2018.

³⁹ Botero, Camilleri y Cortés, *supra* nota 6, p. XXX. Este es el estándar del sistema interamericano según el cual las limitaciones a la libertad de expresión no solo deben ser necesarias en una sociedad democrática, sino también estrictamente proporcionales. Según la Corte Interamericana de Derechos Humanos, el sacrificio inherente a la limitación de la libertad de expresión no debe resultar desmedido frente a las ventajas que se obtienen de la limitación de este derecho (Caso Kimel vs. Argentina, sentencia del 2 de mayo de 2008, § 83).

siones técnicas contradictorias o de imposible cumplimiento, un medio de comunicación puede optar, en el mejor de los casos, por modificar textos del archivo periodístico más allá del alcance del mandato judicial. Y, en el peor, simplemente puede preferir eliminar la información.

El antecedente colombiano del caso “Gloria” pone de presente el problema de no tomar en cuenta esa órbita técnica. Por esa misma vía –pero en una dimensión más preocupante– la autoridad de protección de datos peruana viene tomando decisiones sobre “olvido digital” sin un fundamento técnico adecuado. En una orden reciente contra el diario *El Comercio*, la autoridad de ese país ordenó la actualización del título de una nota, pero parece asumir también que es el medio de comunicación el agente que lleva a cabo la indexación.⁴⁰

Además de los problemas de fondo de estas decisiones –que suelen impactar denuncias de corrupción que no han sido resueltas judicialmente–, las órdenes están incentivando que los medios opten por eliminar el contenido para zanjar la controversia judicial.⁴¹ Ante la confusión y la ambigüedad, el periodismo encuentra en la sustracción de contenido una estrategia de defensa.

Las acciones para garantizar un “olvido digital” implementadas directamente desde las salas de redacción, no deben descartarse. Puede pensarse en alternativas que resulten menos gravosas para la libertad de expresión y el ejercicio periodístico. Una de ellas sería, por ejemplo, que el medio de comunicación publicara tanto una versión parcial –o seudonomizada– del artículo como una completa. Mientras que la primera sería indexada por el motor de búsqueda, la segunda incluiría alguna instrucción de no

⁴⁰ Dice un aparte: “Dicha actualización ha sido colocada líneas abajo del título de la noticia, lo que origina que el buscador de internet ‘Google’ –dado que el ‘El Comercio’ ha indexado el título de la noticia– dé como primer resultado, al digitar el nombre del reclamante, el referido título”. Dirección de Protección de Datos Personales, Ministerio de Justicia y Derechos Humanos de Perú, resolución directorial N° 453 del 12 de marzo de 2018.

⁴¹ Una búsqueda en Google con el titular del caso mencionado –“Renzito: ex asesor ministerial sería uno de sus testaferros”– permite constatar esto. Adicionalmente, en el archivo digital del periódico de la fecha original de publicación tampoco aparece. Ver, <https://bit.ly/2TwOpsn>, último acceso: 28 de diciembre de 2018. La información, sin embargo, sigue existiendo, ya que otros portales la copiaron. Es un ejemplo de un contenido que estaba en el dominio de la memoria transactiva y que, por lo tanto, subsiste a pesar de la eliminación. Este tipo de solución no es la excepción en los medios. En junio pasado, uno de los autores conversó en Lima con abogados y periodistas de varios medios de comunicación sobre el tema. Varios de ellos aceptaron que la eliminación de contenidos se ha convertido en una forma de defensa frente a órdenes administrativas de “olvido digital”.

indexación.⁴² Que la versión completa no sea indexada por el motor de búsqueda afectaría parcialmente la memoria transactiva: la información no podrá encontrarse si se busca el nombre de la persona, pero sí mediante otros términos relacionados porque referenciaría a la versión parcial —o seudonomizada—. No obstante, sí permitiría preservarla en los dominios del sistema y de la organización del medio de comunicación (lo cual al parecer no sucedió en la resolución final del caso “Gloria”). Es decir, la solución no desembocaría en la eliminación total de la información o su indisponibilidad para acceso abierto.

Este tipo de soluciones deben analizarse desde la realidad de las salas de redacción, ya que ejecutarlas demandaría la reorganización de procesos internos y cierto grado de desarrollo en programación. No se trata, por supuesto, de prestaciones de imposible cumplimiento. Una exigencia generalizada puede generar costos altos o barreras de entrada para la generación y la difusión de información.

V. Conclusión

A la ya compleja cuestión de qué es el “derecho al olvido” y cuáles son sus alcances en el entorno digital, se suman las cuestiones técnicas de cómo garantizarlo y qué efectos tendría. Este documento intentó acercarse a esa pregunta por la implementación.

El abordaje desde la memoria transactiva y los dominios de la información subraya dos elementos de análisis relevantes: (i) entre más cerca esté el método de “olvido digital” al núcleo de almacenaje, más dominios serán afectados y, por lo tanto, mayor será la restricción a la libertad de expresión y el acceso a la información, y (ii) en muchos casos el “olvido digital” abarca una “memoria” que se construye de manera colectiva entre muchos actores.

⁴² Estas dos versiones podrían servir para inventarse un sistema de desindexación parcial desde el medio de comunicación que sea menos restrictivo para la libertad de expresión. Como decíamos en el texto, la versión indexada en Google sería aquella con la información parcial o seudonomizada. Sin embargo, cuando alguien hiciera clic en ese enlace, la página del medio de comunicación podría redirigir al usuario a la información completa que no fue indexada por el motor de búsqueda. Este tipo de instrucciones de redireccionamiento son muy comunes en la navegación de internet. Así, si alguien busca a “Gloria” por su nombre propio, la noticia no aparecería en los resultados de Google. Pero si alguien buscara “trata de blancas”, aparecería indexada la nota modificada y, al hacer clic, redirigiría al usuario a la información completa. Esta idea surgió de una conversación con el programador colombiano David Avellaneda.

El abordaje técnico nos ofrece elementos para entender la proporcionalidad desde la dimensión concreta del impacto de una orden judicial o administrativa, y nos permite entender que, en la práctica, no hay una solución absoluta para un “olvido digital”. En desarrollo del rol que juega el Estado en el impulso de cambios tecnológicos que respondan a expectativas sociales, los jueces y las autoridades administrativas deben tomar en cuenta esta parte de la discusión a la hora de tomar decisiones.

La participación del Estado y de la sociedad civil en este debate requiere, sin embargo, de una mayor transparencia de parte de los intermediarios de internet –y, en particular, de los motores de búsqueda–. Hablar de las implicaciones técnicas de una decisión suele confundirse con la inevitabilidad de una configuración presente o con la imposibilidad de impulsar cambios. No obstante, de lo que se trata es de entender el punto de partida para encontrar alternativas. En ese proceso, la responsabilidad de los intermediarios de internet en el diseño de los algoritmos no puede pasarse por alto: “Mientras que los algoritmos representan cálculos y procesamientos que ningún humano podría hacer por sí solo, en última instancia los humanos son los árbitros de los insumos, del diseño del sistema y de los resultados”.⁴³

Integrar la órbita de la implementación del “olvido digital” en la discusión sustantiva permitirá además situarla en el contexto del impacto y de la necesidad de las medidas tecnológicas de olvido. Es decir, posibilitará una conversación también alrededor de la precaución. “La pregunta sobre recordar y olvidar no es una pregunta que deba resolverse en un momento determinado, sino una que la sociedad debe mantener abierta para que las futuras generaciones decidan bajo sus propias circunstancias”.⁴⁴ En medio del presente frágil y el pasado conflictivo de nuestros países, mantener esta pregunta abierta es aún más relevante.

⁴³ Caplan, Robyn y otros, “Algorithmic Accountability: A Primer”, en: *Data and Society*, abril, 2018, p. 10. Traducción informal.

⁴⁴ Thouvenin, Florent y otros, “Remembering and Forgetting in the Digital Age. A Position Paper”, Open Repository and Archive, Zúrich, Universidad de Zúrich, marzo, 2016, p. 5. Traducción informal. Este documento es la base de un libro con el mismo nombre publicado en 2018.

Del cómic a los memes: viejas y nuevas problemáticas en torno al humor y la libertad de expresión

Agustina Del Campo* y Paula Roko**

I. Introducción

En 2018, la campaña “Save the meme” despertó la atención global de la comunidad de activistas por los derechos humanos en internet.¹ A través del lema “Sin memes no hay democracia” numerosas organizaciones internacionales se unieron para recolectar firmas contra la reforma de la directiva europea sobre derechos de autor por considerarla una seria amenaza para la libertad de expresión.² El Relator Especial para la Libertad de Expresión de Naciones Unidas también expresó su preocupación sobre las consecuencias negativas que acarrearía esta norma y llamó públicamente a proteger las

* Agustina Del Campo es directora del CELE. Abogada y magister en Derecho Internacional y Derechos Humanos de American University Washington College of Law. Es profesora de grado y posgrado en derecho internacional, derechos humanos e internet.

** Paula Roko es investigadora del CELE. Abogada y periodista. Se especializó en Periodismo de Investigación en la Universidad del Salvador y actualmente cursa la Maestría en Derecho Constitucional y Derechos Humanos de la Universidad de Palermo.

*** El documento contó con los comentarios de Carolina Botero, directora de Fundación Karisma en Colombia y Vladimir Cortés, Oficial del Programa de Derechos Digitales en Artículo 19 México.

¹ Campaña “Save the meme”, disponible en: <https://savethememe.net/es>, último acceso: 4 de noviembre de 2019.

² Europapress, “Una nueva campaña carga contra la reforma de la ley de copyright de la UE bajo el lema ‘Sin memes, no hay democracia’”, 7 de marzo de 2017, disponible en: <https://bit.ly/36TBYyr>, último acceso: 4 de noviembre de 2019; Cadena Ser, “‘Sin memes no hay democracia’, la nueva campaña contra la reforma de la ley de copyright de la UE”, 8 de marzo de 2017, disponible en: <https://bit.ly/2X4hJJE>, último acceso: 4 de noviembre de 2019.

expresiones humorísticas.³ En junio de 2019, el *New York Times* optó por dejar de publicar caricaturas políticas en su edición internacional luego de una acusación de antisemitismo.⁴ El problema se desató por una viñeta en la que aparecía el primer ministro israelí como un perro lazarillo que conduce a un Donald Trump ciego y con kipá, lo que generó un aluvión de críticas.

Parecería ser que el humor no está en su mejor momento. ¿Acaso se está acotando el espacio para este tipo de expresiones? ¿Cuáles podrían ser las causas de esta constricción y cómo encaja dentro de una perspectiva histórico-jurídica? El presente artículo se propone indagar en las principales tensiones entre el humor y otros derechos, muchas de las cuales parecen surgir o acentuarse a partir de la explosión de los “memes” en internet en los últimos años. En primer lugar, el documento explorará las particularidades del discurso humorístico, desde su concepción más tradicional hasta las más actuales acepciones, que vienen acompañadas de la llamada “cultura del meme”. Luego se analizará el impacto de algunas medidas recientes, tanto estatales como del sector privado, que con distintos propósitos regulan la difusión y la circulación de contenido en internet y por ende están teniendo un importante impacto en la circulación del discurso humorístico.

II. El humor y sus particularidades

El discurso humorístico, y en especial el humor político, ha sido históricamente concebido como un instrumento de denuncia y crítica social desde distintos y variados medios: el periodismo, la literatura, la TV, el cine, la

³ En un documento dirigido a la Comisión europea en el marco de las discusiones previas a la aprobación de la Directiva sobre derechos de autor, el Relator para la Libertad de Expresión y Opinión de la ONU, David Kaye, expresó serias preocupaciones sobre el artículo 13. Y enfatizó que lo que exacerba aún más sus preocupaciones es que las tecnologías de filtrado de contenido no están equipadas para realizar interpretaciones comprensivas de las limitaciones y excepciones a los derechos de autor, como puede ser el contenido educativo, crítico, o la sátira y parodia. Ver: Palais de Nations, “Mandate of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression”, 13 de junio de 2018, p. 7, disponible en: <https://bit.ly/2KccFhh>, último acceso: 4 de noviembre de 2019.

⁴ Chappatté, Patrick, “The end of political cartoons at The New York Times”, 10 de junio de 2019, disponible en: <https://bit.ly/2q5W47Q>, último acceso: 4 de noviembre de 2019; Prieto, Ana, “El New York Times dejará de publicar caricaturas políticas, tras una acusación de antisemitismo”, *Clarín*, 11 de junio de 2019, disponible en: <https://bit.ly/2CBaLT2>, último acceso: 4 de noviembre de 2019.

música, el teatro.⁵ Definir el humor supone una tarea compleja, ya que puede variar según la época, las corrientes teóricas y los autores que lo estudian.⁶ Esto tiene que ver, en parte, con la necesaria distinción que debe hacerse de conceptos como ironía, comicidad, sátira y parodia. No obstante, todos ellos comparten ciertos rasgos semánticos y una finalidad comunicativa: generar risa.⁷ Por otro lado, este tipo de manifestaciones puede cobrar significados distintos –e incluso antagónicos– de acuerdo al sujeto que lo interpreta y el contexto en el que se inserta: lo que a algunos les causa risa, a otros puede incomodarlos. Cuándo y hasta qué punto deben tolerarse ciertas expresiones es un interrogante que parece difícil de responder.

En las últimas décadas hemos sido testigos de un enorme crecimiento del discurso humorístico, en gran parte impulsado por internet.⁸ Lo que hasta hace

⁵ Valero Heredia, Ana, "Libertad de expresión y sátira política: un estudio jurisprudencial", *Revista Internacional de Historia de la Comunicación*, N° 2, Vol. I, 2014, pp. 86-96, disponible en: <https://bit.ly/2wBK8lq>, último acceso: 4 de noviembre de 2019.

⁶ Theofylakti Zavitsanou explica que las diferentes definiciones del humor han ido insertándose en tres teorías básicas: la teoría de la superioridad, la teoría de la descarga y la teoría de la incongruencia. En la primera de ellas, la risa es vista como la expresión de un sentimiento de superioridad ante otros humanos o circunstancias considerados inferiores. Henri Bergson es considerado uno de los principales expositores de esta corriente. La teoría de la descarga se basa principalmente en el trabajo de Sigmund Freud, y se asienta sobre la idea de que el humor sirve para descargar la tensión que crean las inhibiciones y las restricciones sociales en el ser humano. Miguel Billig, Rod Martin y Ofra Nero también han sido seguidores de este enfoque. Finalmente, la teoría de la incongruencia busca definir la esencia del humor y establecer las condiciones que permiten su aparición. Según este enfoque, el humor relaciona ideas dispares entre sí de manera que "rompe las expectativas" de quien recibe el mensaje. Immanuel Kant y Arthur Schopenhauer son quienes han estudiado esta teoría principalmente. Ver, Zavitsanou, Theofylakti, "Humor y discurso político: el humor como recurso de opinión y crítica en la prensa contemporánea griega y española", Repositorio Universidad Pompeu Fabra, 2016, disponible en: <https://www.tdx.cat/handle/10803/385361>, último acceso: 9 de noviembre de 2019.

⁷ Zavitsanou, *ibid.* A los fines de este artículo, se adoptará un concepto de humor genérico, aunque con especial énfasis en la sátira y las problemáticas que suscita. Willibald Ruch, psicólogo y profesor de la Universidad de Zürich, sostiene que esta es la tendencia actual que prevalece a nivel internacional. Ver, Ruch, Willibald, "Humo(ur) Research", artículo presentado en la 14ª Conferencia de la Sociedad Internacional de Estudios de Humor, Italia, 2002.

⁸ De acuerdo a Thomas Kilian, el humor es uno de los elementos esenciales de la comunicación digital ya que satisface la mayoría o casi todas las necesidades por las que las personas utilizan internet. El autor las resume en: (i) informativa: buscar información y consejos, satisfacer un sentido de curiosidad; (ii) integración social: ganar sentido de pertenencia y vínculos sociales, conectarse con amigos, familiares y círculos sociales; (iii) identidad personal; (iv) entretenimiento: distenderse, dispersarse emocionalmente. El hecho de que los usuarios y las usuarias de internet estén constantemente en contacto con contenido humorístico –sobre todo a través de memes–, sumado al fenómeno de la

poco circulaba mayoritariamente a través de revistas y diarios ahora circula también en forma de memes generados y viralizados por los propios usuarios en redes sociales y servicios de mensajería instantánea. El término “meme” nació en 1976 en una publicación del biólogo británico Richard Dawkins, que combinó las palabras en inglés “mimesis” y “gen” y entendió a los memes como unidades de cultura semejantes a los genes.⁹ Actualmente, las acepciones que tiene la palabra lejos están de ser unívocas.¹⁰ Hay quienes sostienen que la principal característica del meme es su viralidad. Cuando se trata de contenido sujeto a derechos de autor, otros agregan que, para considerarse meme, el contenido ha de contar con una modificación sustancial respecto del original. Esta modificación algunos la interpretan como literal (ejemplo, fotomontaje), otros admiten que puede referirse al significado o interpretación que la obra adquiere en un nuevo contexto o formato (ejemplo, el caso de “Pepe the frog”, al que nos referiremos más adelante).¹¹

viralización masiva, los y las ha convertido en verdaderos productores de contenido humorístico. Ver, Kilian, Thomas, Hennigs, Nadine y Langner, Sascha, “Do millennials read books or blogs? Introducing a media usage typology of the internet generation”, *Journal of Consumer Marketing*, Emerald Group Publishing Limited, N° 2, Vol. 29, 2002, pp. 114-124, disponible en: <https://bit.ly/2XYh6nT>, último acceso: 4 de noviembre de 2019. También se ha sostenido que la proliferación del discurso humorístico en internet tiene que ver con que una gran parte de la interacción social de las personas sucede en las redes sociales. Ver, Krotoski, Alex, “What effect has the internet had on comedy?”, *The Guardian*, 3 de abril de 2011, disponible en: <https://bit.ly/2JDws8j>, último acceso: 4 de noviembre de 2019.

⁹ Este término, que surge de la combinación de las palabras en inglés *mimesis* y *gene*, fue acuñado en 1976 por Richard Dawkins, biólogo británico, en su libro *The selfish gene* (*El gen egoísta*). Dawkins sostenía que los “memes” son unidades de información que se encuentran en el cerebro y tienen la capacidad de propagarse por sí mismas y controlar el comportamiento humano. El meme fue entendido, entonces, como una unidad de cultura semejante a un gen. Quienes adscriben a esta teoría consideran que la memética es una aproximación a los modelos evolutivos de transferencia de información cultural. Ver, Nooney, Laine y Portwood-Stacer, Laura, “One does not simply: an introduction to the special issue on internet memes”, *Journal of Visual Cultures*, N° 3, Vol. 13, 16 de diciembre de 2014, pp. 248-252, disponible en: <https://bit.ly/2XKJOES>, último acceso: 4 de noviembre de 2019; MDX Online, “La cultura del meme”, 3 de diciembre de 2018, disponible en: <https://bit.ly/2ZSZqr3>, último acceso: 4 de noviembre de 2019; Dean, Jonathan, “Sorted for memes and gifs: visual media and everyday digital politics”, *Political Studies Review*, 25 de octubre de 2018, disponible en: <https://bit.ly/2Gd2qY1>, último acceso: 4 de noviembre de 2019.

¹⁰ Lantagne, Stacey, “Famous on the Internet: the spectrum of internet memes and the legal challenge of evolving methods of communication”, *University of Richmond Law Review*, 27 de noviembre de 2017, pp. 387-424, disponible en: <https://bit.ly/36UXxyv>, último acceso: 9 de noviembre de 2019. Ver también, Jackson Johnson, Shontavia, “Memetic theory, trademarks & the viral meme mark”, *The John Marshall Review of Intellectual Property Law*, N° 13, Vol. 96, 2013, pp. 104, 106, disponible en: <https://bit.ly/2Q7IVWH>, último acceso: 4 de noviembre de 2019.

¹¹ Lantagne, *op. cit.* La autora señala dos ejemplos concretos propios de la cultura

Hace ya varios años que el universo de estudio sobre las singularidades que poseen los memes y las dinámicas en las que funcionan se ha extendido al ámbito lingüístico y sociológico, lo que se conoce actualmente como la “cultura del meme”.¹² Sin embargo, en el campo del derecho los análisis en esta temática aún son limitados. El explosivo incremento de los memes en internet y su inédita naturaleza discursiva han planteado zonas grises en el derecho. Por un lado, la falta de claridad en su definición torna el concepto omnicomprendido. Pero, además, por la enorme variedad de matices, los institutos del derecho que tradicionalmente podrían aplicarse a algunos de ellos parecen difíciles de adaptar a las peculiaridades de otros. Como destaca la profesora Stacey Lantagne, la teoría del uso justo o *fair use* de derechos de autor, por ejemplo, parecería ser más fácil de aplicar a un meme “estático” —entendido como aquel en el que la imagen no sufrió cambios en su fisonomía ni en su significado al reproducirse— que en un meme “dinámico” cuyo aspecto o significado han evolucionado con su adopción como meme.¹³ Sobre este último punto, hay autores que identifican diferencias entre “usos estáticos” y “usos mutantes”, lo que incide directamente en los argumentos legales que puedan hacerse en su defensa.¹⁴

estadounidense; ver también, Carpenter, Julia, “Meme librarian is a real job, and it’s the best one on the internet”, *The Washington Post*, 21 de diciembre de 2015, disponible en: <https://wapo.st/2O0thK6>, último acceso: 4 de noviembre de 2019; Shifman, Limor, *Memes in the digital culture*, Cambridge, MA, The MIT Press, 2013; Madison, “What makes a meme go viral?”, Medium, 22 de enero de 2018, disponible en: <https://bit.ly/2CxzGHd>, último acceso: 4 de noviembre de 2019.

¹² Zittrain, Jonathan L., “Reflections on internet culture”, *Journals of Visual Culture*, N° 3, Vol. 13, 2014, pp. 388-394, disponible en: <https://bit.ly/2NSqeGM>, último acceso: 4 de noviembre de 2019; Daneci, Marsel, *Understanding media semiotics*, Londres, Bloomsbury Academic, 2° ed., 2018; Wiggins, Bradley E., *The discursive power of memes in digital culture: ideology, semiotics, and intertextuality*, Abingdon, Routledge, 2019.

¹³ Lantagne, *op. cit.*

¹⁴ Ver Lantagne, *ibid.* La autora señala que los “usos estáticos” son aquellos en los que se realiza una mera reproducción de una imagen, sin alterarla ni dotarla de nuevos significados. Si bien puede haberse modificado ligeramente la imagen, en esencia no hay ningún tipo de valor agregado a la original, simplemente se la reproduce con su sentido y contexto original. En contraste, los “usos mutantes” son aquellos que transforman drásticamente la imagen original para dotarla de significado propio. El meme mutante es toda aquella nueva creación imbuida de un nuevo significado que excede al de su creador original, impulsado además por la creatividad colaborativa de diversas comunidades de internet que añaden su propio sello.

III. El humor en tiempos de redes sociales: viejas y nuevas controversias

En general, las expresiones humorísticas a nivel comparado siempre han gozado de altos niveles de protección legal ya que juegan un papel crucial en el control de los abusos de poder y en la formación de la opinión pública. Sin embargo, también se han impuesto ciertos límites, que están permanentemente en disputa, particularmente los derechos a la imagen, reputación y honor, los derechos de autor, o la apología del delito en algunos países. A estas restricciones, que podríamos llamar tradicionales, recientemente se han sumado otras como la desinformación y los discursos discriminatorios y políticamente incorrectos, principalmente motivadas en la presunción de que su amplia propagación en internet podría tener efectos nocivos a nivel social.

En este contexto, han surgido normas que, con la intención de buscar soluciones a las problemáticas del mundo digital, comenzaron a disminuir el alcance de las expresiones humorísticas. Estas varían no solo por su carácter, sino también por los objetivos a los que responden. A continuación abordaremos algunas de las medidas más relevantes que tanto Estados como empresas privadas (plataformas) están tomando y el impacto que pueden tener para el derecho a la libertad de expresión.

1. Protección de la imagen y el honor

Los derechos al honor y la reputación han estado históricamente en tensión con la libertad de expresión y la libertad de prensa, particularmente con el humor, y especialmente el humor político. La jurisprudencia comparada en esta materia es rica y extensa. En Estados Unidos, por ejemplo, la sátira y la parodia gozan de extensa protección judicial. Uno de los casos pioneros sobre sátira fue “*Hustler Magazine, Inc. vs. Falwell*”¹⁵, en donde la Corte decidió que la revista *Hustler* no era responsable por parodiar figuras públicas. El caso se originó cuando la revista publicó una parodia que imitaba una campaña de Campari bajo el título “Jerry Falwell relata su primera vez”. La parodia, parte de una serie, constaba de una falsa entrevista en la que Falwell, un religioso y reconocido comentarista político, relataba que su “primera vez” había sido un encuentro incestuoso con su madre en estado de ebriedad. Al pie de página y en el índice de la revista se advertía que era una parodia.

¹⁵ Corte Suprema de Estados Unidos, “*Hustler Magazine, Inc. vs. Falwell*”, 485 U.S. 46, 1988, disponible en: <https://bit.ly/2CxVMcP>, último acceso: 4 de noviembre de 2019.



En su fallo, el Máximo Tribunal argumentó que el caso no se refería a afirmaciones ciertas o falsas, sino a un supuesto daño como resultado de la publicación de una caricatura cuya misión era claramente satirizar o distorsionar la realidad. Sostuvo que la Constitución de Estados Unidos prohíbe que las figuras públicas y los funcionarios públicos reciban indemnizaciones por daño psicológico intencional por una caricatura si no puede demostrarse que la publicación contiene una afirmación falsa que haya sido realizada con real malicia (*actual malice*), y que incluso este elemento en forma individual no es suficiente para atribuir responsabilidad, ya que “las representaciones gráficas y las caricaturas satíricas han jugado un rol prominente en el debate público y político (...). Desde un punto de vista histórico, es claro que nuestro discurso político habría sido considerablemente más pobre sin ellas”.¹⁶ Este caso, como en general los casos estadounidenses sobre la temática, centran su análisis en la distinción entre personas públicas y privadas, y le atribuyen a las primeras un umbral de protección menor que a las personas privadas. Indudablemente, este enfoque está en línea con la acepción de la libertad de expresión estadounidense. Esta distinción también la siguen muchos otros tribunales en casos vinculados con afectaciones a la imagen y el honor por expresiones humorísticas, incluido muchos de los tribunales en América Latina.

En Europa, el caso de *Charlie Hebdo* fue probablemente el que acaparó mayor atención en los últimos años. En 2006, el semanario francés había

¹⁶ *Ibid.*

lanzando un número especial que mostraba en su tapa una caricatura de Mahoma que se llevaba las manos a la cabeza y decía: “Es duro ser amado por idiotas”, bajo el título “Mahoma, desbordado por los integristas”. Además, la revista reproducía en sus páginas interiores los dibujos del diario danés *Jyllands-Posten*,¹⁷ incluido uno de los más controvertidos en el que aparecía Mahoma con un turbante con forma de bomba. En pocas horas, se vendieron más de 400.000 ejemplares, lo que triplicó las ventas habituales. “Este éxito prueba el interés que la gente tiene por su propia libertad. Es una respuesta ciudadana”, manifestó en esa ocasión el entonces director de la revista.¹⁸



La Unión de Organizaciones Islámicas de Francia y la Gran Mezquita de París denunciaron tres caricaturas: la que aparecía en tapa y otras dos que habían sido reproducidas por la revista pero eran originales del diario

¹⁷ Las caricaturas habían sido originalmente publicadas en 2005 por el diario danés *Jyllands-Posten* y provocaron violentas protestas alrededor de Asia, África y el Medio Oriente que dejaron al menos cincuenta personas muertas. Muchos medios europeos republicaron esas caricaturas como una manera de reafirmar y pedir respeto por la libertad de expresión. Ver, De Andrés, Francisco, “La ruta mortal de las caricaturas”, *ABC Internacional*, 8 de enero de 2015, disponible en: <https://bit.ly/2KJx7V5>, último acceso: 4 de noviembre de 2019.

¹⁸ “Chirac tacha de ‘provocación’ el número especial con caricaturas de Mahoma de una revista francesa”, *El Mundo*, 9 de febrero de 2006, disponible en: <https://bit.ly/32Db2PS>, último acceso: 4 de noviembre de 2019.

Jyllands-Posten.¹⁹ Las dos entidades musulmanas pidieron que se reconociera el delito de injurias por motivos religiosos, con penas de hasta seis meses de cárcel y multa.

En 2007, el Tribunal Correccional de París absolvió al director de *Charlie Hebdo*, entre otros motivos, por tratarse de una publicación de naturaleza eminentemente satírica.²⁰ Señaló que nadie tiene la obligación de comprar o leer ese tipo de revistas, lo que lo diferencia de otros medios, como pueden ser los carteles de la vía pública. Enfatizó además que el fin último de la protección a la libertad de expresión incluye también aquellos discursos que puedan ofender o chocar a algunas personas. En cuanto a las caricaturas en sí mismas, adoptó la postura clásica sobre el interés público en juego, y sostuvo que no constituían difamación ni un ataque personal y directo contra un grupo de personas por motivos religiosos. Resulta interesante el énfasis de la sentencia en cuanto al contexto en el que deben analizarse las caricaturas, que en este caso se trataba de un número especial dedicado al fundamentalismo religioso. En general, el fallo no supuso ninguna sorpresa ya que la propia fiscalía había pedido la absolución por considerar que las caricaturas no atacaban al islam sino a los fundamentalistas.²¹ Algunos analistas repararon en la extensa tradición de la postura de los tribunales franceses con respecto a la separación de la Iglesia y el Estado, al igual que sus repetidos pronunciamientos a favor de la libertad de expresión ante los reclamos fundados en el respeto a la religión.²²

Este caso constituye uno de los últimos hitos mundiales en materia de sátira y ha calado hondo en la jurisprudencia y la doctrina en el mundo, en

¹⁹ En una de ellas se ve a Mahoma con un turbante del cual sale la mecha de una bomba, en la otra se ve al profeta que está pidiendo a los terroristas que “no se inmolen porque ya no quedan más vírgenes en el paraíso”. Esto tiene que ver con la doctrina más tradicional del islam que señala a sus fieles que si se inmolan y dan la vida por Alá recibirán en el paraíso 72 mujeres vírgenes a modo de bendición especial. Las mujeres, en cambio, recibirán un solo hombre “con el que estarán satisfechas”.

