

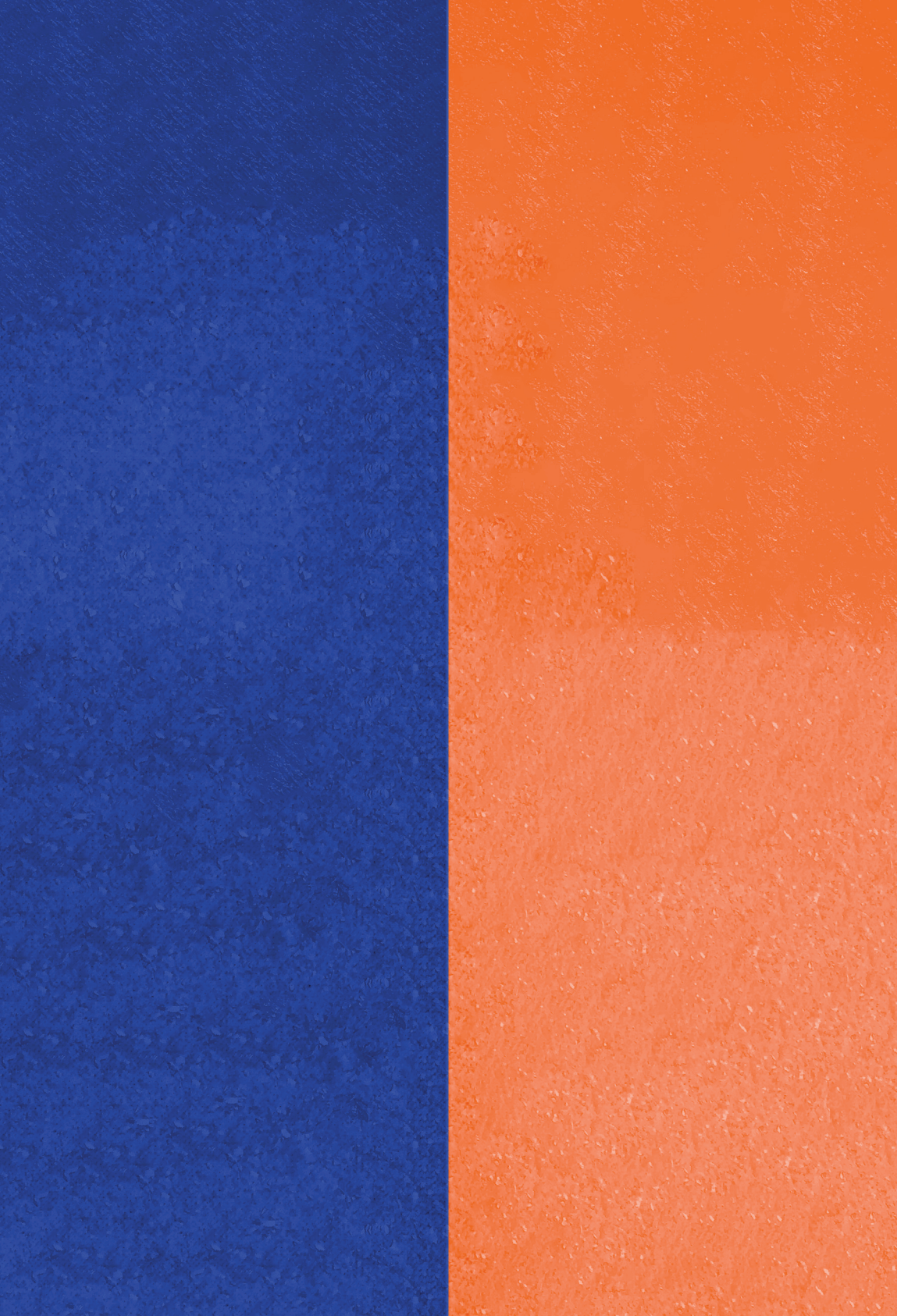