²⁰ La síntesis de los argumentos más importantes fue extraída de Noorlander, Peter, “When satire incites hatred: Charlie Hebdo and the freedom of expression debate”, Center for Media, Data and Society, 2 de febrero de 2015, disponible en: <https://bit.ly/2IPo3fv>, último acceso: 4 de noviembre de 2019; y Furlon, Armelle, “The Charlie Hebdo case: freedom of expression and respect for religious faith”, International Law Office (ILO), 26 de junio de 2008, disponible en: <https://bit.ly/2z5Yoj2>, último acceso: 4 de noviembre de 2019.

²¹ Marti Font, José María, “Un tribunal francés absuelve a la revista que publicó las caricaturas de Mahoma”, *El País*, 23 de marzo de 2007, disponible en: <https://bit.ly/2MFuvYP>, último acceso: 4 de noviembre de 2019.

²² Leveque, Thierry, “French Court clears weekly in Mohammad cartoon row”, *Reuters*, 22 de marzo de 2007, disponible en <https://reut.rs/2IS6ZFA>, último acceso: 4 de noviembre de 2019.

gran medida debido al atentado que sufrió la redacción de *Charlie Hebdo* años más tarde (un tema que ha sido largamente explorado y excede el marco de este análisis). Las reconocidas “viñetas de Mahoma” luego se replicaron en otros medios a nivel mundial,²³ algunos de los cuales tuvieron consecuencias judiciales.²⁴

El Tribunal Europeo de Derechos Humanos se expidió sobre este tema en el caso “Vereinigung Bildender Künstler c. Austria” en 2007.²⁵ Este caso se originó en la exposición titulada “El siglo de la libertad artística” en una reconocida galería independiente en Viena. Una de las pinturas exhibidas constaba de un collage de figuras públicas inmersas en actividades sexuales, entre ellas la Madre Teresa de Calcuta, un cardenal austríaco y reconocidos representantes del Partido Liberal austríaco. El cuadro ocasionó controversias e incidentes, e incluso llegó a ser dañado por un visitante. El ex secretario general del Partido Liberal solicitó la prohibición de la exhibición del cuadro como medida cautelar en base a la Ley de Copyright de Austria, que extiende este derecho no solo al fotógrafo sino a la persona retratada, y una indemnización. Sin perjuicio del encuadre legal, los tribunales nacionales resolvieron el caso a la luz de los derechos personalísimos del demandante y sostuvieron que la obra constituía una acción difamatoria. El tribunal europeo señaló que “es de común comprensión de los tribunales nacionales en todas sus instancias que las pinturas en cuestión no tenían como fin reflejar o sugerir una realidad” y, por ello, “se trataba de caricaturas que utilizaban elementos satíricos”²⁶. Describió a la sátira como una forma de expresión artística y crítica social que por sus características inherentes de exageración y distorsión de la realidad está destinada a “provocar y agitar”. Por ende, cualquier interferencia con el derecho de un artista a tal expresión debe ser examinada bajo estricto y especial cuidado.

En el plano latinoamericano, la Corte Interamericana de Derechos Hu-

²³ En Estados Unidos, por ejemplo, algunas de estas caricaturas aparecieron en el diario *Philadelphia Inquirer* y en *Philadelphia Jewish Voice*, además de ser exhibidas en programas de TV y en boletines universitarios como el de la Universidad de Harvard. Ver, Freedman, Leonard, *The offensive art: political satire and its censorship around the world from Beerbohm to Borat*, Santa Barbara, CA, Praeger, 2008, p. 53 y 54.

²⁴ Ver, Mourenza, Andrés, “Cárcel para dos periodistas turcos por reproducir una caricatura de Mahoma de ‘Charlie Hebdo’”, *El País*, 28 de abril de 2016, disponible en: <https://bit.ly/2KFgGsN>, último acceso: 4 de noviembre de 2019.

²⁵ Tribunal Europeo de Derechos Humanos, disponible en: <https://bit.ly/33Bdd80>, último acceso: 4 de noviembre de 2019.

²⁶ *Ibid.*

manos no tiene casos vinculados a la temática de humor. Sin embargo, el caso “La última tentación de Cristo contra Chile” evalúa el derecho al honor de un grupo frente a una expresión artística, y cuenta con componentes interesantes y válidos para analizar también el discurso humorístico.²⁷ El caso en cuestión se generó a raíz del recurso de protección interpuesto por un grupo de personas en nombre y representación de Jesucristo que solicitaron al Consejo de Calificación Cinematográfica en Chile²⁸ que prohibiera la difusión de la película *La última tentación de Cristo* por considerarla contraria al honor y la reputación de Jesucristo y contraria a la libertad religiosa. El Consejo aceptó la solicitud y prohibió la difusión de la película, y la Corte Suprema de Chile avaló tal prohibición. La Corte Interamericana en su sentencia entendió que la libertad de expresión es la piedra angular de la sociedad democrática y que ampara aquellas expresiones que resulten chocantes. Concluyó que en el caso hubo censura previa incompatible con la Convención Americana y ordenó la reforma de la Constitución chilena. Además de este caso, la Corte Interamericana tiene varios casos vinculados al honor y la reputación como posibles límites a la libertad de expresión donde adoptó la doctrina que distingue entre figuras públicas y privadas e información de interés público, según la línea de las principales Cortes Supremas de América Latina.²⁹

Argentina y Colombia tienen casos interesantes en materia de humor y libertad de expresión, incluso recientes. El caso de Cecilia Pando (2017) en Argentina surgió a raíz de un fotomontaje publicado por la revista *Barcelona* en su contratapa en la que se veía un cuerpo desnudo atado con sogas con el rostro de Pando, activista por los derechos de los condenados por crímenes de lesa humanidad en Argentina. Pando inició una demanda civil por daños y perjuicios pues consideró que habían lesionado su honor e imagen.

La primera instancia fue favorable a la actora y en 2017 la Cámara Nacional de Apelaciones en lo Civil confirmó la sentencia.³⁰ El tribunal entendió

²⁷ Corte IDH, caso “La última tentación de Cristo (Olmedo Bustos y otros) c. Chile”, sentencia del 5 de febrero de 2001, Serie C, Nº 73.

²⁸ Órgano encargado constitucionalmente de la evaluación y aprobación o reprobación de difusión de cinematografía en Chile.

²⁹ Corte IDH, “Herrera Ulloa vs. Costa Rica”, sentencia del 2 de julio de 2004, Excepciones Preliminares, Fondo, Reparaciones y Costas; Corte IDH, “Ricardo Canese vs. Paraguay”, sentencia del 31 de agosto de 2004, Fondo, Reparaciones y Costas; Corte IDH, “Kimel vs. Argentina”, sentencia del 2 de mayo de 2008; entre otros.

³⁰ Cámara Nacional de Apelaciones en lo Civil, Sala D, “Pando de Mercado, María Cecilia c/ Gente Grossa SRL s/daños y perjuicios”, 23 de marzo de 2017, disponible en: <https://bit.ly/2IN0H0Z>, último acceso: 4 de noviembre de 2019. La sentencia ordenaba

que la doctrina de la real malicia no resultaba procedente ya que no se trataba de una noticia publicada en forma falsa o inexacta, sino que parodiaba a Pando a través de una imagen modificada y con títulos que excedían los límites de la prensa. Afirmó que se trataba de un ejercicio imprudente de la libertad de expresión, al punto de afectar derechos personalísimos.

La sentencia fue apelada y se encuentra pendiente de resolución. En 2018, el procurador fiscal ante la Corte Suprema de Justicia de la Nación emitió su dictamen sobre el caso y entendió que la Corte debería revertir la decisión de la Cámara.³¹ Señaló que se trata de una sátira sobre un tema de interés público y referida a una persona pública, por lo que su margen de tolerancia frente a la crítica debería ser mayor. Recalcó que la sátira social o política se encuentra especialmente tutelada por la libertad de expresión pues enriquece el debate público, y que en Argentina tiene una larga tradición.

En Colombia un caso que trascendió fue el del caricaturista César González, más conocido como “Matador”, con su caricatura del presidente Iván Duque en el diario *El Tiempo*. En el dibujo aparece el entonces precandidato presidencial convertido en un cerdo, con una viñeta de texto que dice: “¡¡Ay No!! Soy el único uribista que no está ‘cochino’”.



a la revista a pagar la suma de 70 mil pesos argentinos en concepto de indemnización.

³¹ Dictamen del procurador fiscal ante la Corte Suprema, “Pando c/ Gente Grossa SRL s/ daños y perjuicios”, 20 de febrero de 2018, disponible en: <https://bit.ly/2z5Frqn>, último acceso: 4 de noviembre de 2019.

Un abogado colombiano simpatizante del partido Centro Democrático interpuso una acción de tutela contra la Editorial El Tiempo y el caricaturista, alegando una ofensa y una amenaza de daño a la imagen de los candidatos del partido político. Señaló que la caricatura estaba destinada a influir “de manera arbitraria, y con el pretexto de realizar una manifestación humorística o jocosa, en la libertad de conciencia de los colombianos y en su derecho a ejercer su facultad de elegir sin interferencias, manipulaciones o presiones de los medios de comunicación”.³² La jueza que conoció en el caso negó la tutela.³³ Destacó que lo que se pretendía era “un modo de censura” y que “no se puede calificar de violatorio de derechos fundamentales cuando se publica una caricatura de un candidato en un periódico, pues calificarlo, o limitarlo, o disminuir su potencial crítico, sería violar el derecho de los lectores al pluralismo de ideas, de opiniones, y sería contrario al estado social de derecho y por ende a la democracia”.³⁴ Mencionó, además, que la creatividad y el ingenio del caricaturista no pueden verse limitados por quien ve la caricatura y simplemente no le gusta o no comparte la opinión, que la caricatura solo tuvo la intención de realizar una opinión libre respecto de un tema de actualidad en un contexto de elecciones presidenciales inminentes, de las que se pueden desprender críticas favorables o desfavorables sin que ello constituya una vulneración al honor.

En Ecuador existen múltiples casos de caricaturas dirigidas al expresidente Rafael Correa, quien ha considerado públicamente a este tipo de críticas como “parte de una campaña de desprestigio sistemático con ironía”.³⁵ Una

³² Se solicitó al juez que se ordene un pedido de disculpas público en el medio de comunicación, para proteger “los derechos fundamentales al buen nombre, libertad de conciencia y la libertad de elegir de los militantes del partido Centro Democrático (llamados ‘uribistas’) y de los electores colombianos en general”. Ver, Redacción judicial, “Tutela contra el caricaturista Matador: libertad de expresión en época electoral”, *El Espectador*, 8 de marzo de 2018, disponible en: <https://bit.ly/2Hk4ADk>, último acceso: 4 de noviembre de 2019.

³³ Algunas organizaciones de la sociedad civil como la Fundación para la Libertad de Prensa (FLIP) y el Centro de Estudios de Derecho, Justicia y Sociedad (Dejusticia) presentaron escritos como coadyuvantes en el proceso de acción de tutela, y argumentaron en defensa del caricaturista. Ver, escrito de la FLIP, 10 de marzo de 2018, disponible en: <https://bit.ly/2GANv9f>, último acceso: 4 de noviembre de 2019; y de Dejusticia, 12 de marzo de 2018, disponible en: <https://bit.ly/2SitNnM>, último acceso: 4 de noviembre de 2019.

³⁴ Juzgado 4 Civil Circuito de Ejecución de Sentencias, Bogotá, 20 de marzo de 2018. Tutela instaurada por José Luis Reyes Villamizar contra Julio César Quiceno “Matador” y Casa Editorial El Tiempo-CEET S.A.

³⁵ “Suspenden por varias horas cuenta de Twitter de página crítica de Correa”, *Panam Post*, 29 de enero de 2015, disponible en: <https://bit.ly/2rrJSyD>, último acceso: 11 de noviembre de 2019; Fundamedios, “Fundamedios condena censura de Twitter tras suspensión de cuenta de Crudo en Ecuador”, 11 de agosto de 2017, disponible en:

de las historias que sobresale es la de “Bonil”, un caricaturista del diario El Universo. En 2013, Bonil publicó una viñeta sobre el allanamiento a la casa de Fernando Villavicencio, quien había sido acusado por Correa de hackear sus correos electrónicos, los del vicepresidente Jorge Glas y los del exsecretario jurídico de la Presidencia, Alexis Mera. Las gráficas muestran personajes armados y uniformados y el mensaje “Fiscalía y Policía allanan la casa de Fernando Villavicencio y se llevan denuncias de corrupción”. A raíz de la publicación, la Superintendencia de la Información y Comunicación (Supercom) sancionó a Bonilla y a Diario El Universo. Al caricaturista le dio un plazo de 72 hs. para rectificarse, mientras que al periódico le aplicó una multa equivalente al 2% de la facturación promedio de los últimos tres meses.³⁶ Casi cuatro años después, el Tribunal Distrital número 2 de lo Contencioso Administrativo de Ecuador declaró nula la resolución emitida por la Supercom. Sostuvo que “lo que se ha sancionado es un contenido de opinión” y agregó que “la caricatura deja entrever una posición crítica respecto de una fuente previa (...)”. La sentencia sostiene que en tanto juicio de valor, la caricatura no es pasible de veracidad y no es ni puede ser comprobable o jurídicamente reproachable.³⁷

Indudablemente la tensión entre el honor y la reputación con la libertad de expresión es un tema que, a pesar de estar abordado en la jurisprudencia comparada, sigue estando en la agenda de los tribunales. Interesantemente, los casos permiten vislumbrar la diversidad de acepciones que el honor y la reputación tienen. En algunos casos se lo destaca como derechos personalísimos, en otros se alega la violación de derechos colectivos, como el honor de un grupo (sea religioso o de otro tenor).

Si bien muchos de los ejemplos citados responden a publicaciones o expresiones analógicas, la misma línea siguen los casos que abordan expresiones digitales. Cabe citar, por ejemplo, el caso de Ali Ziggi Mossilmani de Australia, radicado en el Tribunal de Primera Instancia de Sydney.³⁸ Ziggi, como se hace

<https://bit.ly/2q4bNnV>, último acceso: 11 de noviembre de 2019.

³⁶ Ver artículo 25 de la Ley Orgánica de Comunicación de Ecuador. Por esta norma el medio se ve obligado a tomar posición institucional respecto a la inocencia o culpabilidad de Xavier Bonilla.

³⁷ “Jueces declaran nula actuación de Supercom sobre caricature de Bonil”, *El Universo*, 23 de agosto de 2017, disponible en: <https://bit.ly/2Q8PxUP>, último acceso: 11 de noviembre de 2019.

³⁸ Whitbourn, Michaela, “Mullet defamation case gets a haircut as Sydney teen Ali Ziggi Mossilmani suffers setback”, *The Sydney Morning Herald*, 18 de diciembre de 2016, disponible en: <https://bit.ly/2qEo2Yz>, último acceso: 4 de noviembre de 2019. Ver fallo, “District Court New South Wales”, 2016, disponible en: <https://bit.ly/36TITbZ>, último

llamar, es un adolescente de 18 años que adquirió fama mundial en internet en 2016 por una foto suya que se volvió viral en la cual luce un corte de pelo extravagante. La foto fue citada y circulada con diferentes leyendas que hacían alusión al corte de pelo y distintos juegos de palabras. El autor decidió demandar a tres medios australianos por la inclusión de la foto en sus ediciones tanto en papel como en internet. La jueza del tribunal desestimó la demanda por las publicaciones en internet y solo mantuvo la demanda por la publicación en papel de uno de los medios. En su razonamiento argumentó que “tratar de inferir una caracterización de estupidez de alguno de estos artículos, cuando claramente son adaptaciones humorísticas del corte de pelo del demandante sin ningún significado ulterior, es un ejercicio fútil”.³⁹

En Argentina, el senador y excandidato a vicepresidente Miguel Ángel Pichetto denunció en 2018 ante la Justicia Federal haber sufrido un “perjuicio personal, moral y político” por parte de una cuenta de Twitter que, a modo de parodia, utilizaba su imagen y su nombre. La cuenta, que se presentaba como “no oficial” en la biografía, publicaba supuestas frases del legislador que se replicaron rápidamente como chistes entre los usuarios. Si bien la acción judicial no se enmarcó en la normativa local de calumnias e injurias y, por el contrario, lo hizo en base a la “sustitución de identidad”, se trata de una forma más de ataque y restricción a las expresiones humorísticas en internet en aras de proteger la imagen y la honra de un funcionario público. El senador Pichetto argumentó, además, que a la afectación individual se sumaba “el engaño (...) que en su buena fe sufre cada uno de sus seguidores”. En julio de 2019, el juez decidió sobreseer al creador de la cuenta falsa ya que la sustitución de identidad en redes sociales no se encuentra tipificada en el Código Penal argentino.⁴⁰ No obstante, el tema continúa en la agenda pública pues derivó la presentación de diversos proyectos de ley para tipificar penalmente la “usurpación” o “suplantación” de la identidad digital, fundados esencialmente en que el uso de una identidad ajena en redes sociales podría ser la antesala de delitos como la estafa, las “calumnias e injurias” o el abuso de menores.⁴¹

acceso: 4 de noviembre de 2019.

³⁹ En su redacción original: “To attempt to draw an imputation of stupidity from one or more of these [articles] when they are all humorous adaptations of the plaintiff’s hairstyles with no deeper meaning is an exercise in futility”. Withbourn, Michaela, *op.cit.*

⁴⁰ Poder Judicial de la Nación, Juzgado Criminal y Correccional Federal 4, 27 de marzo de 2018, disponible en: <https://bit.ly/2Czo0DP>, último acceso: 4 de noviembre de 2019.

⁴¹ Ver, CELE, “Argentina Proyecto de Ley Usurpación Digital (2449/18)”, 2018, disponible en: <https://bit.ly/36TGlnl>, último acceso: 4 de noviembre de 2019.

Como se ve en los casos mencionados, el humor goza de amplia protección frente a la acción por difamación o injurias. A diferencia de otros temas, existe un consenso en torno a que la protección y el balance en estos casos debe necesariamente ser judicial. Así, en la mayoría de los países de América Latina y Estados Unidos está descartada la responsabilidad objetiva de las plataformas por contenido difamatorio. Y los términos y las condiciones de servicios de las grandes plataformas de contenido no prevén medidas de autorregulación por difamación⁴², a pesar de que no son pocas las insistencias de personas y funcionarios públicos para que las incluyan.

2. Protección de los derechos de autor

La utilización de textos, imágenes, música y, en general, contenidos audiovisuales sujetos a derechos de autor sin la debida autorización constituye otro de los motivos tradicionales por los que el humor ha entrado en tensión con la libertad de expresión. Estos marcos regulatorios le conceden al titular de creaciones literarias, musicales, científicas o artísticas una serie de facultades o prerrogativas para fomentar la creación artística y proteger a los autores en sus derechos patrimoniales sobre las obras y sus reproducciones. Sin perjuicio de ello, la legislación a nivel comparado establece limitaciones a los derechos individuales de los titulares de las obras a fin de armonizarlos con el derecho de acceso al conocimiento y a la información, los usos educativos, e incluso la parodia y la sátira, entre otros intereses primordialmente colectivos.⁴³

En numerosas legislaciones del mundo, la sátira y la parodia están contempladas como excepciones a la exclusividad de uso que propone el derecho de autor. Al tener un valor intrínsecamente educativo y cultural, en general, se considera que las obras objeto de la sátira y la parodia pueden reproducirse sin licencias o autorizaciones. Esta excepción encuentra distintos nombres, como la denominada doctrina de uso justo o *fair use*.

La distinción entre la sátira o la parodia y los usos abusivos de material sujeto a derechos de autor no siempre es clara. Muy por el contrario, los límites

⁴² Por ejemplo los términos de servicio de plataformas como Facebook, YouTube o Twitter.

⁴³ En Colombia, por ejemplo, la Ley 23 sobre Derechos de Autor (1982) contempla una serie de excepciones y limitaciones que permiten el uso del material protegido sin autorización en casos como la cita textual (art. 31), la ilustración en obras destinadas a la enseñanza (art. 32) o la toma de apuntes de conferencias o lecciones en establecimientos educativos (art. 39). También, en Costa Rica, la ley N° 6.683 (1982), luego modificada por la ley N° 8.686 (2008) y en México en la Ley Federal de Derechos de Autor se prevén excepciones de uso justo, solo por citar algunos ejemplos.

están permanentemente en disputa. A diferencia de los casos de difamación, en estos casos las empresas de internet han tenido mayor protagonismo que los tribunales. Existen casos documentados de usos y abusos de la política de protección de derechos de autor para limitar la sátira, la parodia y la crítica humorística en general.⁴⁴ Eduardo Bertoni y Sophia Sadinsky mencionan el caso de la oficina de Turismo de Alberta, en Canadá, que emitió una notificación para que YouTube cancelara un video satírico que utilizaba parte de un comercial para criticar la degradación ambiental en la ciudad. También en Canadá, el Correo nacional solicitó a YouTube que removiera un video difundido por miembros de un sindicato en el que se burlaban de su directora ejecutiva y sus políticas corporativas, al mostrar una fotografía retocada de ella.⁴⁵

A nivel mundial, existen países que no contaban con la excepción de uso justo y la incorporaron recientemente. En Nueva Zelanda, por ejemplo, se presentó en 2017 un proyecto ante el Parlamento para ampliar las excepciones existentes a las infracciones de derechos de autor e incluir la sátira y la parodia en casos de críticas artísticas y *reviews*.⁴⁶ En Australia también se abrió este debate, y se incluyeron finalmente en 2006 las excepciones de uso justo.⁴⁷

En el plano latinoamericano, México modificó en 2018 su ley federal e implementó medidas cautelares en casos de utilización de material sujeto a derechos de autor sin el consentimiento del titular. Por su amplitud y vaguedad, la ley deja abierta la posibilidad de que los memes queden sujetos a los derechos de autor. La norma fue aprobada a pesar de haber sido severamente criticada por organizaciones dedicadas a la promoción de los derechos humanos en el entorno digital.⁴⁸

Desde 2012 y hasta la actualidad, Colombia también ha intentado modificar y actualizar el régimen de derechos de autor pero hasta el momento no ha teni-

⁴⁴ Bertoni, Eduardo y Sadinsky, Sophia, "El uso de la DMCA para limitar la libertad de expresión", *Internet y Derechos Humanos II*, Buenos Aires, CELE, Universidad de Palermo, 2016. Ver también, Keller, Daphne, "Empirical evidence of 'over-removal' by Internet companies under intermediary liability laws", Center for Internet and Society, Stanford University, 12 de octubre de 2015, disponible en: <https://stanford.io/2fBMNhk>, último acceso: 4 de noviembre de 2019.

⁴⁵ Bertoni y Sadinsky, *op. cit.*, p. 63.

⁴⁶ Parlamento de Nueva Zelanda, "Copyright (Parody and Satire) Amendment Bill", disponible en: <https://bit.ly/2JAQWQH>, último acceso: 4 de noviembre de 2019. En particular, ver la sección 42.

⁴⁷ Austin, Graeme, "A copyright exemption for parody and satire", *Newsroom*, 11 de mayo de 2018, disponible en: <https://bit.ly/2GcrAWG>, último acceso: 4 de noviembre de 2019.

⁴⁸ Sinembargo.MX, "Nueva ley censura hasta a los memes", *Noroeste*, 30 de abril de 2018, disponible en: <https://bit.ly/2X4H46L>, último acceso: 4 de noviembre de 2019; Notimex, "Por esta razón podrían demandarte si usás memes", *Milenio*, 6 de marzo de 2019, disponible en: <https://bit.ly/2Y2Ft41>, último acceso: 4 de noviembre de 2019.

do éxito. La última iniciativa que ingresó al Congreso, y que aún se encuentra radicada en el Senado, es la conocida “Ley Lleras 6”.⁴⁹ Este proyecto incluye la excepción de parodia, una deuda pendiente según los activistas, dado que los anteriores mantenían como regla la prohibición de reproducción de obras protegidas y no establecían ninguna excepción a la crítica, la parodia o la caricatura.⁵⁰

Recientemente, la discusión sobre sátira, parodia y derechos de autor se intensificó a raíz de la propuesta y aprobación de la directiva europea sobre derechos de autor,⁵¹ y en América Latina por el impacto que la directiva podría tener en las empresas y en los usuarios en esta región.⁵² Esta norma modifica la directiva de 2001⁵³ y actualiza el marco jurídico de propiedad intelectual. Si bien aún no entra en vigencia y existe un plazo de dos años para que los distintos Estados miembros de la Unión Europea ajusten su legislación, su redacción ha generado importantes y severas críticas.⁵⁴ El artículo 17, uno de los más controvertidos, les exige a las plataformas que ofrecen servicios de publicación e intercambio de contenido en internet (como YouTube, Twitter o Facebook) que supervisen y tomen todas las medidas proactivas necesarias para evitar que los contenidos protegidos por derechos de autor sean alojados en sus servidores y compartidos por usuarios. Un amplio sector de la comunidad se opuso drásticamente a la medida alegando principalmente que dicha redacción promueve el monitoreo activo y filtrado de contenido, incluso mediante filtros de carga o *pre-upload* (en

⁴⁹ Ley N° 206 de 2018.

⁵⁰ Botero, Carolina, “Una de las Ley Lleras por fin será Ley y reformará el derecho de autor en Colombia”, Fundación Karisma, 24 de mayo de 2018, disponible en: <https://bit.ly/2qF16ts>, último acceso: 4 de noviembre de 2019.

⁵¹ Unión Europea, directiva N° 2019/790 del Parlamento Europeo y del Consejo sobre los derechos de autor y derechos afines en el mercado único digital y por la que se modifican las directivas N° 96/9/CE y N° 2001/29/CE, 17 de abril de 2019, disponible en: <https://bit.ly/30D1qEb>, último acceso: 4 de noviembre de 2019.

⁵² CELE, “La Directiva Europea de Derecho de Autor y su impacto en los usuarios de América Latina y el Caribe: una perspectiva desde las organizaciones de la sociedad civil”, abril de 2019, disponible en: <https://bit.ly/36YtNRA>, último acceso: 4 de noviembre de 2019.

⁵³ Unión Europea, directiva N° 2001/29/CE del Parlamento Europeo y del Consejo relativa a la armonización de determinados aspectos de los derechos de autor y derechos afines a los derechos de autor en la sociedad de la información, 22 de mayo de 2001, disponible en: <https://bit.ly/2SPkUm8>, último acceso: 4 de noviembre de 2019.

⁵⁴ Naciones Unidas, Oficina del Alto Comisionado para los Derechos Humanos, “EU must align copyright reform with international human rights standards, says expert”, 11 de marzo de 2019, disponible en: <https://bit.ly/2LtKlch>, último acceso: 4 de noviembre de 2019. Ver también, Lara, Juan Carlos, “Directiva de Derechos de Autor de la UE: avanza la internet filtrada en Europa”, Derechos Digitales, 14 de febrero de 2019, disponible en: <https://bit.ly/2LVp1aD>, último acceso: 4 de noviembre de 2019.

otras palabras, que filtran el contenido antes de que sea publicado).⁵⁵ Sobre los filtros, Wikipedia, uno de los actores más activos en este debate, sostuvo:

El mundo debería preocuparse por las nuevas propuestas para introducir un sistema que filtre automáticamente la información antes de que aparezca en línea. A través de filtros de carga obligatorios, las plataformas se verían obligadas a crear sistemas costosos y a menudo sesgados para revisar y filtrar automáticamente las posibles violaciones de los derechos de autor en sus sitios web.⁵⁶

El filtrado en general supone por lo menos dos problemas. Por un lado, tiene problemas de transparencia y *accountability* difíciles de subsanar. Por otro lado, por la complejidad de este tipo de sistemas, existe la posibilidad de que muy pocas empresas tengan la capacidad técnica para desarrollar esta tecnología, con lo cual estas medidas incrementarían el poderío y la concentración del sector en unos pocos y grandes actores. Trasladado a nuestra región, el filtrado previo que incentiva esta regulación sería directamente equiparable a un sistema de censura directa e incompatible con los estándares interamericanos de libertad de expresión.⁵⁷

Según la propia norma y las regulaciones en esta materia, la sátira y la parodia –incluidos aquí los memes– quedarían fuera del alcance de esta regla por la excepción de uso justo. Sin embargo, la capacidad de detectar y de distinguir este contenido de otros *a priori* no parece sencilla. Tampoco parece claro que todos los titulares de los derechos de autor quieran ejercer este derecho de esta manera en todos los casos. Pensemos, por ejemplo, en el caso de “Pepe the frog” en Estados Unidos. La caricatura fue creada en 2005 por un dibujante estadounidense, Matt Furie, y publicada inicialmente en su webcómic. Se popularizó rápidamente, adaptándose su cara a diferentes escenarios y emociones y convirtiéndose en un meme utilizable en cualquier contexto. En los

⁵⁵ Ver, European Digital Rights (Edri), “Copyright: open letter calling for the deletion of articles 11 and 13”, 29 de enero de 2019, disponible en: <https://bit.ly/2FWqlXR>, último acceso: 4 de noviembre de 2019; Doctorow, Cory, “More than 130 European businesses tell the European Parliament: reject the #CopyrightDirective”, Electronic Frontier Foundation (EFF), 20 de marzo de 2019, disponible en: <https://bit.ly/2ulSZ2k>, último acceso: 4 de noviembre de 2019.

⁵⁶ Sefidari, María, “Tu internet está bajo amenaza. Estas son las razones por las que deberías preocuparte por la Reforma Europea de Derechos de Autor”, Wikimedia, 4 de septiembre de 2018, disponible en: <https://bit.ly/2xQex8K>, último acceso: 4 de noviembre de 2019.

⁵⁷ CELE, “La Directiva Europea de Derecho de Autor y su impacto en los usuarios de América Latina y el Caribe: una perspectiva desde las organizaciones de la sociedad civil”, *op. cit.*

años que siguieron ya era viral en MySpace, Tumblr y otras redes sociales como una forma graciosa de expresar múltiples estados de ánimo. A pesar de ser una obra sujeta a derechos de autor, “Pepe the frog” circuló libremente sin que su autor iniciara acción alguna para exigir el pago de cánones de uso o el cese del mismo. Indudablemente, ya existía ahí cierta tensión con los derechos de autor en tanto el personaje era utilizado y reutilizado por usuarios de distintas latitudes. En 2016, sin embargo, la imagen del Sapo tomó un giro inesperado cuando el movimiento de la derecha alternativa en Estados Unidos se apropió del personaje e hizo de él su símbolo para promover la campaña del entonces precandidato presidencial Donald Trump, al vestirlo con túnicas del Ku Klux Klan y cascos militares. Esta polémica versión del Sapo Pepe se reprodujo en revistas y posters. Fue entonces cuando Furie, su creador, demandó al fundador de un sitio de extrema derecha que lo utilizó con fines políticos por infringir las leyes de copyright.⁵⁸

Como este último caso existen muchos otros que exponen la dificultad para detectar: 1) de qué tipo de uso se trata; y 2) qué connotaciones tienen ciertos usos a diferencia de otros, sobre todo cuando en gran parte dependen del contexto.⁵⁹ Cabe recordar además que, en el caso de los memes, la aplicación de la doctrina del uso justo varía de acuerdo a que se trate de memes de “uso estático” o memes de “uso dinámico”, mencionados inicialmente.⁶⁰ El temor por los filtros automáticos es que estos no sean capaces de identificar ni los unos ni los otros, ni puedan reconocer ironías o expresiones puramente satíricas, o distinguirlas de otro tipo de material protegido por derechos de autor, y que en definitiva sean removidos de la web.⁶¹

⁵⁸ Leon, Harmon, “Alex Jones’ Pepe the Frog Battle & the Co-Opting of Other Innocent Symbols”, *Observer*, 23 de mayo de 2019, disponible en: <https://bit.ly/2O0T6th>, último acceso: 4 de noviembre de 2019; Thomser, Jacqueline, “Judge rules Pepe the Frog copyright lawsuit against InfoWars will go to trial”, *The Hill*, 17 de mayo de 2019, disponible en: <https://bit.ly/2KbV4pM>, último acceso: 4 de noviembre de 2019.

⁵⁹ Dewey, Caitlin, “How copyright is killing your favorite memes”, *The Washington Post*, 8 de septiembre de 2015, disponible en: <https://wapo.st/33Nqgnb>, último acceso: 4 de noviembre de 2019.

⁶⁰ Lantagne, *op. cit.*, p. 395. La autora menciona que mientras que la doctrina de uso justo podría encuadrar perfectamente en los usos de memes estáticos, donde lo que se reproduce es una imitación idéntica o casi idéntica de una imagen, podría en cambio no ser fácilmente aplicable si el uso es mutante y hay una transformación drástica de la imagen original.

⁶¹ Para ver más información sobre las críticas a la directiva: Reynolds, Matt, “What is Article 13? The EU’s divisive new copyright plan explained”, *Wired*, 24 de mayo de 2019, disponible en: <https://bit.ly/2xlQdM4>, último acceso: 9 de noviembre de 2019; Schaeffer, Joe, “The EU’s Article 13 signals the death knell of political satire”, 8 de octubre de 2018, en *Liberty Nation*, disponible en: <https://bit.ly/2LmSmjm>, último acceso: 9 de noviembre de 2019.

La directiva europea constituye una medida sin precedentes, al invertir la lógica por la cual hasta ahora se regían las normas de responsabilidad de intermediarios –lógicas reactivas– y avanzar en cambio hacia una lógica proactiva. Los sistemas vigentes hasta ahora siguen un modelo parecido en más o en menos al sistema estadounidense de la Ley de Derechos de Autor del Milenio Digital (Digital Millennium Copyright Act o DMCA) que prevé un sistema de “puerto seguro”. En este tipo de sistemas, el intermediario está en principio exento de responsabilidad por el contenido de terceros hasta tanto se le haga saber a través de una notificación que el material existe y viola derechos de autor de un tercero. Existen diversos estudios que dan cuenta de los abusos de estos sistemas, dado que no requieren orden judicial para la remoción de contenidos y por su propia arquitectura generan incentivos para que la misma plataforma bloquee contenidos y enlaces frente a la mera notificación.⁶² Sin embargo, rige el principio de inmunidad de las plataformas frente al contenido de terceros hasta tanto haya una notificación de la parte supuestamente agraviada. Por otra parte, los usuarios tienen la posibilidad de apelar estas remociones invocando el uso justo. A diferencia de este tipo de sistemas, la directiva europea avanza en una nueva dirección, haciendo directamente responsables a las plataformas por alojar contenido sujeto a derechos de autor incluso en ausencia de notificación alguna, y obligándolas a “tomar todas las medidas posibles” para bloquear el acceso a contenidos compartidos que infringen las leyes en esta materia. Sumado a ello, la implementación de filtros automáticos y de carga corre el riesgo de que su dimensión, opacidad y falta de precisión genere niveles desproporcionados de remoción de contenidos legítimos, y sin dudas constituirá un nuevo escenario para la libertad de expresión en internet y para el humor en todas sus formas.

3. Desinformación

La discusión sobre las noticias falsas se instaló en 2016 tras la sorpresiva victoria de Donald Trump.⁶³ Semanas antes de las últimas elecciones,

⁶² Electronic Frontier Foundation (EFF), “NBC issues takedown on viral Obama ad”, 30 de septiembre de 2008, disponible en: <https://bit.ly/1SVR8bC>, último acceso: 9 de noviembre de 2019.

⁶³ El término *fake news* tiene diversos significados y hay quienes los han clasificado bajo criterios disímiles. A los fines del presente artículo, nos referiremos a noticias falsas como aquel contenido deliberadamente falso, con apariencia de autenticidad y con intenciones de engañar al usuario o a la usuaria. También consideramos noticia falsa a aquellos contenidos difundidos a través de servicios de mensajería instantánea en forma de imagen, video o meme, sin autor identificado ni identificable.

titulares como “El Papa apoya la candidatura de Donald Trump” o “Hillary Clinton le vendió armas a ISIS”⁶⁴ circularon a través de cuentas falsas en Twitter, propagandas en Facebook, e incluso aparecieron como resultados de búsqueda en Google.⁶⁵ La discusión adquirió otra dimensión luego de que se difundiera la historia de una pizzería del centro de Washington DC en la que supuestamente existía una red de tráfico sexual infantil dirigida por Hillary Clinton y otros miembros del Partido Demócrata. A raíz de ello, en un episodio luego conocido como “Pizzagate”, un hombre de 28 años de Carolina del Norte viajó hasta allí a rescatar a los niños que estaban siendo explotados y disparó tres veces contra la pizzería.⁶⁶ Diversos medios de comunicación advirtieron que la desinformación había sido determinante en el resultado de las elecciones⁶⁷ y pronto derivó en medidas de censura en algunos países⁶⁸, e incluso comenzó a ser utilizado por figuras políticas para denostar a aquellos medios de comunicación cuyas coberturas periodísticas no les agradaban.⁶⁹

Especialmente en períodos electorales, la desinformación constituye

⁶⁴ Ver, por ejemplo, Ritchie, Hannah, “Read all about it: the biggest fake news stories of 2016”, CNBC, 30 de diciembre de 2016, disponible en: <https://cnb.cx/2wE3PDI>, último acceso: 9 de noviembre de 2019; Roberts, Hannah, “This is what fake news actually looks like – we ranked 11 election stories that went viral on Facebook”, *Business Insider*, 17 de noviembre de 2016, disponible en: <https://bit.ly/2xNhwPI>, último acceso: 9 de noviembre de 2019; Silverman, Craig, “This analysis shows how viral fake election news stories outperformed real news on Facebook”, BuzzFeed News, 16 de noviembre de 2016, disponible en: <https://bit.ly/2Ocl1dp>, último acceso: 9 de noviembre de 2019.

⁶⁵ 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: <https://nyti.ms/2o4NhOA>, último acceso: 9 de noviembre de 2019.

⁶⁶ Siddiqui, Faiz y Svrluga, Susan, “N.C. man told police he went to D.C. pizzeria with gun to investigate conspiracy theory”, *The Washington Post*, 5 de diciembre de 2016, disponible en: <https://wapo.st/2k2clCg>, último acceso: 9 de noviembre de 2019.

⁶⁷ Cortés, Carlos e Isaza, Luisa, “Noticias falsas en internet: la estrategia para combatir la desinformación”, CELE, diciembre de 2017, citando a: Read, Max, “Donald Trump won because of Facebook”, *New York Magazine. Intelligencer*, 9 de noviembre de 2016, disponible en: <https://nym.ag/2XQCK9N>, último acceso: 9 de noviembre de 2019; 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>, último acceso: 9 de noviembre de 2019; Parkinson, Hannah Jane, “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>, último acceso: 9 de noviembre de 2019.

⁶⁸ Fraser, Matthew, “Debate: the legal fight against ‘fake news’ must not veer into censorship”, *The Conversation*, 11 de junio de 2018, disponible en: <https://bit.ly/2WFeQCn>, último acceso: 9 de noviembre de 2019.

⁶⁹ Annenberg School for Communication, *Understanding and addressing the disinformation ecosystem*, Philadelphia, PA, 15-16 diciembre de 2017, disponible en: <https://bit.ly/2GbeyJ2>, último acceso: 9 de noviembre de 2019.

un tema de agenda que impacta en el discurso público y, por ende, en las expresiones de humor, especialmente el humor político, cuya relevancia en periodos electorales puede ser aún mayor. Está claro que cuando hablamos de noticias falsas el humor está exento, sin embargo, desde que las noticias falsas cobraron relevancia a nivel global, la distinción entre sátira e información falsa ha generado incertidumbre, y se ha convertido en un área gris para periodistas, investigadores y *fact-checkers*.⁷⁰ Para algunos especialistas, la sátira y la parodia constituyen un tipo de desinformación por tratarse de formas artísticas que pueden convertirse en desinformación cuando las audiencias malinterpretan el mensaje.⁷¹ Aunque no tienen la intención de causar un daño concreto, tienen ínsito un potencial para desinformar.⁷² Un informe reciente de UNESCO advierte que con frecuencia la sátira puede ser malentendida por los usuarios en las redes sociales, que a su vez la difunden como si se tratase de información verídica.⁷³ En efecto, hay quienes llegaron a preguntarse si aún es posible hacer humor y sátira en la era de las *fake news*.⁷⁴

El año pasado, por ejemplo, Snopes, una reconocida organización de verificación de datos o *fact-checking*, verificó una historia de un portal de sátira estadounidense que acusaba a la cadena de noticias CNN de sesgar sus publicaciones.⁷⁵ La nota decía: “CNN compra una lavadora industrial para darle nuevos giros a sus noticias antes de publicarlas”.

⁷⁰ Smith, Justin E.H., “The end of satire. The toxic disinformation of social media has rendered traditional forms of humor quaint and futile”, *The New York Times*, 8 de abril de 2019, disponible en: <https://nyti.ms/2G7QM0V>, último acceso: 9 de noviembre de 2019.

⁷¹ Wardle, Claire, “Fake news. It’s complicated”, *Medium*, 16 de febrero de 2017, disponible en: <https://bit.ly/2HRC2HJ>, último acceso: 9 de noviembre de 2019.

⁷² Wardle, *ibid.*; ver también el caso en: Ireton, Cherylyn y Posetti, Julie (eds.), *Journalism, ‘fake news’ & disinformation*, UNESCO, 2018, p. 46, disponible en: <https://bit.ly/2QmgNwM>, último acceso: 9 de noviembre de 2019.

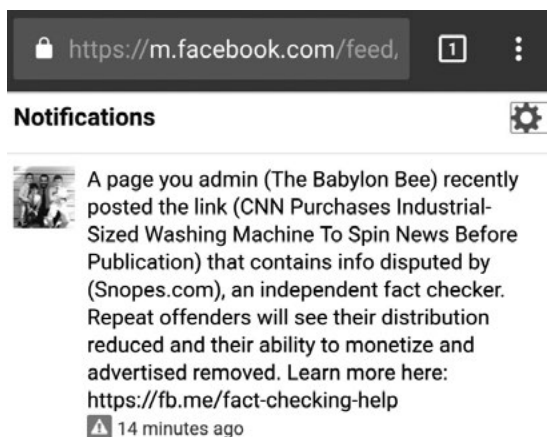
⁷³ Ireton y Posetti, *op. cit.*, p. 17.

⁷⁴ Semley, John, “There’s a deep, dark joke at the heart of Sacha Baron Cohen’s ‘Who is America?’”, *Maclean’s*, 25 de julio de 2018, disponible en: <https://bit.ly/2GIdQ7w>, último acceso: 9 de noviembre de 2019.

⁷⁵ Mikkelsen, David, “Did CNN purchase an industrial-sized washing machine to spin news?”, Snopes, 1º de marzo de 2018, disponible en: <https://bit.ly/2NDMD8T>, último acceso: 9 de noviembre de 2019.



La nota recibió más de 22 mil interacciones en Facebook. Horas después, los usuarios comenzaron a recibir advertencias previas al momento de compartir la nota y finalmente los administradores de la página recibieron una notificación de que su publicación podría verse reducida en alcance.⁷⁶ Según el fundador de Snopes, la política del sitio es verificar cualquier contenido que pueda ser interpretado como verídico.⁷⁷



A raíz de este y otros incidentes similares existe un debate en torno a la necesidad o no de chequear o “flaggear” contenido satírico o de parodia. La mera existencia de este debate impacta en el humor; en definitiva, el hecho de tener que explicar el chiste antes incluso de que genere una reacción atenta contra su propia naturaleza y fines.

⁷⁶ Poynter, “Should satire be flagged on Facebook? A Snopes debunk sparks controversy”, 2 de marzo de 2018, disponible en: <https://bit.ly/2JCO7Ow>, último acceso: 9 de noviembre de 2019; Alcazaren de Leon, Marguerite y Hapal, Kevin, “Satire vs. fake news. Can you tell the difference?”, *Rappler*, 15 de mayo de 2018, disponible en: <https://bit.ly/2LtvKxS>, último acceso: 9 de noviembre de 2019.

⁷⁷ Poynter, *op. cit.*

El caso de Snopes es un ejemplo más de las medidas que las plataformas están tomando en respuesta a la creciente demanda por transparencia y chequeo, algunas plataformas también están adoptando distintas medidas para mitigar la propagación de noticias falsas.⁷⁸ Facebook lanzó una prueba piloto para verificar noticias reportadas como falsas por los usuarios, que actualmente está vigente en catorce países.⁷⁹ Por su parte, Google ajustó recientemente sus algoritmos para que los resultados que arroje el motor de búsqueda sean los más precisos posibles. Para ello, no solo evalúa que los resultados coincidan con las necesidades del usuario, sino también la calidad de los sitios que arroja (entendiendo por “calidad” a los sitios de contenido con mayor experiencia, autoridad, confianza y reputación).⁸⁰ En este sentido, intenta mantener fuera de los resultados de búsqueda el contenido falso o que promueve desinformación, sobre todo luego de que hace algunos años diversos medios mostraran cómo el motor de búsqueda arrojaba contenido extremista y racista entre los primeros resultados ante la pregunta “¿sucedió realmente el Holocausto?”.⁸¹ Algunos elementos comunes a estas distintas iniciativas, tanto públicas como privadas, incluyen dificultades para definir claramente qué constituye desinformación; problemas para determinar la naturaleza de los contenidos que entrarían dentro de esta categoría; y un enfoque que parecería, por lo menos en algunos casos, no mirar el contenido sino las posibles interpretaciones que del mismo se hagan.

Además de las medidas de autorregulación, frente a este fenómeno aparecieron, por otro lado, propuestas legislativas y políticas públicas promovidas por distintos Estados para lidiar con este fenómeno. En Alemania, por ejemplo, el Congreso adoptó en 2018 la NetzDG, una norma que prohíbe, entre otras cosas, la desinformación y sanciona con cuantiosas multas a las empresas en caso de no remover dichos contenidos en un plazo no mayor a 24 horas posteriores a la notificación del usuario.⁸² Otros ejemplos de leyes con fines

⁷⁸ Cortés e Isaza, *op. cit.*

⁷⁹ Lyons, Tessa, “Hard questions: what’s Facebook’s strategy for stopping false news?”, Facebook Newsroom, 23 de mayo de 2018, disponible en: <https://bit.ly/2KP8JB2>, último acceso: 9 de noviembre de 2019.

⁸⁰ Hern, Alex, “Google tweaked algorithm after rise in US shootings”, *The Guardian*, 2 de julio de 2019, disponible en: <https://bit.ly/2Y958YR>, último acceso: 9 de noviembre de 2019.

⁸¹ Cadwalladr, Carole, “Google is not ‘just’ a platform. It frames, shapes and distorts how we see the world”, *The Guardian*, 11 de diciembre de 2016, disponible en: <https://bit.ly/2hzO7Ot>, último acceso: 9 de noviembre de 2019.

⁸² La “Network Enforcement Act (Netzdurchsetzungsgesetz, NetzDG)” o “Ley de aplicación de redes” fue aprobada en 2017 y entró en vigencia el 1º de enero de 2018, disponible en: <https://bit.ly/2qNxOHB>, último acceso: 9 de noviembre de 2019.

similares aparecen en Francia, cuya Asamblea Nacional aprobó en 2018 una “ley sobre la manipulación de la información”⁸³ y en Singapur, donde la legislación contra las noticias falsas acaba de entrar en vigencia.⁸⁴ Todas estas iniciativas fueron duramente criticadas por parte de la comunidad internacional y particularmente los defensores de la libertad de expresión.⁸⁵ Las principales críticas radican en la dificultad para definir a las noticias falsas; la posibilidad de que sea una autoridad estatal quien constituya el “discurso verdadero”; la tercerización de la Justicia que algunas de las medidas implican; y los plazos fijados para que las empresas adopten medidas al respecto.

A nivel regional, en septiembre de 2018 la Unión Europea anunció su “Código de buenas prácticas para combatir la desinformación”, por el cual representantes de las principales compañías (Google, Facebook, Twitter) se comprometieron a adoptar ciertos estándares de autorregulación sobre el tema.⁸⁶ El Código destaca la relevancia de la transparencia en el debate público y la necesidad de combatir la desinformación en tanto importa una amenaza al sistema democrático. Además, establece una amplia gama de compromisos que comprenden desde garantizar la transparencia sobre contenidos patrocinados –en especial la publicidad de partidos políticos– hasta medidas para identificar y cerrar cuentas falsas. También prevé mecanismos de seguimiento, por lo que se realizan estudios para determinar el estado de situación y respecto de los cuales los firmantes deben reportar y comparar sus esfuerzos mensualmente.⁸⁷ Las empresas signatarias se comprometen a invertir en tecnología que les permita resaltar aquellas noticias de fuentes “más fiables y certeras” y “diluir” la visibilidad de aquellas con contenidos falsos o “engañosos” (*misleading*). Si bien la parodia y el humor en general están explícitamente exentos de la definición de desinformación adoptada por el Código, en definitiva, los elementos que el Código utiliza

⁸³ Ayuso, Silvia, “La Asamblea Nacional francesa aprueba la ley contra las ‘fake news’”, *El País*, 4 de julio de 2018, disponible en: <https://bit.ly/34Oo7Yj>, último acceso: 9 de noviembre de 2019; J.M.S., “La ley francesa contra las ‘noticias falsas’ se vuelve contra el propio gobierno”, *ABC*, 5 de abril de 2019, disponible en: <https://bit.ly/32xjssc>, último acceso: 9 de noviembre de 2019.

⁸⁴ Agence France-Presse, “‘Chilling’: Singapore’s ‘fake news’ law comes into effect”, *The Guardian*, 2 de octubre de 2019, disponible en: <https://bit.ly/32BpwQo>, último acceso: 9 de noviembre de 2019.

⁸⁵ Ver, por ejemplo, McAuley, James, “France weighs a law to rein in ‘fake news,’ raising fears for freedom of speech”, *The Washington Post*, 10 de enero de 2018, disponible en: <https://wapo.st/2rwWWCV>, último acceso: 9 de noviembre de 2019.

⁸⁶ Comisión Europea, “Código de buenas prácticas para combatir la desinformación”, disponible en: <https://bit.ly/2xEjvpw>, último acceso: 9 de noviembre de 2019.

⁸⁷ Ver reportes mensuales, *ibid*.

para caracterizar la desinformación no parecen excluirla: 1) información falsa y verificable; 2) creada, presentada o difundida con fines de lucro o para engañar intencionalmente al público; y 3) que pueden atentar contra el interés público.⁸⁸ Gran parte de los casos citados a lo largo de este artículo, precisamente, versan sobre publicaciones y contenidos que cumplen con los tres elementos del test propuesto.

En América Latina, a pesar de algunos proyectos de ley que hay en circulación, no existen leyes “anti *fake news*” ni códigos de conducta en la materia. Sin embargo, Estados y empresas se aliaron durante las campañas electorales de los años 2018 y 2019 a fin de generar sistemas de alertas, verificación y difusión de contenidos chequeados durante los períodos electorales. Ejemplos de ello son Reverso en Argentina (2019), Comprova en Brasil (2018) y Verificado en México (2018). Los problemas de estas iniciativas son similares a las de los demás esfuerzos de *fact-checking* que mencionamos anteriormente.

Por otra parte, y sin perjuicio de dónde o cómo se originen, las políticas de autorregulación de las empresas impactan de manera directa en los usuarios de las plataformas a nivel global. Iniciativas como el código de conducta o la directiva europea sobre derechos de autor, en tanto fomentan la incorporación de tecnologías y prácticas a nivel empresarial, impactan indirectamente en los usuarios de estas empresas mucho más allá de la Unión Europea. Esta realidad indudablemente deja expuesta una desigualdad enorme entre quienes adoptan y promueven las políticas y quienes sufren las consecuencias.

⁸⁸ “Desinformación: según lo dispuesto en el comunicado de la Comisión, a los efectos de este Código, la Comisión y el Grupo de Expertos definen en su informe a la ‘desinformación’ como ‘información verificablemente falsa o engañosa’ que: a) ‘es creada, presentada y difundida con fines económicos o para engañar intencionalmente al público’; b) ‘puede causar daño público’, concebido como ‘amenazas a los procesos políticos democráticos, así como a bienes públicos como la protección de la salud, el medio ambiente o la seguridad de los ciudadanos de la Unión Europea’”. Ver Unión Europea, “Code of Practice on Disinformation”, disponible en: <https://bit.ly/33Dv3r4>.

En su texto original: “Disinformation: as provided under the Commission’s Communication, for the purpose of this Code, the Commission as well as the High Level Expert Group in its report define ‘Disinformation’ as ‘verifiably false or misleading information’ which, cumulatively, (a) ‘Is created, presented and disseminated for economic gain or to intentionally deceive the public’; and (b) ‘May cause public harm’, intended as ‘threats to democratic political and policymaking processes as well as public goods such as the protection of EU citizens’ health, the environment or security’”.

4. Discursos discriminatorios, ofensivos y políticamente incorrectos

“El problema con la sátira es que a todos nos gusta cuando está dirigida a nuestros enemigos, y a aquellos que son objetivamente ridículos”,⁸⁹ señala una periodista en *The Guardian*. ¿Qué sucede cuando la sátira se dirige a grupos vulnerables y minorías, o ataca los principios fundamentales de una creencia religiosa? En tiempos de #NiUnaMenos, #MeToo, de viejos modelos patriarcales y estereotipos que han comenzado a deconstruirse progresivamente, ¿es posible hacer el mismo humor que algunas décadas atrás?

“Nos hemos vuelto más cautos que al principio. La gente está muy sensible”, dice Mike Reiss, el guionista y productor de “Los Simpsons”, una de las series más influyentes de los últimos tiempos y famosa por su singular ironía.⁹⁰ El humor utilizado para burlar a determinadas personas o grupos de personas ya no parece ser recibido de la misma manera que antes. Los lectores, espectadores y usuarios son cada vez más críticos con los chistes machistas, racistas u homófobos. Frente a este escenario, hay quienes invitan a reconsiderar cuidadosamente las limitaciones legales a la libertad de expresión, alegando además que esta debe ser “rigurosamente circunscrita dentro de límites más estrechos cuando está perjudicialmente en conflicto con la integridad social”.⁹¹

Existen numerosos ejemplos de expresiones en las redes sociales que han despertado polémica por su contenido discriminatorio. Hace algunos años, una estadounidense que viajaba hacia Sudáfrica publicó un tweet que recorrió el mundo y se convirtió en blanco de duras críticas: “Rumbo a África. Espero no contraer sida. Es broma. ¡Soy blanca!”. Pronto, su cuenta fue eliminada y ella despedida de la empresa para la que trabajaba. En México se viralizaron memes que utilizaban imágenes de personas indígenas acompañadas de mensajes que burlaban su lengua.⁹² En Argentina, uno de

⁸⁹ Groskopf, Viv, “A new satire must emerge – one that breaks out of the liberal bubble”, *The Guardian*, 13 de febrero de 2017, disponible en: <https://bit.ly/2O43nZ3>, último acceso: 9 de noviembre de 2019.

⁹⁰ Ruiz Jiménez, Eneko, “Los guionistas de ‘Los Simpsons’ nos hemos vuelto más cautos. La gente está muy sensible”, *El País*, 12 de julio de 2019, disponible en: <https://bit.ly/2CxXHxh>, último acceso: 9 de noviembre de 2019.

⁹¹ Yoo, Kwanghyuk, “When does cultural satire cross the line in the global human rights regime?: the Charlie Hebdo controversy and its implication for creating a new paradigm to assess the bounds of freedom of expression”, *Brooklyn Journal of International Law*, N° 2, Vol. 42, 2017, p. 764, disponible en: <https://bit.ly/30GMmpf>, último acceso: 9 de noviembre de 2019.

⁹² Gaona, Pável M., “¿Son discriminatorios los memes que usan imágenes de

los personajes sarcásticos más polémicos de los últimos años es la “Doctora Alcira Pignata”, una cuenta anónima de Twitter que se presenta hacia sus seguidores como “en contra de: árabes, hebreos, homosexuales, negros, peronistas y lacra en general”.⁹³

Frente a este fenómeno que recientemente ha despertado mayor clamor social, han surgido tanto iniciativas estatales como privadas. En Europa, cabe destacar la NetzDG en Alemania, que aborda el tema de discurso de odio y obliga a las plataformas a bajar contenido en 24 horas so pena de incurrir en altas multas. Como mencionamos antes, la norma fue criticada por la vaguedad de sus términos y los incentivos que genera para las empresas. A pocos días de su entrada en vigencia, la cuenta de Twitter de una revista satírica alemana fue bloqueada luego de parodiar a Beatriz von Storch, miembro del partido de Extrema Derecha Alternativa para Alemania, quien unos días antes había criticado a la policía por “apaciguar a las hordas de hombres bárbaros, musulmanes y violadores”.⁹⁴ Twitter suspendió por aproximadamente 48 horas su cuenta por violación al supuesto de discurso de odio que incluye la NetzDG. La Asociación de Periodistas Alemanes dijo que tal medida implicaba censura y advirtió que la normativa podría llegar a ser sumamente peligrosa.⁹⁵

En América Latina, el tema del discurso de odio y discurso discriminatorio es de gran interés en los congresos locales. En México, se presentó en 2018 un proyecto “para prevenir y eliminar la discriminación”.⁹⁶ En un lenguaje legislativo similar, en Argentina desde 2016 y hasta la actualidad se han presentado múltiples iniciativas para modificar la Ley de Actos Discriminatorios vigente e incluir explícitamente la “promoción de la no discriminación en internet”.⁹⁷

indígenas?”, Consejo Nacional para Prevenir la Discriminación (Conapred), México, disponible en: <https://bit.ly/2O56WuZ>, último acceso: 9 de noviembre de 2019.

⁹³ El tweet fue publicado el 29 de mayo de 2014 en la cuenta @drapignata, y posteriormente fue eliminado. Asimismo dicha cuenta se encuentra suspendida.

⁹⁴ Thomasson, Emma, “German hate speech law tested as Twitter blocks satire account”, *Reuters*, 3 de enero de 2018, disponible en: <https://reut.rs/2rm6AI2>, último acceso: 9 de noviembre de 2019.

⁹⁵ *Ibid.*; ver también, “German satire magazine Titanic back on Twitter following ‘hate speech’ ban”, *Deutsche Welle*, 6 de enero de 2018, disponible en: <https://bit.ly/2X5FdOL>, último acceso: 11 de noviembre de 2019.

⁹⁶ Ver Observatorio Legislativo CELE, “México proyecto de ley federal para prevenir y eliminar la discriminación”, 2018, disponible en: <https://bit.ly/2CEECKh>, último acceso: 9 de noviembre de 2019.

⁹⁷ Ver Observatorio Legislativo CELE, “Argentina proyecto de ley sobre modificación de Ley de Actos Discriminatorios”, 2018, disponible en: <https://bit.ly/33Nr1N3>, último acceso: 9 de noviembre de 2019.

En Ecuador, el ex presidente Rafael Correa presentó un proyecto para regular “los actos de odio y discriminación en redes sociales en internet”, incluida una obligación expresa a las plataformas para que remuevan y bloqueen contenidos en un plazo de 24 horas, incluso las réplicas que se hayan hecho del mismo.⁹⁸ Todas estas propuestas legislativas parten de una idea común: que el avance de internet ha venido acompañado de un creciente ánimo de confrontación y que es un medio que permite llevar fácilmente a cabo actos discriminatorios y de odio dada la posibilidad del anonimato, y con un enorme impacto por su capacidad de viralización inmediata. A pesar de perseguir objetivos legítimos, las medidas que plantean todas estas iniciativas adolecen de graves problemas de legalidad y proporcionalidad. Estas acarrearán riesgos muy altos para la circulación del discurso, entre ellos situaciones de censura previa, persecución política y excesiva discrecionalidad de los organismos públicos en su aplicación.

Por otro lado, una característica propia de la actividad legislativa regional es su tendencia a ser reactiva. En este sentido, desde el Observatorio Legislativo Regional en Libertad de Expresión del CELE observamos poco debate previo en torno a las normas que proponen regular contenido en el espacio digital. Dada la complejidad de la arquitectura, diseño y funcionamiento de internet, sería deseable que cualquier intento de regulación cuente con un debate amplio y robusto, y que garantice la participación de todos los actores relevantes, en especial el sector técnico, la sociedad civil y las instituciones académicas.

A nivel privado, las empresas han adoptado políticas explícitas de rechazo a la segregación, la exclusión y la discriminación. Facebook, por ejemplo, tiene una política sobre discurso de odio que cataloga el contenido en tres niveles.⁹⁹ En el primer nivel se incluye contenido violento o deshumanizante en base a una de las condiciones del grupo protegido o cuasi-protegido (género, nacionalidad, religión, país de origen, etc.)¹⁰⁰; en el segundo nivel se incluye contenido con generalizaciones peyorativas hacia grupos o personas protegidas con referencias a deficiencias físicas (higiene, por ejemplo), mentales, morales, o que referencian inferioridad, u ofendan al grupo de otra manera; y en el tercer nivel lenguaje que excluya (por ejemplo, que no se admitan mujeres). La política es compleja y excesivamente detallada con

⁹⁸ Ver Observatorio Legislativo CELE, “Ecuador proyecto de ley que regula los actos de odio y discriminación en redes sociales e internet”, 2017, disponible en: <https://bit.ly/2Q7F6kc>, último acceso: 9 de noviembre de 2019.

⁹⁹ Facebook, “Lenguaje que incita al odio”, disponible en: <https://bit.ly/2KeosMa>, último acceso: 9 de noviembre de 2019.

¹⁰⁰ Facebook distingue entre características protegidas y cuasi protegidas. Ver *ibid*.

ejemplos del tipo de contenido que no sería tolerado en la plataforma. Su aplicación ha generado controversias importantes, particularmente en tanto no distingue por nivel de vulnerabilidad entre los grupos. Bajo esta política, un comentario que excluya a mujeres sería equivalente a un comentario que excluya a hombres. En la misma línea, las expresiones de los supremacistas blancos respecto de afrodescendientes serían tan reprimidas como las expresiones de activistas afrodescendientes respecto de supremacistas blancos o incluso historias de racismo.¹⁰¹ Si bien la política contempla explícitamente una excepción para el humor, en muchos de los casos denunciados públicamente, la empresa no supo captar ni las más gruesas ni las más finas ironías.

Por otra parte, entre 2018 y 2019 Twitter llevó adelante una consulta basada en la propuesta para modificar sus términos y condiciones de servicio en torno a “contenido deshumanizante”. Después de un año de consultas sobre el tema, y habiendo recibido más de mil respuestas de académicos, activistas, organizaciones no gubernamentales, expertos y funcionarios, la compañía decidió limitar el alcance de la norma propuesta y aplicarla solamente a aquellos tweets que utilicen lenguaje “deshumanizante” sobre bases religiosas.¹⁰²

Este abanico de medidas pone en evidencia que hay ciertas expresiones a las que tanto Estados como empresas están optando por no darle cabida. Existe un claro interés en proteger valores de igualdad y no discriminación, y el humor parecería no permear esta barrera. En este contexto, es importante reforzar el respeto por los principios básicos de legalidad, necesidad y proporcionalidad para cualquier tipo de restricción a la libertad de expresión, sobre todo por el rol trascendental que las expresiones de humor político tienen en el discurso de interés público.

IV. Conclusiones

El humor, en tanto expresión que por naturaleza intenta chocar, cuestionar y provocar, ha generado grandes discrepancias a lo largo de la historia. La viralización inmediata, la amplificación de los mensajes y la posibilidad de anonimato que brinda internet han contribuido a exacerbar algunos de estos conflictos. Sin embargo, gran parte de las “nuevas” amenazas al humor vie-

¹⁰¹ Ver, por ejemplo, Guynn, Jessica, “Facebook while black: users call it getting ‘Zucked’, say talking about racism is censored as hate speech”, 24 de abril de 2019, disponible en: <https://bit.ly/2QiZ3F1>, último acceso: 9 de noviembre de 2019.

¹⁰² Twitter Blog, “Updating our rules against hateful conduct”, 9 de julio de 2019, disponible en: <https://bit.ly/2YJdBPt>, último acceso: 9 de noviembre de 2019.

nen de la mano de políticas y prácticas que afectan nuestro objeto de estudio más como un daño colateral que como una medida deliberada.

La velocidad, la permanencia y la accesibilidad de contenidos en internet, sumados a una estructura descentralizada y universal, han producido planteos en torno a los medios para proteger derechos individuales como la honra, la dignidad, los derechos de autor frente a noticias falsas, la desinformación y la discriminación. Estos debates están dando lugar a políticas –públicas y privadas– que atentan contra la deliberación profunda de los casos. La imposición de plazos cortos al estilo NetzDG, la resolución de controversias en manos de empresas y no de justicia estatal, y más recientemente los incentivos para la adopción de filtros (previos y posteriores a su carga) sin denuncia o notificación previa, están ocasionando fuertes incentivos para la automatización de procesos de detección, remoción y moderación de contenidos. En este marco, los contextos, los simbolismos y los modismos que caracterizan a las expresiones satíricas, la parodia y el humor en general –pero particularmente el humor político– se ven indudablemente postergados, invisibilizados, asfixiados. El humor, como muchos otros tipos de expresiones, incluida la crítica y la reflexión, requiere necesariamente ser evaluado y posicionarse en su propio escenario para cumplir con sus propios fines.

La privatización del análisis y la decisión en torno a la remoción, la visibilidad y el alcance de los contenidos en las plataformas es un factor común por lo menos en tres de los supuestos analizados en este artículo: desinformación, discurso discriminatorio y derechos de autor. Esta tercerización, avalada y fomentada desde ciertos Estados, actualmente está sujeta a requisitos meramente parciales de rendición de cuentas y transparencia. Las empresas muchas veces no pueden siquiera notificar a los autores acerca de las limitaciones que aplican sobre su contenido. Por otro lado, el éxito de estas medidas muchas veces se mide solo respecto de la disminución de las demandas formales frente a los tribunales, y de ellas depende la facultad de revisar, evaluar, rechazar o implementar.

Este documento intentó hacer un repaso por las medidas, tanto públicas como privadas, que afectan al discurso humorístico. No obstante, muchas de las interrogantes y las conclusiones que esboza este artículo son trasladables a otras múltiples problemáticas del espacio digital.

Noticias falsas en Internet: la estrategia para combatir la desinformación

Carlos Cortés* y Luisa Isaza**

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

* Carlos Cortés es investigador 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 la London School of Economics. Es consultor en libertad de expresión y regulación de internet.

** Luisa Isaza es 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.

*** Este artículo fue publicado por el CELE en diciembre de 2017.

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>.

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 *news-feed* 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*— hicieron 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,

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>.

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>.

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

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”, *Mediaite*, 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/2f3SyADH>.

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).

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



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.

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>.

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 Dwoskin, 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>.

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

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/2zwWONF>, último acceso: 14 de diciembre de 2017.

el tiempo, ii) el impacto, iii) el rol de la sociedad civil, y iv) la transparencia.

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.

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>.

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.



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

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

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/2swFxKz>.



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

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.

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>.

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 importancia: a través de la herramienta de anuncios –especialmente en

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>.

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 coordinada 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

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

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 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 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/2w3l0ed>.

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>. (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>.

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 insinuado 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

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>.



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 acrítico 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

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>.

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 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;

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>.

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>.

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.

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>.

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, 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

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>.

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.

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

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.

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 al moderación de la desinformación.

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:

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.

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

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

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

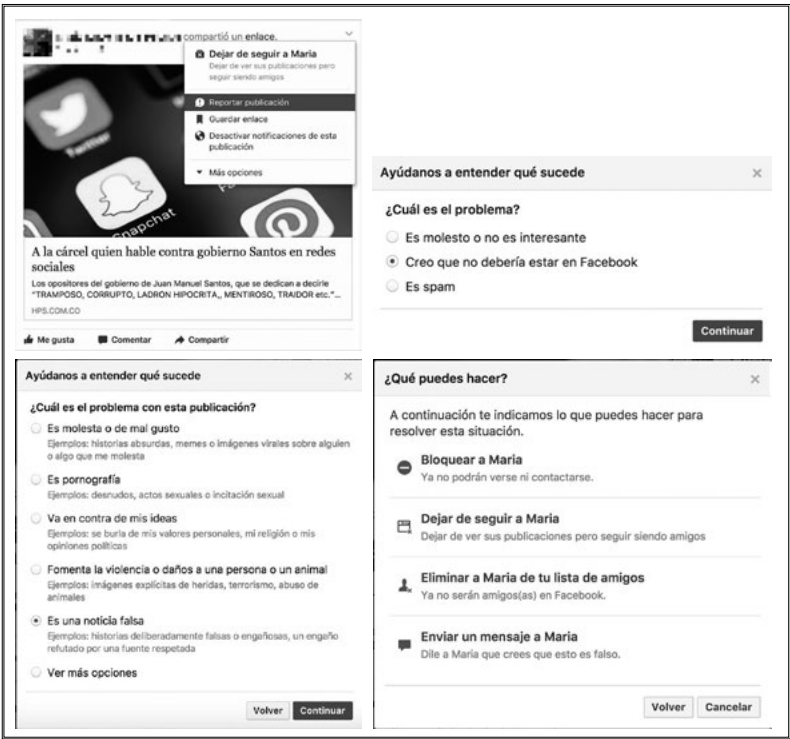
⁵⁵ <http://bit.ly/2fOsvha>

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/2gFFvVw>; Facebook, “Umgang mit Falschmeldungen”, 15 de enero de 2017, disponible en: <http://bit.ly/2gFveyn>; 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>.

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

⁵⁹ Facebook, “Enfrentando engaños y noticias falsas”, 15 de diciembre de 2016,

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 publicación 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

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>.

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

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

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

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>.

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).

2.1. 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

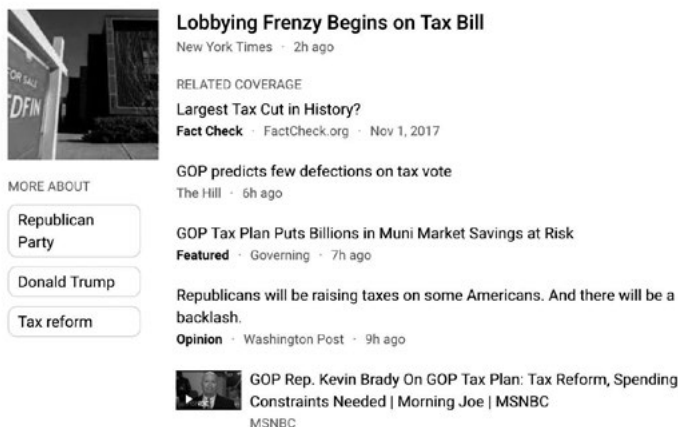
⁶⁷ Wakabayashi, Daisuke y Qiu, Linda, "Google Serves Fake News Ads in an Unlikely Place: Fact-Checking Sites", *The New York Times*, 17 de octubre de 2017, disponible en: <http://nyti.ms/2hO00Ca>.

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.

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.



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)

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.

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.



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 Goo-

⁷⁴ Ver: Schema.org, “ClaimReview”, disponible en: <http://bit.ly/2EpwlDP>, 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>.

gle 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

The image shows a Google search page for Colombia. The search bar contains the text 'la tierra es'. Below the search bar, a dropdown menu displays four autocomplete suggestions: 'la tierra es plana', 'la tierra es redonda', 'la tierra es plana thomas friedman', and 'la tierra es la casa de todos'. To the right of the search bar, there is a sidebar titled '¿Qué predicciones fueron ofensivas?' (Which predictions were offensive?). It contains four checkboxes: 'la tierra es plana', 'la tierra es redonda', 'la tierra es plana thomas friedman', and 'la tierra es la casa de todos'. Below these, there is a section titled 'Las predicciones seleccionadas más arriba son:' (The predictions selected above are:), followed by five radio button options: 'Incita al odio' (Incites to hatred), 'Es sexualmente explícito' (Is sexually explicit), 'Violento' (Violent), 'Actividad peligrosa y dañina' (Dangerous and harmful activity), and 'Otro Razón (obligatoria)' (Other Reason (mandatory)). At the bottom of the sidebar, there is a text input field for 'Comentarios adicionales (opcional)' (Additional comments (optional)) and a link to 'Ver la página de ayuda legal para solicitar cambios de contenido por motivos legales.' (See the legal help page to request content changes for legal reasons). At the very bottom of the sidebar, there are two buttons: 'CANCELAR' and 'ENVIAR'.

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>).

The image shows a Google search interface. The search bar contains the word "snippet". Below the search bar, there are tabs for "Todo", "Imágenes", "Noticias", "Videos", "Maps", "Más", "Preferencias", and "Herramientas". The search results show "Cerca de 58.600.000 resultados (0,44 segundos)". A featured snippet is displayed, defining "Snippet" as a term from English programming. Below the snippet, there is a link to the Wikipedia page. At the bottom of the search results, there is a link to "Comentarios" (Comments). Below the search results, there is a feedback form titled "¿Cuál es tu opinión?". The form has five radio button options: "Es útil", "No me gusta esto", "El contenido es racista, ofensivo o incita al odio", "El contenido es vulgar o sexualmente explícito", and "El contenido es dañino, peligroso o violento". The fifth option, "El contenido es engañoso o incorrecto", is selected. Below the options, there is a section for "¿Tienes comentarios o sugerencias?" with an "Opcional" label and a text input field. At the bottom of the form, there is a note about data usage and a link for legal requests. The form has two buttons: "CANCELAR" and "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

2.2. 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,

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.

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 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

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>.

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

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 cloakers*, 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 inde-

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/2hoiMRc>.

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>.

seable, 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 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

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.

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

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.

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.

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>.

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.

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

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.

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

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.

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>.

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 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.

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

Servicios Over-the-Top: principios fundamentales para su tratamiento regulatorio en Argentina

Maia Levy Daniel*

I. Introducción

Los servicios de comunicación y los medios de comunicación están sujetos a distintas regulaciones, sea por el dominio que tienen sobre infraestructura crítica, por el uso que hacen de recursos limitados cuya titularidad muchas veces detenta el Estado, por razones económicas, o en aras de proteger los derechos humanos potencialmente afectados por estas actividades. En Argentina, al igual que en muchos otros países del mundo, existen leyes tendientes a garantizar la competencia, evitar los monopolios y oligopolios, garantizar el acceso y la no discriminación de las poblaciones minoritarias, entre otras. En este sentido, la Ley de Servicios de Comunicación Audiovisual -con las derogaciones específicas sufridas en diciembre de 2015- regula distintos aspectos del mercado de los medios de comunicación para evitar la existencia de monopolios que constituyan violaciones indirectas a la libertad de expresión y el acceso a la información, y establece pautas en torno a los contenidos. Además, la Ley Argentina Digital regula el mercado de la banda ancha en el país. La regulación tiene distintos aspectos y distintas finalidades. Por un lado, se regula la infraestructura y, por el otro, el contenido, a fin de garantizar la pluralidad de voces que requiere una sociedad democrática. El crecimiento y la expansión de internet, sin embargo, fue considerada solo parcialmente en ambas leyes.

* Maia Levy Daniel es abogada graduada de la Universidad Torcuato Di Tella, magíster en Políticas Públicas por la misma universidad y magíster en Derecho (LL.M.) por la Universidad de Harvard. Se desempeñó como investigadora en el CELE hasta marzo de 2018. La elaboración del documento fue dirigida y contó con los comentarios de Agustina Del Campo, directora del CELE.

** Este artículo fue publicado por el CELE en noviembre de 2016.

Internet ha potenciado la difusión de contenidos, la creatividad y el espíritu emprendedor que derivaron en la incorporación al mercado de nuevos servicios y actores paralelos o complementarios a los existentes. Al crearse nuevos servicios, distintos en muchos aspectos de los ofrecidos tradicionalmente por empresas especializadas en comunicación y transmisión de información, los modelos empresariales también han cambiado. Las redes de alta velocidad y la multiplicación de dispositivos inteligentes, incluyendo teléfonos, tabletas, laptops, etc., coadyuvan a la rápida expansión de estos nuevos servicios, desdibujando las fronteras entre los dispositivos digitales y generando que muchos servicios de comunicación pasaran del mundo físico al digital.¹

Aunque técnicamente no existe una definición unánime del término, ciertos sectores están de acuerdo en llamar servicios “Over the Top” u “OTTs” a los “servicios en línea que pueden sustituir en cierto nivel a los medios de comunicación y servicios de telecomunicaciones tradicionales”.²

Algunos autores definen a los OTTs como un grupo de actores,³ mientras que otros consideran que se trata de un término para calificar una categoría de servicios⁴. En algunas ocasiones se los ha definido como “el contenido, un servicio o una aplicación que es provisto al usuario final en la internet abierta”.⁵ Sin embargo, el común denominador entre las distintas definiciones es que las OTTs se caracterizan por el método de provisión de servicios -léase internet- y no por el tipo de servicio que ofrezca.⁶ Los OTTs dependen plenamente de internet para la entrega del servicio.⁷

El surgimiento de estas nuevas industrias ha generado grandes desafíos para las empresas de servicios tradicionales y para el Estado en vistas a su regulación.⁸ Netflix, YouTube, Skype, Airbnb y WhatsApp, entre muchos otros, constituyen

¹ Comisión Europea, Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, *Una Agenda Digital para Europa*, COM(2010) 245 final/2, Bruselas, 26 de agosto de 2010, p. 4.

² Parlamento europeo, Policy Department A: Economic and Scientific policy, Directorate-general for internal policies, *Over-the-top players (OTTs)*, Study for the IMCO Committee, 2015, p. 11. Traducción libre.

³ Organismo de Reguladores Europeos de las Comunicaciones Electrónicas, *Report on OTT services*, BoR (16) 35, enero de 2016, p. 6. Traducción libre.

⁴ *Ídem*.

⁵ Organismo de Reguladores Europeos de las Comunicaciones Electrónicas, *supra* nota 3, p. 3.

⁶ Parlamento europeo, *supra* nota 2, p. 20.

⁷ *Ídem*.

⁸ Ganuza, Juan José y Vicens, María Fernanda, *Over-the-top (OTT) applications, services and content: implications for broadband infrastructure*, Documento de trabajo N° 13, Centro de Tecnología y Sociedad, Universidad de San Andrés, febrero de 2013, p. 9.

ejemplos típicos de servicios OTT ofrecidos alrededor del mundo. Muchos de estos servicios han reemplazado en distintos aspectos a las formas tradicionales de comunicación y de telecomunicaciones, poniendo a disposición de los usuarios nuevos servicios de mensajería, audio y video, entre otros. Además, han otorgado a los usuarios la posibilidad de interactuar con el contenido, dándoles un rol mucho más activo que el que tenían frente a los servicios de comunicación tradicionales. Los servicios OTT ofrecen el valor añadido de “poder elegir”, frente al modelo tradicional de televisión lineal, por ejemplo.⁹

Existen distintos tipos de aplicaciones en el ámbito digital que se refieren a servicios muy variados. A medida que los servicios en internet crecen en los mercados locales a nivel mundial, distintos sectores alzan sus voces para exigir que se evalúe su regulación legal y/o la compatibilidad de la misma con los derechos humanos potencialmente afectados por estos nuevos actores.

Este artículo busca señalar los principios fundamentales que han de tenerse en cuenta al momento de regular internet, y la base compartida por los servicios OTT, aplicables a todos por igual. Para una mejor comprensión del tema, este artículo ofrece una descripción sobre la situación regulatoria actual en Argentina, así como argumentos que son normalmente esgrimidos a favor y en contra de la regulación de OTTs. Por último, incluye una sección de conclusiones y recomendaciones útiles para la discusión legislativa.

II. Los servicios OTT: qué son y qué significa su presencia en el mercado

Los servicios OTT surgieron recientemente. Skype, por ejemplo, fue diseñado en el año 2003 y WhatsApp fue creado en el año 2009. Netflix, por su parte, entró en América Latina recién en septiembre de 2011 y tiene menos de dos décadas en el mercado estadounidense que lo creó. A pesar de su corta trayectoria, la importancia de las OTTs ha crecido rápidamente en los últimos años, llegando a tener un rol muy relevante para las industrias de tecnología de la información y la comunicación, así como para los consumidores y las empresas.¹⁰ Tanto es así que, en el año 2015, siete de

⁹ Conseil de l' Audiovisual de Catalunya, *Las Plataformas OTT para la distribución de contenidos audiovisuales: ¿una amenaza para el duopolio de la televisión en abierto en España?*, en *Audiovisual OTT, nuevas fronteras y desafíos*, Quaderns del CAC N° 42, Vol. XIX, julio de 2016, p. 24.

¹⁰ Organismo de Reguladores Europeos de las Comunicaciones Electrónicas, *supra* nota 3, p. 3.

cada diez usuarios de internet en Estados Unidos accedían a videos online a través de proveedores OTT.¹¹

En el funcionamiento de un servicio OTT confluyen tres actores distintos e igualmente necesarios: los proveedores de servicios OTT, los fabricantes de artefactos y, por último, los proveedores del servicio de internet (PSI).¹² No existe servicio OTT sin los aparatos y dispositivos (por ejemplo, el módem, teléfono celular o computadora) que facilitan la conexión a internet, o sin un proveedor de servicio de internet para poder conectarse a la red. Para el consumidor, entonces, estos tres servicios son complementarios.¹³

Las redes de banda ancha han posibilitado la existencia de estos servicios. A medida que estas redes aumenten en su alcance y calidad, es razonable pensar que las empresas que ofrecen servicios OTT irán expandiéndose hacia nuevos sectores. En particular, se puede prever que el número de consumidores aumentará a medida que la velocidad de descarga mejore con el tiempo y pueda ofrecerles una mejor experiencia al utilizar estos servicios. En la actualidad, América Latina no cuenta con la infraestructura y tecnología necesarias para brindar un servicio OTT de calidad como en otras regiones, por lo que todavía hoy las empresas proveedoras de servicios tradicionales (por ejemplo, de telefonía, televisión, etc.) no se sienten realmente amenazadas por los servicios OTT¹⁴ como competencia. Distinto es el caso particular de los servicios de mensajería como WhatsApp, que ya se han expandido enormemente de la mano de los *smartphones*.

III. El escenario actual en Argentina

En Argentina no hay hasta el día de hoy normativa específica sobre regulación de servicios OTT. La Ley Argentina Digital (Ley 27.078),¹⁵ dictada en el año 2014, regula las Tecnologías de la Información y las Comunicaciones (TICs), las Telecomunicaciones y sus recursos asociados. El dictado de dicha ley se fundamentó en “(...) las innovaciones tecnológicas, las necesidades de la población, el crecimiento demográfico y el crecimiento del sector de las telecomunicaciones”. En su artículo 2, la ley establece que sus disposiciones

¹¹ Conseil de l' Audiovisual de Catalunya, *supra* nota 9, p. 24. Ver, a su vez, *Seven in 10 US Internet Users Watch OTT Video*, en <http://bit.ly/1ORVChp>

¹² Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, p. 7.

¹³ Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, p. 7.

¹⁴ Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, p. 27.

¹⁵ Ver Ley Argentina Digital: <http://bit.ly/2fDLTfS>

“(…) tienen como finalidad garantizar el derecho humano a las comunicaciones y a las telecomunicaciones, reconocer a las Tecnologías de la Información y las Comunicaciones (TIC) como un factor preponderante en la independencia tecnológica y productiva de nuestra Nación, promover el rol del Estado como planificador, incentivando la función social que dichas tecnologías poseen, como así también la competencia y la generación de empleo mediante el establecimiento de pautas claras y transparentes que favorezcan el desarrollo sustentable del sector, procurando la accesibilidad y asequibilidad de las tecnologías de la información y las comunicaciones para el pueblo.” Sin embargo, excluye explícitamente en su artículo 1 “(…) cualquier tipo de regulación de los contenidos, cualquiera fuere su medio de transmisión.”, por lo que no estipula regulación alguna respecto de los servicios OTT.

Paralelamente, se encuentra vigente la Ley de Servicios de Comunicación Audiovisual, con las modificaciones realizadas por el Decreto N° 267/2015. Asimismo, se está llevando adelante un proceso de debate en torno a la redacción de una nueva ley de servicios de comunicación que, como novedad, contendrá también regulación sobre convergencia.

En el año 2015, por medio del Decreto N° 267/2015¹⁶ se disolvió la Autoridad Federal de Servicios de Comunicación Audiovisual (AFSCA) y el Consejo Federal de Comunicación Audiovisual, creados por la Ley 26.522, y la Autoridad Federal de Tecnologías de la Información y las Comunicaciones (AFTIC) y el Consejo Federal de Tecnologías de las Telecomunicaciones y la Digitalización. Asimismo, se creó el Ente Nacional de Comunicaciones (ENACOM), en la órbita del Ministerio de Comunicaciones, como Autoridad de Aplicación de las leyes Argentina Digital y de Servicios de Comunicación Audiovisual (Ley 26.522), y continuador de los organismos antes mencionados. El ENACOM creó, en el ámbito del Ministerio de Comunicaciones, la Comisión para la Elaboración del Proyecto de Reforma, Actualización y Unificación de las Leyes 26.522 y 27.078. Dicha Comisión tiene por objeto redactar el anteproyecto de la ley, que luego será elevado al Congreso para su discusión y eventual adopción. Metodológicamente, la Comisión se propone convocar “(…) a representantes de consumidores, sindicatos, cámaras, periodistas y a diversos intelectuales y especialistas para la redacción de un marco regulatorio plural y moderno que contemple las nuevas Tecnologías de la Información y las Comunicaciones (TIC), internet, redes de telecomunicaciones y los medios audiovisuales de manera integral y convergente.”

¹⁶ Ver Decreto N° 267/2015: <http://bit.ly/2eLG5mX>

Al momento de redacción de este artículo la Comisión aún trabaja en la elaboración de un proyecto de ley al respecto. En julio de 2016, luego de reuniones con distintos sectores que han expuesto sus puntos de vista y recomendaciones sobre el tema, la Comisión dio a conocer los principios que guiarán la redacción de dicho proyecto.¹⁷ El Principio N° 13 del documento difundido podría ser aplicable a servicios OTT, en tanto establece que “Las aplicaciones de intermediación en la prestación de servicios en el ámbito de las Comunicaciones Convergentes, sin importar el soporte que utilicen, deberán respetar las normas locales que los regulen, respondiendo por los daños y perjuicios que la actividad de intermediación produzca si una vez notificados no cesan con la actividad dañosa. Las mismas deberán registrarse conforme lo establezca la autoridad de aplicación respetando los derechos de propiedad intelectual.” No obstante, hasta tanto no exista un grado mayor de especificidad en torno al contenido específico del proyecto, difícilmente pueda hacerse un análisis adecuado en esta materia.

IV. Argumentos a favor y en contra de su regulación

Conflictos como el relacionado con el arribo de Uber a la Argentina¹⁸ o disyuntivas en torno a cómo cobrar impuestos sobre servicios provistos por internet¹⁹ dan cuenta de la necesidad de promover una discusión en torno a la situación regulatoria de las OTTs. La intención gubernamental de regular los servicios de comunicación convergentes, incluyendo los medios de comunicación tradicionales y los servicios ofrecidos en internet, le dio cierta urgencia al debate.²⁰ La discusión, sin embargo, no radica aún en el tipo de regulación aplicable, sino que abarca también la conveniencia de

¹⁷ Ver los 17 principios: <http://bit.ly/29PiYUA>

¹⁸ Ver, por ejemplo: <http://bit.ly/2ahlAOq> y <http://bit.ly/2ahlLcJ>

¹⁹ La discusión que se ha dado en el año 2014 en la Ciudad de Buenos Aires sobre la posibilidad de cobrar impuestos locales a proveedoras de servicios OTT como Netflix, Spotify y otras es otro ejemplo de la necesidad de fomentar el análisis sobre estos temas. Ver, por ejemplo: <http://bit.ly/2cwkbIR> y <http://bit.ly/2cwiOnl>

²⁰ Al momento de redacción de este artículo no se conocía un texto del proyecto de Ley de Comunicaciones Convergentes. Sin embargo, se puede acceder a los 17 Principios que se propusieron desde la Comisión Redactora de la Nueva Ley de Comunicaciones del Ente Nacional de Comunicaciones (ENACOM) que regirán la redacción de la ley: <http://bit.ly/29PiYUA>

Se encuentran disponibles los comentarios presentados por el CELE a los Principios: <http://bit.ly/2fPPjBk>

regular este tipo de servicios, disyuntiva previa que merece cierta reflexión.

El primer punto que se debe analizar es la necesidad de regular los servicios OTT. Los sectores que se muestran a favor de su regulación argumentan que:

- Los servicios OTT ofrecen los mismos servicios que las empresas de comunicación tradicionales y, por ende, deberían ser regulados de la misma manera a efectos de lograr balancear la situación regulatoria.
- Los servicios OTT son *free-riders* de los proveedores del servicio de internet. Al momento no comparten los costos de las obligaciones con ellos, pagándoles a los proveedores por el uso de sus redes. Quienes argumentan en esta línea sostienen que las empresas de comunicación tradicional han invertido en infraestructura que ahora utilizan las empresas de OTT sin realizar ningún aporte.
- Los servicios OTT tienen un impacto económico negativo en los proveedores de internet, lo que obstaculiza la inversión.²¹
- Por último, cabe destacar que “(...) el interés público reclama (...) del Estado la promoción de políticas públicas que estimulen diversidad de actores, diversidad de perspectivas y puntos de vista, diversidad cultural, diversidad geográfica, diversidad de servicios y aplicaciones. Cuando la política pública no promueve la diversidad, la lógica comercial se desentiende de las necesidades de los sectores más vulnerables (acceso y contenidos)”.²² Por lo tanto, es necesario que, también en internet, el Estado participe implementando políticas que garanticen la diversidad y resguarden los derechos humanos en juego.

Estos argumentos en general encuentran sustento en la necesidad de equilibrar las condiciones de juego (el concepto de “level playing field”). Según esta idea, los servicios que tienen la misma función y compiten entre sí deberían estar sujetos al mismo tratamiento regulatorio.²³ Este concepto es el que destacan los operadores de redes tradicionales que se ven afectados por la expansión de las empresas que ofrecen sus servicios en forma *online*. Se hace hincapié en que los servicios son los mismos, más allá de las diferencias respecto del medio a través del cual se provean, destacando

²¹ Asia Internet Coalition, *Smart Regulation for OTT Growth*, octubre de 2015, p. 5. Disponible en: <http://bit.ly/1mT2e3K>

²² Ver “La convergencia es más que un proyecto de ley” por Martín Becerra. Disponible en: <http://bit.ly/23X2YDF>

²³ Organismo de Reguladores Europeos de las Comunicaciones Electrónicas, *supra* nota 3, p. 22.

la naturaleza de los servicios y no las tecnologías.²⁴

En la contracara, y abogando por no regular o regular de manera distinta a las OTTs respecto de los actores tradicionales en materia de medios y telecomunicaciones, se argumenta que si bien el servicio puede ser similar, la tecnología por medio de la cual se provee es radicalmente distinta y su integridad depende de que se respete esta naturaleza. Entre los argumentos más significativos en esta línea cabe destacar:

- La Declaración Conjunta sobre Libertad de Expresión e Internet, firmada por representantes de distintos organismos internacionales de derechos humanos afirma que “Los enfoques de reglamentación desarrollados para otros medios de comunicación —como telefonía o radio y televisión— no pueden transferirse sin más a Internet, sino que deben ser diseñados específicamente para este medio, atendiendo a sus particularidades”.²⁵ La regulación de los servicios ofrecidos en internet debe considerar los distintos niveles de esta tecnología, y respetar su carácter abierto, dinámico y descentralizado, características con las que no cuentan otros medios de comunicación tradicionales.²⁶ En el caso de la televisión, las diferencias también son esenciales al intentar regular este servicio. Por ejemplo, hasta hace algunos años sólo era posible ver en una pantalla concreta contenidos y productos específicos que eran distribuidos mediante una tecnología dedicada y actores identificados. Hoy en día con los servicios OTT esto cambió radicalmente: la televisión se encuentra disponible en múltiples dispositivos con acceso multipantalla y multicanal, y los contenidos están disponibles en múltiples formas.²⁷ El usuario tiene la posibilidad de

²⁴ Ver documento presentado por Asociación Interamericana de Empresas de Telecomunicaciones (ASIET) ante la Comisión Redactora para la Nueva Ley de Comunicaciones con fecha 18 de mayo de 2016, p. 21. Documento disponible en: <http://bit.ly/1sxiXg6>

²⁵ El Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP), *Declaración Conjunta sobre Libertad de Expresión e Internet*, 1 de junio de 2011, párrafo 1) c.

²⁶ Ver “Internet plural y abierta – Principios fundamentales para la regulación” en <http://bit.ly/1XksnIW>

²⁷ Conseil de l’ Audiovisual de Catalunya, *Estrategias y normativas de los servicios OTT en el marco de los EE.UU. (2005-2015)*, en *Audiovisual OTT, nuevas fronteras y desafíos*, Quaderns del CAC N° 42, Vol. XIX, julio de 2016, p. 16.

interactuar con el contenido y, además, la oferta de contenido ya no es de dominio exclusivo del organismo de radiodifusión tradicional. Cualquier editor de contenido o servicios puede ofrecer sus contenidos en igualdad de condiciones.²⁸

- Las empresas OTT, a diferencia de los operadores tradicionales, no controlan “infraestructura crítica”, tienen costos bajos de entrada al mercado y mucha competencia. Los operadores tradicionales, en cambio, tienen un costo de entrada al mercado muy alto y muy poca competencia. Por ende, el concepto del *level playing field* fracasa desde su concepción.²⁹
- Un servicio OTT forma parte de internet como una aplicación más, por lo que no es necesario regularlo, al igual que sucede con las aplicaciones en general. No tendría sentido tratar de imponer viejas estructuras regulatorias en las nuevas redes de banda ancha. Según Vinton Cerf, “Internet fue diseñada como una plataforma neutral para el intercambio de paquetes sobre la cual todo lo que pudiera ser empaquetado, podría ser transmitido y recibido.” Así, la “noción de OTT es simplemente un malentendido del diseño de este sistema, ya que cualquier aplicación se realiza sobre esta red de alta velocidad de conmutación de paquetes subyacente”.³⁰
- En cuanto al argumento que sostiene que las empresas OTT son *free-riders*, algunos autores señalan que el acceso y las aplicaciones son complementarios. Aplicaciones más complejas generan mayor demanda y disposición a pagar por mejor acceso a la red, y una cobertura de acceso de mejor calidad permite un mayor uso de servicios de mensajería y otras aplicaciones. Por lo tanto, el problema de *free rider* no existe.³¹
- Por último, “muchas de las regulaciones que afectan al sector de las telecomunicaciones tenían sentido cuando existían monopolios, pero con el surgimiento de nuevas tecnologías y la aparición de nuevos

²⁸ *Ídem*.

²⁹ Teniendo en cuenta los costos que la regulación acarrea, para ciertos sectores la solución debería consistir en la eliminación o reestructuración de la regulación de los operadores de redes a fin de promover la inversión, el crecimiento y el acceso al ancho de banda. El *level playing field* se lograría, entonces, a través de la desregulación de todos los proveedores.

³⁰ Vinton Cerf en el panel “¿Somos todos OTT? Los peligros de regular un concepto indefinido” organizado por la Asociación Latinoamericana de Internet (ALAI) en el Foro de Gobernanza de Internet (IGF) de 2016 en Guadalajara, México. Se puede acceder al resumen de la sesión aquí: <http://bit.ly/2hx09MX>

³¹ Williamson, Brian, *Next Generation communications & the level playing field – what should be done?*, junio de 2016, p. 17. Disponible en: <http://bit.ly/2guYerk>

proveedores, la necesidad de regulación para promover la competencia es menos relevante”.³²

Los servicios tradicionales son ofrecidos a nivel local en cada país, mientras que los servicios de internet son ofrecidos en forma global y traspasan los límites territoriales.³³ El regulador ha de tener particularmente en cuenta que las OTTs no están solo sujetas a la normativa local, sino también a la de todos los demás países donde operan, balanceando las ventajas de la tecnología y la innovación global con la necesidad de imponer regulación específica local a quienes la generan. En muchos casos, los servicios OTT son ofrecidos por empresas que no tienen presencia física en el país que impone la regulación, en cuyo caso surgen importantes problemas en torno a la implementación de la ley.

La naturaleza global de internet y la posibilidad de generar fragmentaciones que perjudiquen a los usuarios también son argumentos ofrecidos para no regular los servicios OTT.

V. OTTs y la libertad de expresión

Lo que hace a los servicios OTT distintos de los servicios de comunicación tradicionales es el hecho de que se ofrecen por medio de internet. Tal como afirma Naciones Unidas, los derechos humanos, incluyendo la libertad de expresión, se encuentran plenamente vigentes en internet.³⁴ Internet tiene un gran potencial para permitir a las personas expresarse, participar en las discusiones públicas, así como acceder a bienes y servicios, lo que fomenta los principios básicos de una sociedad democrática. Internet y los servicios disponibles en ella hacen a las comunicaciones más económicas, accesibles y rápidas, y potencian la audiencia para quien se expresa, extendiendo las fronteras de la propia ciudad, provincia, estado, región, etc. La Declaración Conjunta sobre Libertad de Expresión en internet, atendiendo precisamente a este factor distintivo y considerando la

³² Alexander Riobó, Director de Regulación de Latinoamérica en Telefónica, en el panel “¿Somos todos OTT? Los peligros de regular un concepto indefinido” organizado por la Asociación Latinoamericana de Internet (ALAI) en el Foro de Gobernanza de Internet (IGF) de 2016 en Guadalajara, México. Se puede acceder al resumen de la sesión aquí: <http://bit.ly/2hx09MX>

³³ Asia Internet Coalition, *supra* nota 21, p. 7.

³⁴ ONU, Consejo de Derechos Humanos, *Promoción, protección y disfrute de los derechos humanos en internet*, A/HRC/20/L.13, 29 de junio de 2012, párr. 1. Disponible en <http://bit.ly/1XksnlW>

importancia fundamental de internet para la expresión y la participación de la discusión pública, establece como principio general la obligación positiva de los Estados de facilitar el acceso universal a internet.³⁵

Precisamente por la naturaleza y el potencial de internet, conviene evaluar la regulación del medio –internet- desde una perspectiva de derechos humanos y, particularmente, desde la perspectiva de la libertad de expresión. Toda limitación a la libertad de expresión debe ser evaluada en torno al test tripartito, a saber: las limitaciones han de establecerse claramente en una ley, deben ser necesarias en una sociedad democrática y ser proporcionales a los fines legítimos que la medida persiga. Las medidas tendientes a remover contenidos o bloquear el acceso a determinadas aplicaciones en internet deben ser excepcionales, y adoptarse en procesos que garanticen el debido proceso y la notificación adecuada al usuario.³⁶

En todos los casos, la regulación de las OTTS debe evitar transformarse en una intromisión directa del Estado en los contenidos ofrecidos en internet, lo que podría traer consecuencias graves para el derecho a la libertad de expresión de los ciudadanos, tanto en su sentido individual como social, como para el ejercicio de otros derechos humanos fundamentales, tales como el derecho de asociación y de reunión, de participación política, entre otros.³⁷

La vinculación entre las dimensiones individual y social de la libertad de expresión la resumió la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos diciendo que la libertad de expresión es una “condición para que la comunidad, a la hora de ejercer sus opciones, esté suficientemente informada. Por ende, es posible afirmar

³⁵ El Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP), *supra* nota 25, párrafo 6.

³⁶ OEA, CIDH, Relatoría Especial para la Libertad de Expresión, *Libertad de expresión e Internet*, OEA/Ser.L/V/II., CIDH/RELE/INF. 11/13, 31 de diciembre de 2013, párr. 55.

³⁷ Estas cuestiones han sido destacadas recientemente en Argentina por distintas organizaciones de la sociedad civil como consecuencia de la presentación de un proyecto de ley de la Legislatura de la Ciudad Autónoma de Buenos Aires (proyecto de ley D-2298) que modificaba la Ley de Procedimiento Contravencional de la Ciudad. Dicho proyecto de ley permitía, en forma preventiva, a un fiscal o autoridad preventora, “(...) ordenar a las empresas prestatarias del servicio de “internet” bloquear, o denegar el acceso al dominio o aplicación de que se trate en forma parcial o total según la conducta ilícita sea generada dentro de la Ciudad Autónoma de Buenos Aires o produzca sus consecuencias dentro de ésta”.

que una sociedad que no está bien informada no es plenamente libre”³⁸. La Corte Interamericana de Derechos Humanos agregó que “el entorno en línea no solo ha facilitado que los ciudadanos se expresen libre y abiertamente, sino que también ofrece condiciones inmejorables para la innovación y el ejercicio de otros derechos fundamentales como el derecho a la educación y a la libre asociación”.³⁹ Las redes sociales constituyen un claro ejemplo de esto. Individuos de todas partes del mundo han encontrado en ellas un espacio con el que no contaban anteriormente para expresarse, informarse y conectarse con otras personas en forma fácil y rápida. Facebook, Twitter e incluso Whatsapp, entre otras redes sociales y servicios de mensajería instantánea, han devenido esenciales en el ejercicio de derechos fundamentales, en el ejercicio del periodismo, y como plataformas de reunión, de asociación, de protesta, de activismo o de denuncia. Para grupos que nunca habían tenido acceso a los medios de comunicación tradicionales, las redes sociales han venido a proveer una alternativa económica y accesible que les permite, quizás por primera vez, expresar sus opiniones e ideas, y difundirlas.

Los medios de comunicación en internet constituyen un instrumento particular que permite y promueve, en una forma nunca antes experimentada, el ejercicio tanto del derecho a la libertad de expresión como de otros derechos humanos fundamentales. Atento a su importancia, las Relatorías de Naciones Unidas y de la Comisión Interamericana para la Libertad de Expresión han sostenido que la decisión de bloquear el acceso de las personas a sitios web o filtrar cierto contenido de internet debe ser una medida extrema, solo justificada según los estándares internacionales (por ejemplo, cuando sea necesaria para proteger a menores del abuso sexual).⁴⁰ Al funcionar en internet, esta restricción aplica también a cualquier plataforma o aplicación que ofrezca servicios OTT. El bloqueo de un servicio brindado a través de internet tiene que ser una medida de última *ratio* y debe cumplir con los estándares antes mencionados, ya que podría conllevar una violación al derecho a la libertad de expresión y de acceso a la información de las personas, así como a derechos humanos en general.

³⁸ OEA, Corte Interamericana de Derechos Humanos, *Opinión Consultiva OC-5/85, La colegiación obligatoria de periodistas (arts. 13 y 29 de la Convención Americana sobre Derechos Humanos)*, 13 de noviembre de 1985, párrafo 70.

³⁹ OEA, CIDH, Relatoría Especial para la Libertad de Expresión, *supra* nota 36, párrafo 2.

⁴⁰ El Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP), *supra* nota 25, párrafo 3).

VI. Principios fundamentales a tener en cuenta a la hora de regular

Cada servicio OTT cuenta con ciertas particularidades que lo distinguen del resto: Spotify no ofrece el mismo servicio que Netflix (en el caso de Netflix se trata de contenido audiovisual, mientras que Spotify ofrece sólo audio). De la misma manera, la regulación específica de las telefónicas no podría ser directamente trasladada al servicio que ofrece Facebook. Asimismo, son también relevantes las diferencias en los modelos de negocio que actualmente actúan en internet y que requieren distinciones específicas en cuanto a su regulación. Los modelos más conocidos son los de micropagos,⁴¹ el pago por consumo,⁴² suscripción,⁴³ membresía,⁴⁴ freemium/premium,⁴⁵ por publicidad insertada⁴⁶ y de acceso abierto.⁴⁷ Cada modelo de negocios es distinto y rige de manera distinta también en cada tipo de servicio OTT, y estas diferencias importan al momento de regular. Por ejemplo, si se decidiera establecer un impuesto a las transacciones en internet y no a las suscripciones, se estaría perjudicando a un tipo específico de servicios OTT frente a otros que no deberían pagar. No

⁴¹ Micropagos/contenido fraccionado: se realiza una transacción para acceder a algún tipo de contenido (por ejemplo, un artículo en una web, una canción o el acceso al siguiente nivel en un videojuego).

⁴² Pago por consumo: un modelo que nació en el ámbito de la televisión ("pay per view") y luego se extendió a otros espacios. Es un sistema de pago en el que el usuario paga sólo por lo que ve.

⁴³ Suscripción: se cuenta con una base de clientes fija en un tiempo concreto (semanal, mensual, anual) y con un flujo de ingresos también fijos. Por ejemplo, el caso de Netflix, Hulu o Spotify.

⁴⁴ Membresía: consiste en la pertenencia a un grupo que puede ser también cualquier tipo de empresa que oferte servicios o contenidos. Un usuario puede ser miembro de un club de fans o de una comunidad de lectores, pero sería la suscripción la que llevaría implícita un pago por algún tipo de servicio o acceso. Ser miembro supone la pertenencia a algo (un club, periódico, seguro, comunidad, etc.) Un ejemplo es el caso del videojuego Second Life.

⁴⁵ Freemium/Premium: consiste en un modelo que ofrece algún producto o contenido de manera gratuita, mientras que para tener acceso a otra parte de los contenidos hay que pagar (el modelo Premium). Muchas veces este modelo incluye publicidad o marketing agregado a los contenidos (o sólo en los modelos Freemium), a fin de que sustenten el negocio junto con los usuarios Premium. Es el caso de Spotify, Skype, iCloud, Dropbox, entre muchos otros.

⁴⁶ Por publicidad insertada: es otra versión del modelo Freemium/Premium. Consiste en ofrecer contenido gratuito pero con publicidad insertada en los eBooks, frente al servicio Premium, sin publicidad y con otras ventajas.

⁴⁷ Acceso abierto (*Open Access*): se refiere a todo tipo de acceso sin necesidad de suscripción previa o pago. Suele utilizarse para ofrecer material de tipo educativo, científico o académico que está en relación directa con la gestión de adquisiciones y préstamos en las bibliotecas. Nuevos modelos de negocio en la era digital, LIBER, 2014. Disponible en: <http://bit.ly/1tdtTlv>

obstante las diferencias, que pueden abordarse desde distintas perspectivas en una eventual regulación, hay aspectos generales fundamentales que subyacen a todas las empresas que ofrecen servicios OTT distintos. El Estado, al decidir regular la prestación de los servicios OTT, debe considerar una serie de principios básicos que surgen directamente de las características propias de internet, su apertura, descentralización y universalidad que guiarán la labor legislativa en esta materia.

Teniendo en cuenta el análisis desarrollado en las secciones previas, y sobre la base de la bibliografía consultada, la opinión de expertos que abordaron distintos aspectos de la temática y consensos en la materia, a continuación presentaremos los principios generales que consideramos fundamentales a la hora de pensar estos temas:

1. Definir precisamente el concepto de “servicio Over-the-Top” en caso de decidir regularlo

Las definiciones de servicios OTT que se conocen hoy en día abarcan una gran cantidad de servicios variados y radicalmente distintos en cuanto a procedimientos, prestaciones y modelos económicos. Con las distintas definiciones actuales, servicios como los ofrecidos por Netflix o Facebook podrían ser regulados de la misma manera, a pesar de sus radicales diferencias.

La falta de claridad en la definición del concepto otorga discrecionalidad al Estado para decidir, incluso arbitrariamente, qué reglas se aplican a determinados contenidos o servicios y discriminarlos de otros similares o incluso iguales. Esta facultad debería ser lo más restrictiva posible, por las obvias consecuencias negativas que ello traería en términos de libertad.

Para suplir la deficiencia terminológica, algunos países han optado por agrupar determinados servicios OTT de acuerdo a la función que cumplen, por ejemplo. Así, en países como Hong Kong, Singapur y Estados Unidos, los reguladores han diferenciado específicamente los servicios sustitutos de la telefonía tradicional de los servicios emergentes,⁴⁸ estableciendo para ellos una regulación particular. En el caso de Singapur, por ejemplo, la regulación no requiere que todos los proveedores de servicios OTT ofrezcan acceso a servicios de emergencia, pero requiere que quienes no lo provean informen al usuario al respecto.⁴⁹

⁴⁸ Asia Internet Coalition, *Smart Regulation for OTT Growth*, October 2015, p. 7. Disponible en: <http://bit.ly/1mT2e3K>

⁴⁹ *Ídem*.

2. Fomentar el acceso universal a internet

El Estado debe asegurar que la regulación en cualquier caso no genere consecuencias perjudiciales para el acceso de los ciudadanos a internet, entendido no solo como el acceso a la infraestructura, sino como el conjunto de aptitudes requeridas para acceder plenamente a la red, incluyendo la accesibilidad de los contenidos, alfabetización digital, pluralidad de lenguajes, asequibilidad, etc.

Los Estados tienen la obligación de promover el acceso universal a internet a fin de garantizar el ejercicio efectivo del derecho a la libertad de expresión y de otros derechos humanos fundamentales, tales como el derecho a la educación, a la atención de la salud y el trabajo, el derecho de reunión y de asociación, y el derecho a elecciones libres.⁵⁰ En palabras de la Asamblea General de la OEA, “las tecnologías de la información y la comunicación (TIC) son cruciales para el desarrollo político, económico, social y cultural, así como un factor esencial para la reducción de la pobreza, la creación de empleo, la protección ambiental y para la prevención y mitigación de catástrofes naturales”.⁵¹

La regulación en materia de servicios OTT debe considerar el rol fundamental que tiene internet hoy en día tanto para las comunicaciones como para la participación pública, y la consecuente obligación de promover su acceso universal. Las obligaciones establecidas por los Estados a los servicios OTT no deberían constituir una barrera en el acceso universal a internet, sino propender y coadyuvar al cumplimiento del objetivo de acceso universal. Establecer obligaciones regulatorias desproporcionadas o innecesarias a los servicios OTT podría implicar la imposición de barreras que pueden impactar en la naturaleza abierta de internet, fragmentándola e imposibilitando el acceso de usuarios a contenido valioso.⁵²

En este sentido, la CIDH ha expresado, por ejemplo, que si un Estado decidiera imponer obligaciones de registro u otros requisitos a los proveedores de servicios en internet, las mismas podrán considerarse legítimas

⁵⁰ El Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP), *supra* nota 25, Punto 6. a).

⁵¹ OEA, Asamblea General, *Utilización de las Telecomunicaciones/Tecnologías de la Información y la Comunicación para crear una sociedad de la información integradora*, AG/RES. 2702 (XLII-O/12), 4 de junio de 2012.

⁵² Asia Internet Coalition, *supra* nota 21, p. 4.

sólo si superan la prueba establecida por el derecho internacional para las restricciones a la libertad de expresión, resumidas en el test tripartito,⁵³ previamente mencionado.

Dadas las características particulares de internet, la regulación de los servicios OTT puede incidir e impactar negativa o positivamente en la innovación, el crecimiento económico y la disponibilidad de los servicios en el país. La regulación tradicional en general, por ejemplo, impone altos costos a la entrada y requerimientos locales específicos.⁵⁴ Si se siguiera ese modelo, desincentivado por las Relatorías Especiales en Libertad de Expresión de la ONU y de la CIDH, cada país podría establecer sus propias reglas para proteger a sus operadores o usuarios, imponiendo barreras para la obtención de productos innovadores.⁵⁵ De ahí que deba atenderse al objetivo de acceso universal y a la naturaleza de internet al momento de debatir cuestiones regulatorias en materia de OTTs.

3. Respetar el principio de neutralidad de la red

La neutralidad de la red es “el principio según el cual el tráfico de Internet debe ser tratado con igualdad, sin discriminación, restricción o interferencia independientemente de su remitente, destinatario, tipo o contenido, para que la libertad de elección de los usuarios de Internet no esté restringida por favorecer o desfavorecer la transmisión de tráfico de Internet asociado con determinado contenido, servicios, aplicaciones o aparatos”.⁵⁶ Aun cuando no existe una definición unívoca sobre el concepto de neutralidad de la red, al menos hay consenso sobre un piso común: la neutralidad incluye la no discriminación de contenidos, la garantía de no bloqueo y la libertad de uso de dispositivos.⁵⁷

Los contenidos deben ser tratados de igual forma por los proveedores de

⁵³ El Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP), *supra* nota 25, Punto 6. d).

⁵⁴ Asia Internet Coalition, *supra* nota 21, p. 4.

⁵⁵ *Idem*.

⁵⁶ Dynamic Coalition on Network Neutrality, “Model Framework on Network Neutrality”, disponible en inglés en: <http://bit.ly/2fY6Oi0>, y en español en <http://bit.ly/1QFWbaC>

⁵⁷ Cortés Castillo, Carlos, *La neutralidad de la red: la tensión entre la no discriminación y la gestión*, en “Internet y derechos humanos: aportes para la discusión en América Latina”, Centro de Estudios en Libertad de Expresión y Acceso a la Información (CELE), Buenos Aires, febrero de 2014, pp. 30 y 31.

infraestructura, sin importar el origen o el tipo de contenido de que se trate.⁵⁸ Además, los consumidores deben poder acceder a servicios y contenidos de su elección en forma gratuita vía internet público.⁵⁹ Por lo tanto, “no debe haber discriminación, restricción, bloqueo o interferencia en la transmisión del tráfico de Internet, a menos que sea estrictamente necesario y proporcional para preservar la integridad y seguridad de la red”.⁶⁰ El objetivo es poder contar con una internet abierta, rápida y justa.

El principio de neutralidad de la red “debe ser aplicado hacia todos los servicios de acceso a Internet ofrecidos por ISPs, sin discriminar contra la tecnología subyacente utilizada para transmitir las señales”.⁶¹ Dicho principio constituye una garantía para que las empresas que entren a un nuevo mercado en internet puedan competir en forma leal con empresas grandes que ya estén en funcionamiento.⁶² Este tipo de medidas podría disminuir la innovación en ciertos sectores, en tanto desaparecen los incentivos para generar productos o servicios nuevos si no existen medios adecuados para garantizar el acceso a los mismos *online*. De no definir y establecer como principio la neutralidad de la red en la legislación, los operadores de internet fijos y móviles, que proveen a los usuarios la conexión a internet a través de la cual se ofrecen los servicios en línea, podrían contar con los incentivos y las habilidades para regular o bloquear a los servicios en línea que constituyan una amenaza comercial a sus propios servicios, como los de mensajería o VoIP.⁶³

La Declaración Conjunta sobre libertad de expresión en internet, alertada sobre estos riesgos, dispone que “Se debe exigir a los intermediarios de Internet que sean transparentes respecto de las prácticas que emplean para la gestión del tráfico o la información, y cualquier información relevante sobre tales prácticas debe ser puesta a disposición del público en un formato que resulte accesible para todos los interesados”.⁶⁴ No respetar el

⁵⁸ Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, p. 10.

⁵⁹ Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, p. 74.

⁶⁰ OEA, CIDH, Relatoría Especial para la Libertad de Expresión, *supra* nota 36, párr. 30.

⁶¹ Dynamic Coalition on Network Neutrality, *supra* nota 56.

⁶² Mahanagar Doorsanchar Bhawan, Jawahar Lal Nehru Marg; Organismo Regulador de Telecomunicaciones de India, *Consultation Paper on Regulatory Framework for Over-the-Top (OTT) services*, Documento de consulta N° 2/2015, 27 de marzo de 2015, Nueva Delhi, p. 78.

⁶³ Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, p. 75.

⁶⁴ El Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de

principio de neutralidad de la red podría otorgar un gran nivel de control y discrecionalidad a los ISPs para discriminar ciertos tipos de contenidos y opiniones, sea por motivos económicos, políticos, o de cualquier otra índole. Esta discrecionalidad podría perjudicar a la sociedad toda, privándola de opiniones e ideas, y a los usuarios en su individualidad, en tanto podría negarse la difusión de expresiones contrarias a las opiniones generales, masivas o de la mayoría.

Si bien algunos proveedores de internet han expresado su intención de generar un esquema de precio adicional por el acceso a internet y tasas diferenciales para otorgar prioridad a cierto contenido⁶⁵ (generalmente con el argumento de que los productores de contenido aprovechan la infraestructura sin pagar por ella y sus servicios amenazan a los tradicionalmente provistos por ellos),⁶⁶ este tipo de esquemas podría tener efecto negativo en la sociedad,⁶⁷ precisamente porque habilitaría la promoción o degradación de determinados servicios en claro contraste con el principio de neutralidad.⁶⁸

Países como Estados Unidos, Brasil y Europa en general han dictado leyes que resguardan la neutralidad de la red.⁶⁹ En Estados Unidos, por ejemplo, la Comisión Federal de Comunicaciones (FCC, por sus siglas en inglés) lanzó en marzo de 2015 sus nuevas reglas de internet sobre la base del Título II del *Communications Act* de 1934 y la Sección 706 del *Communications Act* de 1996. Estas reglas fortalecen el concepto de neutralidad de la red en ese país, y aplican tanto para servicios de internet fijos y móviles⁷⁰. Estas reglas son:

- No bloqueo: los proveedores de servicio no podrán bloquear el acceso a contenido legal, aplicaciones, servicios o a dispositivos no dañinos.
- No regulación: los proveedores de servicio no podrán afectar o degradar el tráfico legal de internet sobre la base de contenido, aplicaciones, servicios o dispositivos no dañinos.
- No priorización paga: los proveedores de servicio no podrán favorecer a algún tráfico de internet legal sobre otro tráfico a cambio de

Derechos Humanos y de los Pueblos (CADHP), *supra* nota 25, Punto 5. b).

⁶⁵ Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, pp. 11.

⁶⁶ Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, pp. 10 y 11.

⁶⁷ Ver el Reglamento (UE) 2015/2120 recientemente aprobado por la Unión Europea sobre neutralidad de la red: <http://bit.ly/2ckebgv>

⁶⁸ ONU, *La nueva regulación digital: de la Internet del consumo a la Internet de la producción*, eLAC 2018, Comisión Económica para América Latina y el Caribe (CEPAL), julio de 2015, p. 85.

⁶⁹ Ganuza, Juan José y Vicens, María Fernanda, *supra* nota 8, p. 75 y 76.

⁷⁰ Mahanagar Doorsanchar Bhawan, Jawahar Lal Nehru Marg, *supra* nota 62, p. 64.

cualquier tipo de consideración – es decir, no debe haber “carriles rápidos”. Esta regla también prohíbe a ISPs priorizar contenido y servicios de sus afiliados.⁷¹

En Argentina, la neutralidad de la red está consagrada en la Ley Argentina Digital de 2014 y está vagamente mencionada en los Principios publicados por la Comisión redactora para la nueva ley de comunicaciones convergentes. Es fundamental que se establezca y garantice expresamente la neutralidad de la red para proteger y promover la libertad de expresión en internet.

4. Respetar la privacidad y fomentar la transparencia

La ONU ha previsto que “(...) toda captura de datos de las comunicaciones es potencialmente una injerencia en la vida privada y, además, la recopilación y conservación de datos de las comunicaciones equivale a una injerencia en la vida privada, independientemente de si posteriormente se consultan o utilizan esos datos”.⁷² Cualquier regulación de servicios OTT debe atender a la necesidad de respetar la privacidad de las personas y fomentar la transparencia de las acciones del Estado.

Atendiendo a los distintos motivos por los cuales puede querer regularse los servicios OTTs, incluyendo motivos que hacen a la seguridad interior y a la protección e investigación del delito, conviene destacar la importancia de garantizar el derecho a la privacidad tanto en internet como fuera de ella. La interceptación de las comunicaciones en internet sólo debe ser permitida bajo orden judicial competente, en cumplimiento de las garantías del debido proceso. En este sentido, y tal como establece el Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, “(...) el Alto Comisionado de las Naciones Unidas para los Derechos Humanos también ha instado a las empresas del sector de la información y las comunicaciones a que revelen los riesgos y las exigencias de los Gobiernos de manera transparente. (véase A/HRC/27/37) (...) Las obligaciones en materia de transparencia empresarial pueden incluir también el deber de revelar los procesos e informes relativos a la aplicación de las condiciones de servicio y las solicitudes privadas de regulación de contenido y de datos

⁷¹ Mahanagar Doorsanchar Bhawan, Jawahar Lal Nehru Marg, *supra* nota 62, p. 64.

⁷² ONU, Asamblea General, *El derecho a la privacidad en la era digital*, Informe de la Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos, A/HRC/27/37, 30 de junio de 2014, párrafo 20.

de usuarios”.⁷³ Es decir, según las obligaciones de las empresas en materia de derechos humanos, aquellas que brinden servicios OTT deberían publicar estadísticas de los pedidos de interceptación, en caso de que haya habido, en aras de lograr una mayor transparencia.

Además, al regular estos servicios conviene que se establezca legalmente una obligación de transparencia por parte del Estado y de las empresas de servicios OTT respecto al tratamiento de los datos personales, a efectos de que los mismos no se utilicen con fines distintos de aquellos para los que los usuarios brindaron su consentimiento.

5. Promover modelos que fomenten la libertad de expresión y no la restrinjan

Es importante tener en cuenta que cualquier servicio OTT es ofrecido en internet, un medio que tiene características propias y distintas a las de otros medios tradicionales. La Relatoría Especial para la Libertad de Expresión de la CIDH sostuvo que las limitaciones a la libertad de expresión en internet deben considerar las consecuencias que dicha limitación trae aparejadas para el ciberespacio en su conjunto.⁷⁴ Modelos regulatorios que habilitan el bloqueo a ciertos dominios y aplicaciones, por ejemplo, podrían ser perjudiciales para el ejercicio del derecho a la libertad de expresión. Internet debe ser un espacio abierto que promueva la difusión de información e ideas, y la solución regulatoria debe tener en cuenta estas características particulares.

Tal como fue mencionado anteriormente, la Declaración Conjunta sobre libertad de expresión e internet establece que, dado el carácter plural y abierto de internet, una medida que determine el bloqueo obligatorio de sitios web, direcciones IP, puertos, protocolos de red o ciertos tipos de usos debe ser extrema.⁷⁵ Por lo tanto, la regulación de servicios OTT podrá considerar

⁷³ ONU, Asamblea General, *Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión*, A/HRC/32/38, 11 de mayo de 2016, párr. 12.

⁷⁴ OEA, CIDH, Relatoría Especial para la Libertad de Expresión, *supra* nota 36, párr. 53. El Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP), *supra* nota 25, párrafo 3) a.

⁷⁵ Aquí es importante mencionar nuevamente el caso reciente del proyecto de ley de la Ciudad Autónoma de Buenos Aires (proyecto de ley D-2298) que otorgaba a diversas

el bloqueo sólo para casos sumamente excepcionales, ya que implica la privación de un medio esencial para el ejercicio del derecho a la libertad de expresión y otros derechos fundamentales.

Conviene que la prioridad en la elección de modelos de regulación de la comunicación se centre en el usuario de los servicios. Cuanto menos costoso sea el servicio, mayor será el acceso que tendrán los consumidores y así podrán ejercer su derecho a la libertad de expresión, plenamente aplicable a las comunicaciones, ideas e informaciones que se difunden y acceden a través de internet.⁷⁶ En general, los servicios OTT tienen menos barreras de entrada y menores costos, por lo que los precios que deben pagar los consumidores suelen ser menores que aquellos ofrecidos por los servicios de comunicación tradicionales.

6. Fomentar la inversión y el avance tecnológico

Aun cuando el surgimiento de servicios OTT podría generar consecuencias negativas en algunos sectores, “a lo largo de la historia, toda revolución tecnológica ha tenido “ganadores” y “perdedores”, y lo que finalmente debe ser tenido en cuenta es el efecto sobre el bienestar agregado de la sociedad. Por consiguiente, los gobiernos deberían facilitar este proceso y no instrumentar medidas que pudieran entorpecerlo (...) Una buena polí-

autoridades locales el poder de bloquear de nombres de dominio y de aplicaciones. En el caso en cuestión, organizaciones de la sociedad civil sostenían que “(...) *el proyecto de Ley de Presti pone en serio riesgo esta herramienta [internet] al permitir que autoridades administrativas -entre ellas la fuerza policial- pueda censurar contenido para prevenir contravenciones locales. También otorga este poder a auxiliares de la justicia y a fiscales, generando masivas oportunidades para la fragmentación de Internet.*” Asimismo, entre otros puntos, la carta presentada por organizaciones de la sociedad civil sostiene que “(...) *los pactos de Derechos Humanos que protegen la Libertad de Expresión tienen rango constitucional y que por tanto, la restricción de contenidos debe adecuarse a estos estándares. Como lo explicamos en nuestra Carta, este proyecto no cumple con ninguno, siendo especialmente peligroso en relación a la investigación de las actividades que todos los usuarios realizamos conectados a la red.*” Ver: <http://bit.ly/2bSdRHU>

Ver “Internet plural y abierta – Principios fundamentales para la regulación” en <http://bit.ly/1XksnlW>

⁷⁶ OEA, CIDH, Relatoría Especial para la Libertad de Expresión, *supra* nota 36, párr. 2. El Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP), *supra* nota 25, párrafo 1) a.

tica económica y regulatoria debería, sin duda, asegurar y potenciar estas posibilidades en la sociedad”.⁷⁷

Por lo tanto, una regulación con esta perspectiva debe tener en cuenta, principalmente, el impacto que las medidas puedan tener en la inversión del país y en el desarrollo de avances tecnológicos. Por ejemplo, estipular medidas como la necesidad de registro de las empresas que brindan servicios OTT podría constituir un impedimento en el desarrollo tecnológico del país y, ulteriormente, perjudicar a los consumidores, lo que no es recomendable como regulación.

VII. Conclusiones

Este trabajo se ha abocado principalmente a sistematizar y desarrollar los principios que conviene tener en cuenta al diseñar regulación en la materia, particularmente a la luz del régimen nacional e internacional de derechos humanos. Los principios propuestos tienen vocación general, sin perjuicio de las prioridades y particularidades locales que el Estado pueda tener o contemplar al momento de regular.

Como se ha destacado a lo largo de este artículo, en la actualidad no hay normativa clara en relación con los servicios OTT. De hecho, hoy en día no hay una definición exacta del término ni desde lo técnico ni desde lo normativo. En este contexto resulta clave priorizar la definición del objeto a regular, teniendo en cuenta que los servicios OTT se ofrecen en internet, un medio y plataforma para el ejercicio de derechos humanos, con características únicas y particulares, como la apertura, descentralización, pluralismo, entre otras. La regulación de OTTs debe atender a la naturaleza de internet, su importancia para el desarrollo y ejercicio pleno de los derechos humanos, y sus particularidades técnicas, sobre todo al momento de evaluar la proporcionalidad de las medidas regulatorias propuestas respecto de los objetivos legítimos que dichas medidas puedan perseguir.

⁷⁷ Ganuza, Juan José y Vicens, María Fernanda, *El desafío de los contenidos y servicios over-the-top*, en Banda Ancha en América Latina: más allá de la conectividad, CEPAL, febrero de 2013, p. 349. Disponible en: <http://bit.ly/2fCFIUH>

Moderación de contenidos y censura privada: estándares extraídos de la jurisprudencia del sistema interamericano de derechos humanos*

I. Introducción

El Centro de Estudios en Libertad de Expresión y Acceso a la Información (CELE) es un centro de investigación afiliado a la Facultad de Derecho de la Universidad de Palermo en Buenos Aires, Argentina. El Centro se dedica a promover la protección y comprensión de la libertad de expresión particularmente en América Latina. Para ello, el CELE realiza investigaciones de vanguardia capaces de enriquecer el debate público sobre temas de interés público y políticas públicas; genera espacios de debate y reflexión; y desarrolla actividades de capacitación y enseñanza.

Esta presentación pretende acercar posibles respuestas a las preguntas formuladas por el Relator Especial sobre Libertad de Opinión y Expresión de Naciones Unidas en su solicitud de contribuciones sobre la regulación privada de contenidos en la era digital desde las normas que podrían extraerse del Sistema Interamericano de Derechos Humanos.

En palabras del Relator, “las empresas privadas facilitan un intercambio global de información e ideas sin precedentes. Particularmente las redes sociales y los motores de búsqueda se han convertido en fuentes primarias de noticias e información (y desinformación) para cientos de millones de personas. Con ese rol, también se han convertido en guardianes de la expresión capaz de despertar pasiones y generar conocimiento, o incitar al odio, la discriminación, la violencia, el hostigamiento y el abuso. En este

* Este documento fue elaborado por Agustina Del Campo, directora del CELE. Fue presentado en diciembre de 2017 al Relator Especial de las Naciones Unidas sobre la Promoción y Protección del Derecho a la Libertad de Opinión y Expresión, David Kaye, en respuesta a la convocatoria abierta de la Relatoría para recibir contribuciones sobre moderación de contenidos en el espacio digital. Ver convocatoria: <https://bit.ly/36TbgWt>.

marco, el Relator pregunta específicamente “¿Qué medidas deben tomar las plataformas, los actores gubernamentales y otros para asegurarse que estos procesos establezcan garantías adecuadas para proteger la libertad de expresión?”. Esta es la pregunta que hemos buscado abordar, tratando de establecer si el Sistema Interamericano de Protección de Derechos Humanos ofrece alguna orientación, ya sea para las empresas y/o los Estados, para afrontar la regulación de contenido en la era digital, y de ser así, cómo son esos estándares. Concluimos que el Sistema Interamericano sí proporciona algunos estándares que podrían servir de referencia en esta temática y ofrecemos recomendaciones concretas para que los Estados y las empresas privadas mejoren aún más la protección de la libertad de expresión en la era digital.

II. Expresión, difusión y censura

La protección de la libertad de expresión goza de un amplio reconocimiento en los diferentes instrumentos internacionales de derechos humanos en todas las regiones, así como en la Declaración Universal y el Pacto Internacional de Derechos Civiles y Políticos. Existe consenso entre los distintos organismos, defensores y expertos, que la expresión y la difusión son indisolubles y una no puede existir sin la otra. Según lo declarado por la Corte Interamericana de Derechos Humanos hace 30 años, “en su dimensión individual, la libertad de expresión no se agota en el reconocimiento teórico del derecho a hablar o escribir, sino que comprende además, inseparablemente, el derecho a utilizar cualquier medio apropiado para difundir el pensamiento y hacerlo llegar al mayor número de destinatarios. Cuando la Convención proclama que la libertad de pensamiento y expresión comprende el derecho de difundir informaciones e ideas ‘por cualquier... procedimiento’, está subrayando que la expresión y la difusión del pensamiento y de la información son indivisibles, de modo que una restricción de las posibilidades de divulgación representa directamente, y en la misma medida, un límite al derecho de expresarse libremente”.¹

Si bien existe un amplio acuerdo en cuanto a la naturaleza indisoluble de la expresión y la difusión, también existe un amplio acuerdo en que la

¹ OEA, Corte Interamericana de Derechos Humanos, Opinión Consultiva OC-5/85, 13 de noviembre de 1985, La Colegiación Obligatoria de Periodistas (Artículos 13 y 29 de la Convención Americana sobre Derechos Humanos), párr. 31.

libertad de expresión no es un derecho absoluto y que ciertas expresiones constituyen abusos. Sin embargo, no existe un acuerdo universal en cuanto al significado de “abusos”, o incluso en cuanto a los medios para abordar tales abusos. Lo que puede considerarse abusivo en un país, puede no serlo en otro. Y lo que podría entenderse como un medio legal para hacer frente a tales abusos en una región, puede no serlo en otra.

Las empresas privadas que median el contenido y la expresión a través de fronteras tienen un poder enorme (y cada vez mayor) para afectar el discurso público, impactar la libertad de expresión y el acceso a la información. También están experimentando una presión cada vez mayor por parte del Estado y de la sociedad civil para ejercer sus facultades de mediación, estableciendo condiciones de servicio (ToS, por sus siglas en inglés) que prohíben ciertos tipos de contenido considerados abusivos. Algunos de esos contenidos abusivos pueden ser ilegales. Algunos sin llegar a ser ilegales pueden ser problemáticos. Algunos otros pueden ser ilegales en algunos países pero no en otros. Un ejemplo de cada uno puede ayudar a ilustrar: a) la pornografía infantil es ilegal en todas partes; b) el lenguaje agresivo o irrespetuoso sin llegar a ser hostigamiento, aunque no deseado, no es ilegal; c) la blasfemia puede ser un delito punible por la ley en algunos países, y no ser considerada abusiva en otro.

En cuanto a los medios para abordar el contenido “abusivo”, los que son aceptados legalmente también varían de una región a otra. La censura previa, por ejemplo, entendida como una medida preventiva para impedir la difusión de información o ideas, puede entenderse como un medio legítimo para lidiar con contenido abusivo en algunas regiones (por ejemplo, en Europa) y no en otras (por ejemplo, en las Américas).

Si bien los gobiernos y las leyes varían de un Estado a otro, algunas empresas, particularmente las grandes redes sociales, constituyen una sola unidad y necesitan normas comunes para funcionar a través de las fronteras en los diferentes regímenes jurídicos. Sin embargo, al hacerlo no solo deben tener en cuenta las diferencias sustantivas entre los sistemas jurídicos, sino también los diferentes enfoques que pueden adoptar legítimamente o hacer cumplir a nivel mundial como posibles medios de reparación.

III. Artículo 13 de la Convención Americana: ¿En qué se diferencia de otros marcos para la protección de la libertad de expresión en el mundo?

El artículo 13 de la Convención Americana sobre Derechos Humanos (CADH), aunque en apariencia similar al artículo 19 del PIDCP, establece estándares que difieren ligeramente de estos últimos y generan ciertas especificidades que podrían ser particularmente relevantes para el estudio que se está llevando a cabo. Según lo establecido en la Opinión Consultiva 5/85, “la forma en que está redactado el artículo 13 de la Convención Americana difiere muy significativamente del artículo 10 de la Convención Europea, que está formulado en términos muy generales. (...) El Pacto contiene una lista más reducida de restricciones que la Convención Europea, sólo sea porque éste no prohíbe expresamente la censura previa”.²

El artículo 13 en sus partes pertinentes establece que:

1. *Toda persona tiene derecho a la libertad de pensamiento y de expresión. Este derecho comprende la libertad de buscar, recibir y difundir informaciones e ideas de toda índole, sin consideración de fronteras, ya sea oralmente, por escrito o en forma impresa o artística, o por cualquier otro procedimiento de su elección.*
2. *El ejercicio del derecho previsto en el inciso precedente no puede estar sujeto a previa censura sino a responsabilidades ulteriores, las que deben estar expresamente fijadas por la ley y ser necesarias para asegurar:*
 - a. *el respeto a los derechos o a la reputación de los demás; y*
 - b. *la protección de la seguridad nacional, el orden público o la salud o la moral públicas.*
3. *No se puede restringir el derecho de expresión por vías o medios indirectos, tales como el abuso de controles oficiales o particulares de papel para periódicos, de frecuencias radioeléctricas, o de enseres y aparatos usados en la difusión de información o por cualesquiera otros medios encaminados a impedir la comunicación y la circulación de ideas y opiniones.*

El texto de la CADH prohíbe expresamente la censura previa y aborda no solo las restricciones indirectas del Estado a la libertad de expresión, sino también las restricciones privadas a estos derechos cuando esas res-

² OEA, Corte Interamericana de Derechos Humanos, Opinión Consultiva OC-5/85, 13 de noviembre de 1985, La Colegiación Obligatoria de Periodistas (Artículos 13 y 29 de la Convención Americana sobre Derechos Humanos), párr. 45.

tricciones puedan llevar a resultados similares a los controles del gobierno. Como lo expresó la Corte Interamericana, “ni la Convención Europea ni el Pacto contienen una cláusula comparable” al artículo 13 (3). Las garantías en la CADH “(...) fueron diseñadas para ser más generosas y para reducir a un mínimo las restricciones que impiden la libre circulación de ideas”.³

La prohibición de censura previa es amplia y abarcativa. La única excepción es la establecida en el artículo 13(4), que permite la censura previa de los entretenimientos públicos con el único propósito de regular el acceso a ellos de los menores.⁴ La Corte tuvo múltiples oportunidades para abordar la censura previa a través de su jurisprudencia y confirmó de forma consistente su postura de que la censura, ya sea que prohíba la expresión o su difusión, constituye una medida inaceptable por parte del Estado.⁵

Existen diferentes puntos de vista sobre lo que constituye censura previa y cómo se aplica con respecto a los actores estatales y no estatales.⁶ El Relator Especial para la Libertad de Expresión en el Sistema Interamericano ha afirmado que las ordenes gubernamentales de filtrado y bloqueo sin determinación judicial de legalidad constituyen censura previa. Y siguiendo esta línea, la profesora Nunziato argumenta que este es el caso independientemente de que el contenido sea eliminado antes de estar disponible públicamente o después de su publicación, pero antes de la determinación judicial de ilegalidad, y cita a la Corte Suprema de los EE.UU. en su conclusión.⁷

La otra diferencia principal entre la CADH y otros instrumentos es la prohibición expresa de restricciones indirectas, incluidas las generadas por controles abusivos por parte del gobierno o de actores privados (Artículo 13 (3)). Estas restricciones indirectas pueden presentarse en formas diferentes

³ OEA, Corte Interamericana de Derechos Humanos, Opinión Consultiva OC-5/85, 13 de noviembre de 1985, La Colegiación Obligatoria de Periodistas (Artículos 13 y 29 de la Convención Americana sobre Derechos Humanos), párr. 50.

⁴ CADH, Art. 13(4):4. Los espectáculos públicos pueden ser sometidos por la ley a censura previa con el exclusivo objeto de regular el acceso a ellos para la protección moral de la infancia y la adolescencia, sin perjuicio de lo establecido en el inciso 2.

⁵ Corte IDH, caso La Última Tentación de Cristo (Chile), 5 de febrero de 2001, Fondo y Reparaciones, Serie C Nº 73, disponible en http://www.corteidh.or.cr/docs/casos/articulos/Seriec_73_esp.pdf; Y Palamara Iribarne vs. Chile, 22 de noviembre de 2005, Excepciones preliminares, méritos y reparaciones, Serie C, Nº 135, disponible en: http://www.corteidh.or.cr/docs/casos/articulos/seriec_135_esp.pdf.

⁶ Nunziato, Dawn C., Preservar la libertad de expresión en América Latina, CELE, Hacia una Internet libre de censura, 2012. pág. 29-30.

⁷ Nunziato, Dawn C., Preservar la libertad de expresión en América Latina, CELE, Hacia una Internet libre de censura, 2012. pág. 33 citando a Bantam Books vs. Sullivan, 372 US 58 (1963).

y múltiples y el Artículo 13 (3) es abierto en cuanto a los ejemplos que cita. El Sistema Interamericano aplicó esta cláusula a diferentes casos a medida que las restricciones a la libertad de expresión en la región pasaron de ser directas y manifiestas a ser de una naturaleza más sutil e indirecta (por ejemplo, la renovación de licencias, procesos de nacionalidad, asignaciones de publicidad oficial, etc.). Siguiendo la misma lógica, la Oficina del Relator Especial de la Comisión Interamericana de Derechos Humanos afirmó en repetidas ocasiones que responsabilizar a los intermediarios por el contenido publicado por terceros podría infringir esta norma en particular y constituir una restricción indirecta según la CADH.⁸

Además, y siguiendo la conceptualización anterior de censura previa, la eliminación de contenidos realizada por empresas privadas de acuerdo a sus términos de servicio podría también ser considerada censura previa según la Convención Americana. Según el art. 13 (3) este tipo de acción privada podría tener un efecto similar en la censura gubernamental. Como lo expresa el profesor Eduardo Bertoni, los Estados podrían ser considerados internacionalmente responsables al otorgar a las entidades privadas la capacidad de censurar contenido, ya que esas entidades privadas de hecho estarían infringiendo la libertad de expresión de sus usuarios.⁹

IV. El deber de respetar y garantizar en la CADH, el PIDCP y la Convención Europea

Al evaluar casos de libertad de expresión, normalmente el enfoque está puesto en las obligaciones de los Estados de respetar los derechos de las personas, entendiendo esto como una obligación negativa, de abstenerse o de no interferir ilegítimamente con la libertad de expresión, ya sea de forma directa o indirecta. Sin embargo, también existen obligaciones positivas para los Estados de garantizar el pleno ejercicio de este derecho a las personas bajo su jurisdicción, que generalmente están establecidas tanto en la Convención Europea y en el PIDCP como en la CADH.

⁸ CIDH, Libertad de expresión en internet, RELE, 2013, disponible en: http://www.oas.org/en/iachr/expression/docs/reports/internet/foe_and_internet_report_2013.pdf

⁹ E. Bertoni, OC 5-85: Su vigencia en la era digital, en Libertad de Expresión a 30 años de la Opinión Consultiva sobre la Colegiación Obligatoria de Periodistas, CIDH RELE, 2017. Ejemplifica con dos casos: Costeja, donde el propio Estado está promoviendo la censura privada; y Delfi, donde la censura privada se basa en esquemas de responsabilidad de intermediarios.

La Observación General 34 del Comité de Derechos Humanos establece desde el principio en el párrafo 7 que “[l]os Estados partes deben cerciorarse de que las personas estén protegidas de los actos de particulares o de entidades privadas que obstan al disfrute de las libertades de opinión y expresión en la medida en que esos derechos del Pacto sean susceptibles de aplicación entre particulares o entidades privadas”.¹⁰

La Convención Europea tiene un lenguaje similar aunque no tan específico en el artículo 1.¹¹ El Tribunal Europeo ha expresado que “el ejercicio genuino y efectivo de esta libertad no depende simplemente del deber del Estado de no interferir, sino que puede requerir medidas positivas de protección, incluso en la esfera de las relaciones entre individuos [...]”.¹² Si bien los detalles específicos de las obligaciones positivas pueden no estar definidos, y el Tribunal Europeo respalda el principio de un amplio margen de apreciación, distintos autores sostienen que para cumplir plenamente con el artículo 10 de la Convención Europea, bajo la doctrina de obligación positiva, los Estados deben “asegurarse que los intermediarios no se impongan a sí mismos filtros capaces de amenazar la libertad de expresión de los usuarios de Internet por temor a reclamos de responsabilidad y para protegerse de cualquier acusación posterior”.¹³

En el marco interamericano, según el artículo 1 de la CADH, los Estados tienen el deber de asegurar y garantizar el pleno ejercicio de los derechos, entendidos como la obligación de adaptar toda su estructura para que las personas bajo su jurisdicción puedan disfrutar plenamente y ejercer pacifi-

¹⁰ Comité de Derechos Humanos, Observación general N°. 34: Artículo 19: Libertad de opinión y libertad de expresión, CCPR/C/GC/34, 12 de septiembre 2011, disponible en <http://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>. Esta obligación se deriva de la Observación general N° 31 del Comité de Derechos Humanos, “The Nature of the General Legal Obligation Imposed on States Parties to the Covenant” [La naturaleza de la obligación jurídica general impuesta a los Estados Partes en el Pacto].

¹¹ Convención Europea de Derechos Humanos, Artículo 1.

¹² *Özgür Gündem v. Turkey*, N°. 23144/93, ECHR 2000-III, párr. 43. a través del Study of fundamental rights limitations for online enforcement through self-regulation [Estudio de las limitaciones de los derechos fundamentales para la aplicación en línea por medio de la autorregulación] realizado por el Institute for Information Law (IViR) [Instituto de Derecho de la Información] Facultad de Derecho de la Universidad de Amsterdam, pág. 35.

¹³ Estudio de las limitaciones de los derechos fundamentales para la aplicación en línea a través de la autorregulación realizado por el Institute for Information Law (IViR) Facultad de Derecho de la Universidad de Amsterdam, pág. 38 citando a E. Montero y Q. Van Enis, “Enabling freedom of expression in light of filtering measures imposed on Internet intermediaries: Squaring the circle” [Permitir la libertad de expresión a la luz de las medidas de filtrado impuestas a los intermediarios de Internet: la cuadratura del círculo], *Computer Law & Security Review* 27 (2011) 21-35, p. 34.

camente sus derechos humanos.¹⁴ Ciertas obligaciones positivas específicas han sido desarrolladas jurisprudencialmente en el marco del derecho a la libertad de expresión. Los estándares para garantizar el pluralismo y la diversidad en los medios de comunicación sirven como ejemplos.

La Opinión Consultiva 5/85 de la Corte Interamericana establece que:

“El artículo 13.3 no sólo trata de las restricciones gubernamentales indirectas, sino que también prohíbe expresamente “controles... particulares” que produzcan el mismo resultado. Esta disposición debe leerse junto con el artículo 1.1 de la Convención, donde los Estados Partes “se comprometen a respetar los derechos y libertades reconocidos (en la Convención)... y a garantizar su libre y pleno ejercicio a toda persona que esté sujeta a su jurisdicción...”. Por ello, la violación de la Convención en este ámbito puede ser producto no sólo de que el Estado imponga por sí mismo restricciones dirigidas a impedir indirectamente “la comunicación y la circulación de ideas y opiniones”, sino también de que no se haya asegurado que la violación no resulte de los controles particulares” mencionamos en el párrafo 3 del artículo 13”.¹⁵

La obligación de garantizar incluye el deber de actuar cuando los Estados adquieren conocimiento de una violación de los derechos humanos y el deber de tomar las medidas adecuadas para prevenirlas. Así, por ejemplo, las leyes que dejan sin condena las violaciones de derechos humanos cometidas por actores privados son incompatibles con la Convención Americana.

Desafortunadamente, aun no hay jurisprudencia sobre este tema en casos de internet y libertad de expresión dentro del Sistema Interamericano. Sin embargo, los principios y estándares que existen sugieren que las prácticas actuales en las empresas de Internet podrían comprometer la responsabilidad internacional del Estado. Como señaló Bertoni, permitir que las entidades privadas censuren contenido puede constituir una violación de las normas del Sistema Interamericano. Podríamos argumentar que permitir a los intermediarios establecer, interpretar y hacer cumplir condiciones de servicio

¹⁴ Corte IDH, Caso Velásquez Rodríguez vs. Honduras, Serie C, Nº. 4. 1989. párr. 166 disponible en http://www.corteidh.or.cr/docs/casos/articulos/seriec_04_ing.pdf. “La segunda obligación de los Estados Partes es la de “garantizar” el libre y pleno ejercicio de los derechos reconocidos en la Convención a toda persona sujeta a su jurisdicción. Esta obligación implica el deber de los Estados Partes de organizar todo el aparato gubernamental y, en general, todas las estructuras a través de las cuales se manifiesta el ejercicio del poder público, de manera tal que sean capaces de asegurar jurídicamente el libre y pleno ejercicio de los derechos humanos.”

¹⁵ OEA, Corte Interamericana de Derechos Humanos, Opinión Consultiva OC-5/85, 13 de noviembre de 1985, La Colegiación Obligatoria de Periodistas (Artículos 13 y 29 de la Convención Americana sobre Derechos Humanos), párr. 48.

de forma arbitraria, oscura o ambigua también podría constituir una violación de los deberes de los Estados de garantizar el derecho a la libertad de expresión, incluida la prevención por medios razonables de cualquier violación de este derecho.

La Corte Interamericana contextualizó el tema claramente con respecto a los medios de comunicación: “[s]i en principio la libertad de expresión requiere que los medios de comunicación social estén virtualmente abiertos a todos sin discriminación, o, más exactamente, que no haya individuos o grupos que, a priori, estén excluidos del acceso a tales medios, exige igualmente ciertas condiciones respecto de éstos, de manera que, en la práctica, sean verdaderos instrumentos de esa libertad y no vehículos para restringirla. Son los medios de comunicación social los que sirven para materializar el ejercicio de la libertad de expresión, de tal modo que sus condiciones de funcionamiento deben adecuarse a los requerimientos de esa libertad”.¹⁶

De la redacción de la Corte se desprende claramente que el objetivo es proteger los medios principales para el ejercicio de la libertad de expresión, que en 1985 eran los medios de comunicación. Siguiendo la lógica de la Corte, en 2018, los medios de comunicación masivos ciertamente son un vehículo, pero Internet se ha convertido en un medio o un vehículo igual o incluso más poderoso para que las personas comunes y los periodistas puedan ejercer este importante derecho. De acuerdo con la norma de la Corte, las condiciones para el uso de Internet también deberían cumplir con los requisitos de esta libertad.

V. Conclusiones y recomendaciones

1. Los estándares son similares pero no iguales en las distintas regiones y en comparación con los estándares universales

Si bien la libertad de expresión es reconocida universalmente como un derecho humano, vinculada directamente a la gobernabilidad y los valores democráticos, la definición y el alcance del derecho varían, aunque sea ligeramente, de un instrumento internacional a otro. Tales variaciones generan diferentes obligaciones para los Estados en las distintas regiones.

¹⁶ OEA, Corte Interamericana de Derechos Humanos, Opinión Consultiva OC-5/85, 13 de noviembre de 1985, La Colegiación Obligatoria de Periodistas (Artículos 13 y 29 de la Convención Americana sobre Derechos Humanos), párr. 34.

Al igual que otros instrumentos, la Convención Americana establece el deber de los Estados de respetar y garantizar los derechos contenidos en la CADH. Y también establece que cuando se enfrentan a diferentes estándares u obligaciones en base a diferentes instrumentos, los Estados deben hacer cumplir los que proporcionen la mayor protección al derecho en cuestión. En el caso de los Estados en las Américas, la CADH estableció las normas más protectoras del derecho a la libertad de expresión y los Estados deben cumplirlas.

Por lo tanto,

- a. Las empresas deben respetar los derechos humanos en las diferentes jurisdicciones donde operan y los Estados deben tomar todas las medidas razonables para garantizar su cumplimiento;
- b. Para determinar las Condiciones de Servicio a nivel global, y verificar si respetan las normas de derechos humanos, las empresas deben adecuarse a los estándares más protectores en lugar de los menos protectores y adaptarlos a nivel regional, cuando sea necesario. Y
- c. Las empresas deben familiarizarse mejor con las normas universales y regionales de derechos humanos y cómo interactúan y dialogan entre sí.

2. Los deberes de respetar la libertad de expresión implican obligaciones de no realizar interferencias ilegítimas, ya sea de forma directa o indirecta

Los Estados no deben vulnerar los derechos a la libertad de expresión de las personas bajo su jurisdicción, ni directamente a través de la censura del gobierno; ni indirectamente a través de otras regulaciones, incluyendo impuestos, nacionalidad, incentivos monetarios como la publicidad oficial; ni a través de imposiciones de regímenes indebidos de responsabilidad de intermediarios de internet.

Los Estados deben abstenerse de imponer restricciones directas o indirectas a la libertad de expresión en línea y offline y deben abstenerse de presionar, sugerir, imponer indirectamente restricciones y, en particular, de censurar a los intermediarios de Internet de cualquier modo que fuera incompatible con sus obligaciones de derechos humanos.

3. Bajo los estándares interamericanos, los Estados podrían ser responsables de las condiciones de servicio abusivas de una empresa si infringen ilegítimamente la libertad de expresión.

Según los estándares establecidos en la CADH como se explicó con anterioridad, los Estados tienen el deber no solo de proteger la libertad de

expresión del abuso o los controles del gobierno, sino también del abuso y los controles privados cuando no cumplieran con las restricciones permitidas por la Convención Americana.

Los Estados también deben garantizar la libertad de expresión a las personas bajo su jurisdicción mediante medidas positivas, incluida la promulgación de leyes que protejan la libertad de expresión y establezcan claramente cualquier restricción permitida, siguiendo los criterios establecidos por los instrumentos regionales y universales. Finalmente, los Estados deben garantizar que los actores privados, incluidas las empresas de Internet, no vulneren arbitrariamente la libertad de expresión de sus usuarios, incluso mediante medidas de moderación, filtrado, bloqueo, suspensión o cancelación.

Esta conclusión tiene dos repercusiones principales:

- a. Los Estados deben abstenerse de presionar, proponer o regular las condiciones de servicio que de otra manera serían consideradas abusivas o incompatibles con los estándares de libertad de expresión dentro de sus propios países y bajo los estándares regionales o globales.
- b. Los Estados deben establecer regulaciones para que las empresas que operan dentro de sus jurisdicciones cumplan con los estándares de libertad de expresión y no controlen abusivamente la circulación de opiniones e ideas ni excluyan a grupos específicos o ideas del debate al mismo tiempo que por supuesto se respete la capacidad de las empresas para realizar negocios. Esto implica que:
 - (i) Los Estados deben garantizar que las Condiciones de Servicio sean claras y transparentes;
 - (ii) Los Estados deben garantizar que la aplicación y el cumplimiento de las Condiciones de Servicio sean transparentes y respeten los derechos humanos, incluida la libertad de expresión, la no discriminación y el debido proceso.
- c. Lo anterior no debe interpretarse como un derecho de acceso a un determinado foro o establecer condiciones de servicio específicas para las empresas de Internet. Si bien nadie tiene derecho a acceder a Gmail, Facebook o Twitter sin aceptar sus condiciones de servicio y cumplir con ellas, los Estados no pueden otorgar a las empresas de Internet el derecho de aplicar de forma arbitraria y oscura tales condiciones de servicio en detrimento de derechos humanos básicos, como los anteriormente enumerados.

4. Las empresas deben cumplir con los estándares internacionales de derechos humanos y ser transparentes con respecto a sus Condiciones de Servicio, reglas y procesos.

Si las empresas establecen condiciones de servicio locales, globales o mixtas dentro de las diferentes jurisdicciones, deben asegurarse de que esas condiciones respeten los derechos humanos, no sólo en su noción teórica sino en su aplicación, ejecución y cumplimiento.

- a. Las empresas deben cumplir con las normas de derechos humanos en todas las jurisdicciones donde operen.
- b. Las empresas deben establecer y publicar claramente sus condiciones de servicio sin ambigüedad.
- c. Las interpretaciones y la aplicación de las condiciones de servicio deben ser transparentes y estar sujetas al control y la comprensión del usuario. La implementación o ejecución arbitraria de las condiciones de servicio podría constituir una violación a la libertad de expresión.

Estudio sobre la Junta de Supervisión de Facebook: modificar las expectativas*

I. Introducción

La Corte Penal Internacional es conocida comúnmente como la “Corte mundial”. Con el anuncio de Facebook de crear un organismo independiente para examinar la aplicación y el cumplimiento de sus condiciones de servicio en sus 2.7 mil millones de usuarios en todo el mundo, el término “Corte Mundial” probablemente ha ganado al menos un nuevo significado más.

Desde la creación de este organismo, ha habido voces a favor y en contra. Los conceptos compartidos por ambos lados son 1) que esta es una iniciativa de autorregulación y no puede reemplazar una instancia judicial de control; 2) tal como es en la actualidad, solo aborda algunos de los problemas que tienen los usuarios con la moderación de contenido; 3) hay más preguntas que respuestas con respecto a la naturaleza, el alcance, el diseño, las expectativas, etc. 4) la creación de esta Junta, aunque es una propuesta interesante, no debe quitar atención a las obligaciones de debido proceso frente a medidas de restricción y curación del contenido, y las obligaciones de transparencia respecto de las normas y la toma de decisiones, así como su implementación, ya sean impuestos por ley o autorregulados.

* Este documento fue elaborado en mayo de 2019 por Agustina Del Campo, directora del CELE, y contó con los comentarios de Franco Serra y Paula Roko, investigadores del CELE. Fue presentado como contribución a la consulta global abierta desarrollada por Facebook en el marco de la creación de su Junta Independiente de Supervisión.

II. Antecedentes e implicancias de la Junta de Supervisión

Después de un par de años de escándalo tras escándalo y un conjunto creciente de reglas, Facebook está buscando legitimidad para gobernar lo que la Corte Suprema de los EE. UU. ha dictaminado como “el nuevo foro público”. Las reglas han adquirido complejidad y ambigüedad a lo largo de los años y con cada apéndice (las *internal guidelines* se modifican aproximadamente cada semana, y se incluyen un total de cerca de 2000 cambios por año, de acuerdo con la información proporcionada por FB). La cantidad de contenido “denunciado” diariamente por los usuarios por estar en contra de esas reglas es significativo.

Debido a la presión manifestada de diferentes maneras por parte de usuarios y gobiernos, la empresa ha cambiado la moderación de contenido de un modelo reactivo a uno proactivo, ha automatizado la mayoría de sus prácticas de detección de contenido no deseado (aunque no está claro qué parte de sus prácticas de eliminación es automática) y enfrenta debates extremadamente complejos sobre ética, responsabilidad social empresarial, responsabilidad civil, daños y perjuicios y libertad de expresión, entre otros temas.

La decisión de crear un consejo independiente o una junta de supervisión para Facebook, con un proceso abierto y consultas, es quizás la decisión más drástica que hemos visto de FB en los últimos años, al menos con respecto a su política de contenido. Y constituye un primer paso de parte de las empresas de Internet para incorporar actores externos en sus procesos de toma de decisiones. Después de cambiar la retórica de un discurso orientado a la libertad de expresión a un discurso sobre seguridad en los últimos años y de ocupar una función de curaduría defensiva de tipologías de contenido cada vez más complejas, las *Blue Print for Content Governance and Enforcement* [NT: el Plan de Gobernanza de Contenido y Aplicación] de Zuckerberg (noviembre 2018) parecen marcar un nuevo punto de partida para la empresa en materia de regulación de contenido.

En primer lugar, la nota de Zuckerberg apela directamente a la necesidad de aumentar la legitimidad de la gestión y la toma de decisiones; en segundo lugar, hay un compromiso de crear un consejo de revisión externo independiente, cuyas decisiones serían vinculantes y públicas. Aún se desconocen muchos de los detalles. En tercer lugar, invita de forma expresa a los Estados a definir lo que esperan de un régimen de moderación de contenido y detalla al menos dos iniciativas concretas con las que FB ya se ha comprometido en Europa: 1) el acuerdo firmado con el Presidente Macron para trabajar en una nueva regulación de contenido; 2) su trabajo hacia un nuevo marco

europeo para la moderación y regulación de contenido en los próximos dos años. Aunque FB había anticipado este cambio con respecto a la legislación durante las audiencias ante el Congreso de los EE. UU. en 2018, esta es la primera comunicación que FB envía a su propia comunidad oficialmente para dar la bienvenida e incluso obligar a los Estados a regular. Como afirma Zuckerberg: “A fin de cuentas, los servicios deben respetar las leyes locales con respecto a contenidos, y creo que todos se beneficiarían de una mayor claridad sobre cómo los gobiernos locales esperan que la moderación de contenido funcione en sus países”.

El reconocimiento de que “un sistema completo debe abordar tanto la gobernanza como la aplicación” es sin duda un avance positivo en los debates de moderación de contenido, en particular los relacionados con transparencia. La sociedad civil y los académicos de todo el mundo expresaron su preocupación por la falta de transparencia y los niveles de discreción que gozaban las empresas de Internet y han estado trabajando con ellos a lo largo de los años para mejorar la transparencia sobre reglas y procesos.¹ Con el nuevo enfoque, FB puede mejorar la transparencia de los procedimientos y su implementación al tiempo que toma decisiones vinculantes y públicas.

Pero quizás lo que es aún más interesante es el cambio de política con respecto a la administración del contenido en la plataforma y lo que parece ser una nueva disposición para “compartir” la responsabilidad (o la culpa) sobre la creación de reglas. “A medida que pensaba en estos problemas de contenido, me fui convenciendo de que Facebook no debería tomar tantas decisiones importantes sobre libertad de expresión y seguridad por nuestra cuenta”, escribió Zuckerberg en noviembre. Los activistas de derechos humanos lo han estado diciendo durante años. ¿Qué cambió? ¿Y cuánto cambió?

La creación de una junta que funcione como una auditoría de las decisiones de esta plataforma sobre la moderación de contenido podría fortalecer el ejercicio de la libertad de expresión en línea tanto como podría perjudicarlo. Si se entiende a la Junta como un proceso interno, destinado a servir a la empresa para tratar temas complejos de libertad de expresión y su equilibrio con otros derechos, unificar criterios y ayudar al sistema a dialogar con las normas internacionales de derechos humanos, los resultados podrían ser más positivos para todo el ecosistema. Si se lo entiende como un sistema de reemplazo para los ya débiles mecanismos de debido proceso dentro de la estructura de toma de decisiones de la plataforma, y como un organismo que crearía e interpretaría solo reglas privadas (derecho contractual), podría

¹ RDR, GNI, por nombrar algunos.

dañar tanto al ecosistema como a la empresa. Tampoco cumpliría el propósito para el cual fue creado. Si lo que FB quiere y necesita es legitimidad, solo se la puede construir a partir del diálogo entre las normas privadas de autorregulación y las normas internacionales de derechos humanos existentes.

Que esta Junta pueda lograr legitimidad o no depende, en gran medida, de cómo esté estructurado el organismo, los objetivos que se establezcan, los requisitos que los candidatos deban cumplir para ser parte de este organismo, la naturaleza de sus decisiones (si de hecho es un órgano de apelación), y la transparencia y publicidad de sus decisiones y razonamientos.

III. Legitimidad en los órganos decisorios:

El concepto de legitimidad puede abordarse desde una dimensión sociológica y normativa, y la distinción podría ser útil para ilustrar este punto. Por un lado, cuanto más positiva sea la actitud del público hacia la facultad de gobernar de una institución, mayor será su legitimidad popular.² Sin embargo, esta legitimidad es frágil. Particularmente para la Junta de FB, que no tendrá una reserva de legitimidad acumulada durante una larga historia a la que recurrir. Por otro lado, la legitimidad también puede tener un fundamento normativo, refiriéndose a si la autoridad está fundada en el derecho.³ Dado el desafío global que enfrenta FB, la legitimidad de la junta de supervisión debe fortalecerse en ambos frentes, pero aún más en este último. La construcción de una sólida legitimidad normativa podría proporcionar un estándar para juzgar a la Junta y decidir si merece apoyo. Además, la legitimidad normativa puede influir en la legitimidad sociológica, o en la percepción de autoridad justificada, y por lo tanto, en la medida en que la sociedad apoyará o socavará el trabajo de la Junta de FB.

Existe abundante literatura sobre qué elementos contribuyen a la legitimidad de los órganos decisorios. Incluso hay escritos concretos sobre la legitimidad de los organismos decisorios internacionales. ¿Qué determina su legitimidad? Los académicos y los profesionales han identificado tres elementos clave para la legitimidad de los mecanismos decisorios internacio-

² Richard H. Fallon, *Legitimacy and the Constitution* [La legitimidad y la constitución], 118 HARV. L. REV. 1787 (2004–2005)

³ Daniel Bodansky, *The Legitimacy of International Governance: A Coming Challenge for International Environmental Law?* [La legitimidad de la gobernanza internacional: ¿un desafío que se avecina para el derecho ambiental internacional?], 93 AM. J. INT'L L. 596, 601 (1999).

nales: 1) decisiones justas e imparciales; 2) una interpretación y aplicación de reglas coherentes con su alcance y propósito; y 3) que el organismo sea transparente, independiente y con normas democráticas.⁴

Las decisiones justas e imparciales deben estar en el centro de cualquier sistema decisorio, no solo internacional, para alcanzar la legitimidad. Tradicionalmente se relaciona lo “imparcial” con la independencia y, por lo tanto, se ha centrado en los procesos de nominación y selección, la calidad y solidez de los informes, decisiones y recomendaciones, el discurso público y los escritos de los miembros del tribunal. Todos estos elementos tienen definiciones concretas y relevantes en una vasta jurisprudencia internacional y comparativa que no debe ignorarse.

La equidad, por otra parte, según la profesora N. Grossman, “no tiene por qué implicar una proporción igual de fallos a favor y en contra de una parte determinada; de hecho, la mayoría de los mecanismos decisorios de los derechos humanos emiten muchas más sentencias contra los Estados que contra las partes actoras y ambas partes pueden seguir considerándolos legítimos”, pero necesariamente requieren equidad procesal, debido proceso y una aplicación e interpretación coherente de la ley.

En cuanto a la interpretación y aplicación de las reglas de acuerdo con su alcance y propósito, este es probablemente uno de los elementos clave para la legitimidad de la junta de supervisión de FB. Las normas comunitarias tienen un alcance y un propósito, y dialogan diariamente con otras normas, incluidas las normas nacionales, regionales e internacionales de derechos humanos. Las sentencias y los fundamentos de las decisiones de casos individuales (si FB lo decide) deben seguir la misma lógica y diálogo entre los estándares de la comunidad y las normas de derechos humanos. Además, esa lógica debe ser transparente y publica. El razonamiento independiente y objetivo detrás de cada decisión es lo que un consejo de expertos agrega a un sistema cerrado como el equipo legal existente en FB.

Finalmente, la transparencia en este contexto puede definirse como una cualidad: una oficina u organismo, ya sea judicial o de otra índole, es transparente “cuando crea las condiciones que permiten a la sociedad comprender de manera plena y clara cómo actúan, las razones detrás de sus actos, así como los costos y recursos asociados con esas acciones”.⁵ Este factor

⁴ N. Grossman, *Legitimacy and International Adjudicative Bodies* [Legitimidad y organismos decisorios internacionales], 41 Geo. Wash. Int'l L. Rev. 107 (2009-2010)

⁵ M. Pulido Jiménez, M. González Armijo, M. Sánchez de Tagle, S. Ruiz Cervantes, J. Sáenz Andujo, Hacia un modelo de transparencia y acceso a la información en el sistema interamericano de derechos humanos, en Desafíos del Sistema Interamericano

afecta también los dos anteriores, ya que sin cierto grado de transparencia (manifestado a través de alguna de las siguientes acciones: publicar las decisiones o que estén a disposición pública; proporcionar razones para las decisiones; identificar a quienes toman las decisiones, sus discrepancias y concurrencias, etc.) no hay forma de evaluar si un organismo está sesgado o no, o si sus decisiones e interpretaciones de las normas están dentro del alcance razonable y el propósito de la ley.

Si bien la legitimidad de los organismos decisorios internacionales proviene tradicionalmente de que los estados acepten su jurisdicción, la legitimidad de la Junta de FB derivaría de: i) la construcción de una base sólida de decisiones justas; ii) una interpretación coherente y persuasiva (la persuasión es una de las funciones de la legitimidad) de las reglas de comunidad en relación con las normas internacionales de derechos humanos; iii) y mecanismos holísticamente transparentes.

IV. Representación vs diversidad

Uno de los objetivos principales de la Junta de FB es proporcionar mayor legitimidad a su sistema de moderación de contenido. A menudo se menciona la diversidad como una dimensión que agrega valor a la legitimidad de un organismo, particularmente cuando el espacio que gobierna el organismo es diverso. Sin embargo, la diversidad tiene muchos significados y no debe confundirse con la representación.

Obviamente, la representación podría aumentar la legitimidad de las decisiones de la Junta ante los ojos de la comunidad en general, particularmente para aquellos que logran estar representados dentro de la Junta. Sin embargo, una Junta de 40 miembros como la que se propone difícilmente podría representar a una comunidad de 2.700 millones de miembros. Es prácticamente imposible representar a todos. No se podría lograr una representación directa, proporcional o incluso asimétrica en una proporción tan desigual y debido a la gran escala de FB, probablemente no podría haber una relación viable con la que trabajar. Este problema debe reconocerse e incorporarse en el diseño de la Junta de supervisión para generar expectativas realistas. Hacer caso omiso de ello probablemente engañaría a los usuarios, creadores y público y socavaría gravemente el objetivo final.

Ante la falta de representación que podría proporcionar la diversidad de voces, la experiencia técnica podría ser la mejor alternativa. No sería una Junta representativa, sino un órgano técnico capaz de supervisar la implementación de FB de su propia política, en cuanto a su relación con las normas de derechos humanos. Esta membresía técnica debería considerar la diversidad cultural, geográfica y de género en su conformación para permitir un diálogo real entre diferentes expertos y evitar un enfoque único de voces / puntos de vista. La diversidad es clave para garantizar el acercamiento y el entendimiento cultural a través de las fronteras; para lidiar con nuevas cuestiones interjurisdiccionales y servir a la vasta comunidad a la que FB sirve. No hay un modelo único para adaptarse a cada necesidad. Decidir cómo será la Junta de FB implica necesariamente definir y aceptar lo que no será.

V. ¿Qué se espera del Consejo en términos generales?

El borrador de la carta constitutiva sugiere que “la Junta será un cuerpo de expertos independientes que revisarán las decisiones de contenido más desafiantes de Facebook, centrándose en casos importantes y en disputa. Entre las primeras preguntas que pueden plantearse se encuentran ¿para qué? ¿Cuál es el objetivo final de que el Consejo revise estas decisiones? Esta pregunta afecta directamente el tipo de estructura que FB está intentando crear. ¿Es este un Tribunal Supremo, un Tribunal de apelaciones, un tribunal popular? ¿O es un panel de arbitraje de pares? ¿Puede ser un organismo asesor o ya se ha decidido que será una instancia de revisión? ¿Cómo interactuará este Consejo con otras estructuras existentes dentro de Facebook (seguridad y protección, políticas y difusión, por nombrar algunas?)

Si de hecho se trata de un “enfoque similar a la Corte Suprema”, lo que parece ser el fundamento de su creación, los Profesores Klonick y Kadri de Yale sostuvieron que “lo que esto significa realmente para la libertad de expresión y el debido proceso en Internet dependerá de la respuesta a una pregunta clave: ¿Cuán similar será el “Tribunal Supremo de Facebook” a la Corte Suprema de Justicia de los Estados Unidos de América?”⁶ Como argumentan en su artículo del New York Times, un elemento clave de la Corte Suprema de los EE. UU. (como en cualquier otra corte suprema) es que está sujeta a un conjunto de reglas que permanecen inalterables a lo largo del tiempo: la Constitución. Pero,

⁶ Klonick, Kate y Kadri, Thomas, “How to Make Facebook’s ‘Supreme Court’ Work”, The New York Times, 17 de noviembre de 2018, disponible en: <https://nyti.ms/37tZ6nk>.

como se describe en la breve sección de antecedentes de este documento, las políticas de FB cambian cada semana y hasta el momento no se ha adoptado de manera formal ninguna norma o regla específica para ilustrar ese proceso. ¿Cuál debería ser el equivalente para la Junta de FB?

El relator especial de la ONU, David Kaye, sostiene que las normas internacionales de derechos humanos deben ser las reglas inamovibles que guíen la moderación de contenido en línea.⁷ La adopción de tales estándares proporcionaría una base universal y un entendimiento compartido por muchos sobre qué significa la libertad de expresión y qué garantías deben tenerse en cuenta al limitarla. También proporcionaría un lenguaje común para definir y comprender algunas restricciones, evitando términos contradictorios y en constante expansión (como “discurso de odio”). En última instancia, la adopción de reglas universales sobre derechos humanos, en particular la libertad de expresión, también garantizaría cierta certeza contra la discriminación y el abuso, ya sea que estos provengan de gobiernos, usuarios o grupos de interés, y contribuiría a garantizar que nadie esté “a priori” excluido del debate público, que es la norma establecida por la Corte Interamericana de Derechos Humanos en la Opinión Consultiva 5 de 1985, una de las interpretaciones internacionales más progresistas y protectoras de la libertad de expresión y el acceso a la información.

Considerar las normas internacionales de derechos humanos como el “equivalente a la Constitución” tiene muchas ventajas. También tiene ciertos límites que no se pueden pasar por alto. Entre estos uno de los más importantes es la libertad de la plataforma para desarrollar y proteger su negocio y adaptarlo a diferentes públicos. Por ejemplo: el entretenimiento para adultos no es ilegal y está protegido por las normas internacionales de derechos humanos. Si se aplica directamente el derecho internacional de derechos humanos, ninguna empresa podría prohibir la distribución o carga de pornografía o contenido violento en sus plataformas.

Sin embargo, cada empresa debe respetar y acatar los Principios Rectores sobre las Empresas y los Derechos Humanos de la ONU, que generan expectativas concretas en relación con el compromiso de la empresa con los derechos humanos, el deber de mitigar y prevenir las violaciones a estos, el deber de reparar, etc.⁸ Debe promoverse, desarrollarse y expandirse la

⁷ Naciones Unidas, Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, A/HRC/38/35 , 6 de abril de 2018, disponible en: <https://bit.ly/2KSgwR7>.

⁸ *UN Guiding Principles on Business and Human Rights* [Principios rectores de la ONU sobre empresas y derechos humanos], disponible en: <https://bit.ly/2sfNkgl>.

relación entre las normas internacionales de libertad de expresión y las reglas de contenido de Facebook. ¿Debería o podría el Consejo ser el órgano para cumplir con ese objetivo? La respuesta a esta pregunta probablemente contribuirá a definir los objetivos de este Consejo.

VI. Si fuera un órgano decisorio, ¿cuáles son los modelos que existen y qué se puede importar a esta nueva estructura: sistemas arbitrales; tribunales judiciales; tribunales internacionales; consejos de medios; y más?

Uno de los objetivos principales de la junta de supervisión de FB es brindar mayor legitimidad al sistema de moderación de contenido; pero crear un órgano lo suficientemente legítimo para esto es uno de los principales desafíos.

Si bien este consejo privado para supervisar y unificar las decisiones de moderación de contenido es el primero de su tipo, existen numerosos modelos diferentes que pueden y deben informar el proceso que está llevando a cabo FB. Primero, la compañía debe aprovechar las mejores prácticas y lecciones que se conocen gracias a más de 100 años de mecanismos decisorios internacionales de diferentes tipos si la idea es, de hecho, crear un órgano decisorio; en segundo lugar, estos organismos gozan de una legitimidad que amerita su influencia para determinar cómo seguir adelante con esta iniciativa; en tercer lugar, a diferencia de los organismos nacionales de toma de decisiones, los sistemas decisorios internacionales están especialmente diseñados para tratar las diferencias culturales y nacionales, habiendo adquirido una experiencia sobre el tema que debería reconocerse y de la cual se podría aprender.

Los mecanismos internacionales decisorios varían de un área a otra y en cada región. Hay varios modelos que se pueden tomar en cuenta, incluida la Corte Internacional de Justicia; los diferentes mecanismos de arbitraje que se crearon para ocuparse de los tratados bilaterales de inversión (por ej. el CIADI); el Comité de Derechos Humanos creado en virtud del Pacto Internacional de Derechos Civiles y Políticos; los diferentes mecanismos regionales decisorios en materia de derechos humanos (ej. el Tribunal Europeo de Derechos Humanos, la Corte Interamericana de Derechos Humanos y la Comisión Africana de Derechos Humanos y de los Pueblos). Estos son solo algunos ejemplos de organismos jurisdiccionales o cuasi jurisdiccionales internacionales existentes y que ya han tratado algunos de los problemas más complejos a los que la Junta de FB probablemente deberá enfrentarse, como ser: las diferencias culturales, idioma, nacionalidad, normas y estándares

globales o fuentes internacionales de derecho.

El diseño elegido para la Junta, por supuesto, determinará una serie de otras cosas, incluida la naturaleza y las facultades de sus miembros, la dinámica que se espera que tenga la Junta, la forma y las herramientas que tendrá para abordar la diversidad en todas sus formas (criterios, idioma, cultura jurídica, etc.). Los paneles de arbitraje, por ejemplo, son específicos para cada caso, diversos, compuestos por individuos seleccionados de una lista pre-existente de árbitros acreditados. Se basan en una secretaría fuerte para mantener el proceso en marcha y garantizar una cohesión procesal mínima y cierta memoria institucional que sea útil en procesos que a menudo son confidenciales. La capacidad de elegir a los árbitros ayuda a las partes a generar confianza en torno al proceso, brinda una oportunidad para que cada parte elija a un árbitro y el tercero generalmente es designado por la Secretaría. Esto también contribuye a garantizar cierta familiaridad entre los responsables de la toma de decisiones y los problemas, el contexto, el idioma y la cultura en que se ha presentado el caso. La diversidad del grupo de árbitros es particularmente relevante para estas estructuras. Aun así, la naturaleza descentralizada del propio órgano puede atentar contra la cohesión de las decisiones que surgen del organismo. El hecho de que el panel sea específico para cada caso puede ser problemático al tratar de definir a través de la interpretación el significado y las implicaciones de una norma internacional amplia o general.

Los tribunales internacionales como la Corte Interamericana de Derechos Humanos o el Tribunal Europeo de Derechos Humanos son radicalmente diferentes a los paneles de arbitraje. Cuentan con una membresía estable a través de los años, lo que garantiza cierta coherencia y seguridad jurídica en cuanto a los criterios de toma de decisiones. También garantiza un mayor nivel de igualdad para sus usuarios. La mayoría de las veces hay cohesión en sus decisiones y su naturaleza permanente proporciona mayor transparencia y responsabilidad del propio panel. La desventaja de estas estructuras es que no es posible garantizar la diversidad y la representación de todas las partes potenciales y es probable que estén más limitados en el número de casos que pueden tratar en un año determinado. La Corte Interamericana, por ejemplo, siempre opera de manera colegiada y con todos los miembros del tribunal. El Tribunal Europeo ha establecido un sistema mediante el cual los miembros del Tribunal ocupan diferentes salas, cada una con 3 jueces designados para ella. Algunas decisiones se elevan a la Gran Sala y, por lo tanto, garantizan la cohesión de su toma de decisiones de manera más amplia.

Como se muestra en los ejemplos breves, hay diferentes modelos potenciales que FB podría adoptar al diseñar su Consejo. Sin embargo, independientemente del que elijan, se deben consultar experiencias comparadas y anteriores y tenerse en cuenta al finalizar este proyecto.

Al tomar en cuenta estos tribunales y sus prácticas, debe prestarse especial atención a las lecciones aprendidas y las peores y mejores prácticas que se derivan de ellas. Uno de los muchos problemas que enfrentan estos tribunales es el retraso que generan. Los casos requieren mucho tiempo para ser litigados ante estos organismos, ya sea que requieran un abogado o no, y el acceso a los tribunales internacionales no es fácil. Entre las preguntas que FB debería hacerse una de las principales es cómo va a funcionar este “cuasi tribunal privado”. ¿Sumará un caso mas a la crisis sistemática de acceso a la justicia que impera a nivel global, o contribuirá a resolver, al menos en parte, ese problema? Las metas y expectativas para la Junta no son una cosa menor: si se trata de un tribunal de apelaciones, un tipo de respuesta a los usuarios, ¿cómo manejará este organismo a los 2,7 mil millones de usuarios y la cantidad de contenido que generan?

FB ya enfrenta críticas por sus procesos internos inmediatos para revisar las decisiones de la empresa sobre la moderación de contenido. Esto incluye no solo las decisiones de eliminación, sino también las decisiones de degradar o fomentar la circulación de ciertos contenidos en comparación con otros. Estos problemas no se resolverán con una junta de supervisión y, de hecho, la creación de una junta de supervisión no debería distraer la atención de ellos (mecanismos de apelación regulares dentro de la empresa), ya que son la base para cualquier posible reparación por prácticas incorrectas o injustas en la moderación de contenido.

Un enfoque similar al de la Corte Suprema (o Corte Constitucional o internacional) podría proporcionar una contribución mucho más sustancial y ser más apropiado con respecto a la escala. Sin embargo, debería haber mucha más claridad en cuanto a las expectativas, los criterios, el proceso, la selección de casos, la postura (en relación a las ONG, los usuarios, las organizaciones de consumidores), etc. para que esta junta reciba y seleccione casos, los trate y los haga públicos. Las decisiones que surjan de la Junta también deberían afectar la resolución de casos similares dentro de las operaciones regulares de moderación de contenido de FB y los procesos de apelación, convirtiendo a la Junta en una referencia interna y un órgano autorizado para interpretar las Condiciones de Servicio en relación con el derecho internacional de los derechos humanos. De lo contrario, y debido a problemas de acceso, escala y relevancia, el ejercicio pronto resultará irrelevante.

VII. ¿Quién formará parte de este Consejo? ¿Cuáles son los requisitos para ser miembros del Consejo?

Al pensar en los organismos decisorios, quién decide es tan importante como la forma en que lo hace. Sin embargo, nuestra pregunta no es, literalmente, quién estará en la junta, sino cuáles serán los criterios para seleccionar a esos miembros. La pregunta está relacionada de forma estrecha con las expectativas que se pueden tener para esta Junta. ¿Quién o qué está supervisando y para qué?

Como se explicó con anterioridad, los criterios de selección son fundamentales para garantizar la legitimidad. Aquí también los criterios de organismos comparables (en la medida de lo posible) y existentes deberían guiar los criterios de selección para los miembros de la Junta. El conocimiento y la experiencia en derecho internacional de los derechos humanos y en particular en libertad de expresión deben estar entre los requisitos de cualquier candidato. Esta recomendación no debe tomarse a la ligera. FB ya cuenta con consejos de seguridad y protección, compuestos por expertos externos y que participan activamente en el diseño de las políticas y las condiciones de servicio de la empresa. Analizar las decisiones sobre moderación de contenido y curación implica necesariamente un equilibrio entre la libertad de expresión y otros derechos.

La mayoría de los organismos decisorios internacionales de derechos humanos requieren que sus miembros cumplan con los requisitos para ser jueces en sus propios países y haber demostrado experiencia en derechos humanos. Si bien no todos los jueces tendrán la misma educación (algunos serán de países que aplican *Common law*, otros de tradiciones de derecho continental; algunos serán del norte, otros del sur, según la división Norte-Sur), todos tendrán algún tipo de educación jurídica; los miembros de La Comisión Interamericana o el Comité de las Naciones Unidas no necesitan ser abogados, pero sí necesitan cierta cantidad de años de experiencia, un conocimiento sólido del derecho internacional de los derechos humanos y altos estándares morales y éticos. Para estar en un panel de arbitraje, los árbitros deben cumplir con ciertos requisitos que incluyen ser versados en el derecho.

A diferencia de otras estructuras comunes a varias empresas diferentes, como la comisión de seguridad y confianza, o el grupo de seguridad para la infancia, la Junta de supervisión que se propone tiene la intención de lidiar con límites y restricciones al discurso. Si bien estos otros organismos están compuestos principalmente por expertos en temas de la infancia, poblaciones vulnerables, manejo de riesgos, violencia y abuso, es probable que

estos grupos no cuenten con muchos expertos en libertad de expresión. Las diferentes áreas requieren habilidades diferentes y una pregunta clave que FB debe plantearse al definir los criterios para la Junta de supervisión es cuál será su papel y cuáles deberían ser sus habilidades.

VIII. Conclusiones

En primer lugar, al diseñar su Junta de supervisión, FB y su equipo debería evaluar el impacto que dicho organismo va a tener en los derechos humanos de sus usuarios, en particular la libertad de expresión, el debido proceso, el acceso a la justicia, la igualdad y la no discriminación. También debería evaluar cómo esta Junta, independientemente de la estructura que tenga, contribuirá con la implementación de los Principios Rectores sobre las Empresas y los Derechos Humanos de la ONU.

La legitimidad es un elemento clave que tanto FB como sus usuarios esperan de la moderación de contenido. Sin embargo, se puede definir legitimidad de diferentes y variadas maneras, y entre los órganos de supervisión la legitimidad requiere ciertas características clave que no deben pasarse por alto para un órgano decisorio. La escala y la diversidad en este caso particular, con 2.700 millones de usuarios en 180 países, plantean desafíos adicionales y complejos para lograr legitimidad anclada en los modelos tradicionales de representación. Estos desafíos deben ser reconocidos y abordados de manera sistémica. FB no debe apuntar a algo que no puede proporcionar.

Continúa habiendo más preguntas que respuestas sobre la creación de la Junta de supervisión de FB. Al definir y responder estas preguntas, debería prestarse especial atención a los objetivos y expectativas para la Junta de supervisión. Después de muchas consultas, está claro que diferentes organizaciones y personas tendrán diferentes expectativas para este organismo. Es fundamental que la empresa aclare cuál es su postura sobre estos puntos y es importante proporcionar nuevos espacios para diálogo y comentarios después de que la empresa genere una propuesta concreta para su Junta y antes de que la implemente.

Si FB decide crear un modelo decisorio privado, debe tener en cuenta las prácticas recomendables y las lecciones aprendidas de las muchas y diversas estructuras que los mecanismos decisorios han adoptado a lo largo de los años, con resultados variables. Los mecanismos jurisdiccionales internacionales son particularmente relevantes, ya que son transfronterizos, sirven a una comunidad multicultural e interjurisdiccional, son tribunales y mecanismos de último recur-

so, y a menudo resuelven problemas jurídicos complejos e interrelacionados.

Siguiendo las obligaciones de FB en materia de derechos humanos, así como el marco aplicable a las restricciones de la expresión a nivel internacional, debería prestarse especial atención a la experiencia técnica requerida para los candidatos de la Junta. El desarrollo de criterios concretos, específicos y claros para la selección de miembros de la Junta es esencial para su fundación, independientemente de quién elija el primer grupo de miembros. Los criterios de selección dependerán de los objetivos y las expectativas concretas de esta Junta.

Acerca de CELE

El Centro de Estudios en Libertad de Expresión y Acceso a la Información es un centro de investigación afiliado a la Universidad de Palermo, en Buenos Aires, Argentina. El Centro proporciona investigación técnica jurídica para promover la comprensión y el desarrollo de la libertad de expresión y el acceso a la información, particularmente en América Latina. Desde 2012, contamos con una Iniciativa por la Libertad de Expresión en Internet (iLEI) en virtud de la cual hemos estudiado y producido investigaciones relacionadas con la libertad de expresión en internet, el acceso a la información y la privacidad, particularmente en el marco del sistema interamericano de derechos humanos. Nuestras estrategias para promover cambios incluyen la investigación; la capacitación; y la promoción de espacios de reflexión y debate de alto nivel. Visítenos en www.palermo.edu/cele y en www.observatoriolegislativocele.com.

Comentarios en respuesta a los cambios propuestos por Twitter a las reglas sobre “contenido deshumanizante”*

I. Introducción

El Centro de Estudios en Libertad de Expresión y Acceso a la Información (CELE) en la Universidad de Palermo preparó este documento en respuesta a la invitación de Twitter para comentar sobre una regla propuesta que aborda el “contenido deshumanizante”.

Desde CELE agradecemos la oportunidad de enviar comentarios y contribuir a mejorar el diálogo sobre las reglas que rigen los espacios como Twitter, que son fundamentales para el ejercicio actual de la libertad de expresión, la libertad de opinión y el acceso a la información. Los procesos de colaboración en la creación de estas reglas brindan una oportunidad para que los usuarios, la comunidad, la sociedad civil y cualquier parte interesada participen en la discusión, la configuración y el marco de las políticas que luego los afectarán.

II. De manera general, y algo que produce cierta preocupación: un enfoque cambiante a la libertad de expresión y discurso. Hacia “un diálogo saludable”.

El famoso objetivo principal de Twitter de proporcionar un espacio para “enfrentarse al poder” ha sido el motor detrás de la expansión de la empresa y el crecimiento cada vez mayor desde su inicio. Las políticas de la compañía sobre intervención limitada y libre flujo de información e ideas

* Este documento fue preparado por Agustina Del Campo, directora del Centro de Estudios en Libertad de Expresión y Acceso a la Información de la Universidad de Palermo. El documento es el resultado de la discusión y el análisis realizado en el marco del curso “Internet y Derechos Humanos”, impartido por la Prof. Del Campo en el Centro de Estudios de Posgrado de la Universidad de Palermo.

de todo tipo han contribuido a que sea un espacio único, particularmente en América Latina, donde no existen otras plataformas y empresas que ofrezcan un medio similar. Recientemente, sin embargo, el enfoque ha pasado de “enfrentar al poder” a contribuir a un “diálogo saludable”.¹ El cambio en la semántica parece particularmente importante para analizar las políticas propuestas y lo que podrían implicar. Y la compatibilidad de políticas que fomentan “diálogos saludables” con la libertad de expresión reconocida a nivel internacional y regional en las Américas probablemente dependerá del alcance y los medios para implementar estas políticas.

El derecho a la libertad de expresión tiene dimensiones individuales y sociales que contribuyen a su importancia decisiva para el desarrollo de la propia identidad, como individuo y como grupo; la libertad de expresión constituye un medio instrumental para el ejercicio de otros derechos (políticos, sociales, económicos y culturales); y es un requisito necesario para las sociedades democráticas.² La libertad de expresión protege no solo la expresión educada y políticamente correcta, sino también la expresión chocante, ofensiva o desagradable. La difusión de información y la expresión de todo tipo están relacionadas y vinculadas de forma indisoluble. Una restricción o interferencia en la difusión necesariamente afectaría el ejercicio de la libertad de expresión e impactaría de forma directa en la dimensión social de este derecho. Por lo tanto, queda implícito que una restricción o interferencia a uno de estos conceptos constituye necesariamente una violación del otro.³

Debido a su importancia tanto a nivel individual como social, los tratados internacionales de derechos humanos estipulan que las limitaciones a la libertad de expresión deben estar claramente establecidas por ley, de acuerdo con un objetivo legítimo, deben ser proporcionadas y necesarias en una sociedad democrática. El Relator Especial de las Naciones Unidas para la Libertad de Opinión y Expresión, David Kaye, en su informe de junio de 2018, recomendó que las empresas de Internet alinearan sus políticas de moderación con el derecho internacional de los derechos humanos y citó los Principios Rectores de la ONU sobre las Empresas y los Derechos Humanos como una fuente autorizada para dicha recomendación,⁴ a fin de fomentar y promover

¹ <https://twitter.com/jack/status/969234275420655616>.

² OEA, Relatoría Especial para la Libertad de Expresión, *Marco jurídico interamericano sobre el derecho a la libertad de expresión*, 2009.

³ Corte IDH, OC-5/85, disponible en: <https://bit.ly/2O2KoMp>.

⁴ Consejo de Derechos Humanos, Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y expresión, A/HRC/38/35, 6 de abril de 2018, disponible en: <https://bit.ly/2qvlHiN>.

la libertad de expresión en tiempos de espacios cada vez más privatizados.

La razón detrás de estas reglas y protecciones no radica en la necesidad de otorgar poder a quienes abusan de la expresión, sino en la necesidad de garantizar que no se excluya “*a priori*” ninguna voz de los debates sociales y políticos. Y aunque a veces las declaraciones impactantes o desagradables no parecen contribuir a los intercambios saludables, existen razones objetivas que sugieren lo contrario: 1) el gusto, la conmoción o la falta de respeto son conceptos subjetivos; 2) el discurso social es acalorado, diverso y apasionado, y no es quirúrgicamente limpio, ni probado académica o científicamente; 3) el discurso y la expresión no ocurren en el vacío y responden a diferentes culturas, contextos y entornos; 4) lo más importante es que una sola expresión no constituye una conversación, sino más bien una declaración que puede dar lugar a una conversación rica y relevante, cuyo beneficio queda por determinar.

Si bien entendemos que el cambio hacia la decisión de “fomentar diálogos saludables” puede estar motivado por un problema válido, nos preocupa que este cambio de política represente un vuelco considerable en la interpretación de libertad de opinión y expresión por parte de la empresa y el rol que juega para conseguir la plena realización de estos importantes derechos.

III. La política propuesta

“Política acerca de la deshumanización de Twitter”

“No se puede deshumanizar a nadie basándose en la pertenencia a un grupo identificable, ya que este lenguaje puede provocar daños fuera de Internet.”

Definiciones:

Deshumanización: Uso de lenguaje que trata a los demás como si fueran menos que humanos. La deshumanización puede ocurrir cuando a otros se les niega la calidad humana (deshumanización animal) o la naturaleza humana (deshumanización mecanicista). Los ejemplos pueden incluir la comparación de grupos con animales y virus (animalistas), o la reducción de grupos a una herramienta para algún otro propósito (mecanicistas).

Grupo identificable: cualquier grupo de personas que pueda distinguirse por sus características compartidas, como ser su raza, etnia, origen nacional, orientación sexual, género, identidad de género, afiliación religiosa, edad, discapacidad, enfermedad grave, ocupación, creencias políticas, ubicación, o prácticas sociales”.

1. ¿Qué constituye la “deshumanización” y cómo debe abordarse?

La política propuesta busca abordar el contenido “deshumanizante” que puede contribuir a la violencia. Argumentamos que, a pesar de la buena intención de los objetivos, la política puede ser demasiado amplia y general y podría convertirse en una herramienta para la censura discrecional. Todo, desde el contenido discriminatorio o inflamatorio hasta el contenido políticamente incorrecto, podría caer bajo la categoría de “contenido deshumanizante”.

Como declara el profesor Timothy Garton Ash, debe haber una clara distinción entre el discurso que incita a la violencia y, por lo tanto, causa daños; y el discurso irrespetuoso u ofensivo. “si no permitimos ningún discurso que ofenda a ninguna persona, va a quedar muy poco de lo que podamos hablar”.⁵

Desde la publicación del blog que anuncia el cambio, la nueva política busca: “expandir nuestra política relacionada con la conducta de odio para incluir contenido que deshumaniza a terceros en función de su pertenencia a un grupo identificable, incluso cuando el material no esté dirigido a alguien en particular”. Continúa estableciendo que “muchos académicos han examinado la relación entre la deshumanización y la violencia. Por ejemplo, Susan Benesch ha descrito el lenguaje deshumanizante como un sello distintivo del discurso de odio, porque puede hacer que la violencia parezca aceptable, y Herbert Kelman ha postulado que la deshumanización le reta fuerza al combate contra la violencia”.

Aun así, como explican los autores citados, qué constituye contenido deshumanizante puede depender del contexto y la cultura. Sin embargo, la política no hace referencia en absoluto al contexto ni a la pertinencia de la prohibición. Como se indica actualmente, la política propuesta elige centrarse selectivamente en lo que Benesch describe como un “sello distintivo del discurso peligroso”, pero no menciona otros aspectos importantes de su argumento: “El discurso peligroso no puede ser identificado únicamente por los sellos distintivos o por algún aspecto de su contenido, ya que su capacidad para inspirar violencia depende mucho del contexto: quién lo transmite, cómo, a quién y en qué contexto social e histórico”. Y que “Tales esfuerzos (para combatirlo) no deben vulnerar la libertad de expresión, ya que es un derecho fundamental, y cuando se impide a las personas “expresar sus quejas es menos probable que las resuelvan pacíficamente y es más

⁵ Tenenbaum, Tamara, “Timothy Garton Ash: Lanzar discursos de odio no debería ser gratis”, *La Nación*, 30 de septiembre de 2018, disponible en: <https://bit.ly/331BhzP>.

probable que recurran a la violencia”.⁶

Sobre la base del lenguaje de la política propuesta, el contenido “des-humanizante” estará prohibido en la plataforma. Esta prohibición absoluta sin excepciones equivale a una prohibición de dicho contenido en la plataforma. El impacto de esta política en los debates de interés público, en los coloquialismos y el habla puede ser incluso de mayor alcance y restrictivo si se busca de manera proactiva el contenido por medio de la inteligencia artificial y las decisiones a tomar se dejan a cargo de algoritmos.

Los ejemplos proporcionados en cuanto a lo que puede constituir contenido deshumanizante incluyen:



[Texto 1: No queremos más [miembros de un grupo religioso] en este país. ¡Ya es suficiente con esos GUSANOS!]

Texto 2: [Los / Las género] sólo sirven para el sexo]

Si se toma literalmente, según esta política, cualquier contenido que haga referencia a los fanáticos del equipo de fútbol argentino River Plate como “gallinas” podría interpretarse como deshumanizante; o cuando alguien se refiere a los mexicanos como “Guey” (que viene de la palabra “buey”) también podría interpretarse como deshumanizante. Al mismo tiempo, en algunas culturas, cuando una persona es extremadamente fuerte se la compara con un “toro” o “buey”, ¿qué se haría en este caso?

Un ejemplo más complejo sería el de los comentarios que se refieren a las mujeres como “potras” o a los hombres como “potros” (designación utilizada

⁶ <https://dangerousspeech.org/the-dangerous-speech-project-preventing-mass-violence/>.

en los años noventa para referirse al atractivo físico) que también se reportarían por violación de la política. En estos casos, aunque relacionar la belleza de una persona con un caballo puede no ser deseable, independientemente de la buena intención, probablemente tampoco debería prohibirse. El ejemplo podría ser aún más problemático si consideramos las referencias a las mujeres como “yeguas” (hembra del caballo), que de hecho tiene connotaciones bastante diferentes, es una denominación ofensiva e implica malvada o malvado. Mismo animal, misma “deshumanización”, diferentes palabras, usos, contextos, significados culturales, diferentes connotaciones. ¿Cabrían dentro de esta política? ¿Deberían prohibirse estas últimas referencias?

Pero incluso si el contenido es deshumanizante, dado que si se lo combinara con otros factores podría naturalizar la violencia y se lo podría excepcionalmente prohibir, mantenemos que tampoco se debería prohibir parte de este contenido y se lo debería preservar para el acceso público por varias razones que incluyen: la adjudicación de responsabilidades, el humor y lo que resulta más importante son las consecuencias positivas de dicha expresión: la réplica. Además, la excepción de interés periodístico debe ser explícitamente aplicable a esta política. Aquí enumeramos un par de ejemplos para ilustrar esto:

- 1) Tan recientemente como la semana pasada, en octubre de 2018, surgió la noticia de que a un funcionario importante del gobierno de Trump le “gustó” un meme que circulaba en Twitter sobre los Obama mirando un plátano; un meme claramente racista destinado a retratar a los afroamericanos como simios o monos. The Huffington Post publicó un artículo interesante al respecto y proporcionó más antecedentes sobre casos pasados de funcionarios dentro de la administración de Trump compartiendo, dándole “me gusta” o difundiendo contenido racista e intolerante.⁷ Este tipo de adjudicación de responsabilidad pública es fundamental para las sociedades democráticas. Es imprescindible que las personas se enteren de lo que sus representantes piensan y defienden, y más aún cuando es algo racista, segregacionista, sexista, etc. Las empresas no deben contribuir a “ocultar” estos casos, ya que estarían ayudando a fomentar la impunidad y la excepción de interés periodístico debe estar incluida explícitamente en la regla.
- 2) Memes y humor: una parte importante de la sátira y la parodia implica burlarse y ridiculizar elementos sociales que son comunes a nuestra

⁷ Kaufman, Alexander, “EPA Chief Andrew Wheeler Engaged With Racist, Conspiratorial Posts On Social Media”, The Huffington Post, 10 de noviembre de 2018, disponible en: <https://bit.ly/330267y>.

realidad, religión, situación política o estatus. Una imagen vale más que mil palabras y, con mucha frecuencia, estas ridiculizaciones implican recurrir a algún tipo de “deshumanización” tal como lo define esta política: ya sea centrándose en la naturaleza mecánica de lo que hace la persona o el grupo o utilizando animales para crear paralelos. Este es un ejemplo de una caricatura que fue judicializada en Colombia por comparar a los simpatizantes del expresidente Uribe con cerdos. El segundo dibujo muestra a los cerdos reclamando al artista por compararlos con políticos. El caso fue desestimado en la corte a favor del caricaturista, pero ¿qué hubiera ocurrido en Twitter bajo esta política?



- 3) Discurso de réplica: el ejemplo proviene de Argentina, donde a una famosa estrella del rock durante un ejercicio académico con un grupo de estudiantes de periodismo se le preguntó su opinión sobre las denuncias de abuso sexual y conducta inapropiada que involucraban a menores dentro de la escena nacional del rock and roll. Esta persona respondió, entre otras cosas, que «hay mujeres que necesitan ser violadas para disfrutar del sexo» y que «las jóvenes de 16 años no deben ser consideradas menores según esta ley en particular». Un estudiante furioso subió la entrevista a Facebook y despertó una gran presión para bloquearla. Sin embargo, la condena social y las réplicas que provocó la entrevista (reacciones que incluyeron la cancelación de recitales, debates nacionales en la televisión, críticas en la prensa y el clamor a nivel nacional sobre tales comentarios sexistas) fue tan importante y planteó con tanta eficacia el tema en la agenda que el contenido bloqueado hubiera significado poner restricciones innecesarias a debates saludables, productivos e interesantes y enfriar una conversación necesaria sobre la igualdad de género y el sexismo en nuestras sociedades.

Estos son solo algunos ejemplos para demostrar que este tipo de políticas no debe tomarse a la ligera, deben diseñarse y redactarse con atención y aplicarse con mucho cuidado para no infringir la libertad de expresión y reducir la capacidad de las personas para debatir las cosas importantes de nuestro tiempo. Estas discusiones suelen ser acaloradas, potencialmente ofensivas, impactantes para algunos y a veces muy polarizadas. Sin embargo, son necesarias y en tiempos de restricciones crecientes en las redes sociales, Twitter, que tiene una larga historia de protección de la expresión, debería proporcionar algo de espacio para estas discusiones.

2. ¿Cómo define la política a los “grupos vulnerables”?

Algunas empresas de Internet, al redactar sus políticas relacionadas con el discurso de odio, las amenazas o la violencia, crean sus propias definiciones de lo que constituye “grupos vulnerables”. La definición de quiénes son vulnerables en algunos casos se superpone con lo que tradicionalmente se encuentra bajo la ley de derechos humanos como grupos o poblaciones vulnerables, pero a menudo hay discrepancias significativas. Facebook, por ejemplo, considera que los funcionarios públicos y los jefes de estado son “grupos vulnerables” para ciertos tipos de contenido.

Aunque la política no incluye a estos grupos dentro de los grupos vulnerables que se pretende proteger, la enumeración tampoco lo excluye explícitamente. Si se interpreta que los funcionarios públicos, los políticos, etc. forman parte de un grupo vulnerable, la política podría impactar el discurso político de manera desproporcionada, generando restricciones injustificadas a la libertad de expresión. El humor político, la parodia y la sátira podrían sufrir un gran golpe en detrimento del debate de interés público y el intercambio democrático.

3. ¿Cómo se implementará esta política?

Debido a la relevancia que el contexto, la cultura y el quién, cómo y cuándo lo dice juegan en calificar el “contenido deshumanizante” que podría naturalizar la violencia o convertirse en un “sello distintivo del habla peligrosa”, Twitter no debería analizar el contenido de forma proactiva para detectarlo, sino que debería mantener la práctica tradicional de acudir a los “requisitos de reporte”. El uso de medios mecánicos o automatizados para monitorear, detectar y actuar frente a este tipo de contenido podría llevar a un bloqueo excesivo y/o una reacción exagerada sobre contenido que nadie encuentra ofensivo, o incluso si fuera ofensivo, no debería estar prohibido.

Si se adopta la política propuesta, debería implementarse cuando los usuarios reportan contenido. Al mismo tiempo, tal vez no debería bastar con un solo reporte. Una alternativa podría ser requerir un número mínimo de reportes para este tipo de contenido que no se refieran a una persona sino a grupos; y se debería requerir que el que realiza el reporte sea un miembro del grupo real que está “ofendido” por el contenido, o una organización que represente los intereses de dicho grupo. Por ejemplo, solo se debería cuestionar el contenido que se refiere a mujeres si está reportado por mujeres u organizaciones que representan a las mujeres.

Se debe prestar especial atención, como se indicó anteriormente, a la definición de grupos vulnerables y distinguir cuidadosamente a quién está dirigido el contenido, en lugar de aplicar la política en abstracto. El discurso de réplica de las minorías puede considerarse ofensivo para los miembros de grupos mayoritarios. El segundo ejemplo citado en el blog establece: “[Los/Las género] solo sirven para el sexo”. Al llenar los espacios en blanco, la frase no significaría lo mismo si pusiéramos «mujeres» que «hombres».

En todos los casos, la política debe interpretarse de manera restrictiva, prestando mucha atención al contexto, a la cultura y a otros aspectos del discurso en sí que podrían hacer que este tipo de contenido fuera “violento” o que incite a la violencia. La empresa debe realizar una supervisión adecuada de cómo se interpreta y aplica la política. Y debería haber alguna discusión sobre quién y cómo podría recurrir a una decisión restrictiva sobre el contenido referido a un grupo.

Finalmente, la transparencia es fundamental para cualquier moderación de contenido y debe haber mecanismos en juego para tener en cuenta las estadísticas sobre bloqueo de contenido, eliminación, suspensión de cuentas y cualquier otra medida adoptada en relación con esta política; así como transparencia en la forma en que se interpreta la política a través del tiempo para evitar decisiones excesivamente discrecionales que prohíben el contenido de la plataforma. El que ha generado el contenido debe recibir una explicación adecuada de por qué se eliminó el contenido y cómo se interpretó la política en su caso particular.

En el caso de esta política, la transparencia puede ser aún más importante, ya que contribuiría al objetivo establecido por la propia política. Si el objetivo es promover un “diálogo saludable”, las personas deberían entender claramente qué significa eso y cómo se comportan sus interacciones con respecto a la salud del intercambio. Twitter se encuentra en una posición privilegiada para implementar esta política de manera transparente y abierta, y así puede educar al mismo tiempo que la hace cumplir.

IV. Conclusión

La política propuesta responde claramente a un nuevo enfoque en la libertad de expresión y el acceso a la información que plantea serias preocupaciones con respecto a su compatibilidad con las normas internacionales y regionales de derechos humanos. Específicamente, la política propuesta es demasiado amplia y no establece en detalle las limitaciones que se impondrán, lo que permite una excesiva discreción en su interpretación e implementación.

Si se adopta, la política debe establecer claramente excepciones a la regla y aclaraciones sobre cómo se debe leer, interpretar e implementar y qué recursos se proporcionarán para cuestionar las decisiones. La transparencia es vital para garantizar buenas prácticas de moderación y, específicamente con respecto a esta política, para contribuir al objetivo general de la norma.

Agradecemos nuevamente a la empresa la oportunidad de enviar estos comentarios. Reiteramos nuestra opinión de que estos ejercicios son de la mayor importancia ya que colaboran en darle forma a las reglas que guiarán los futuros debates en Internet.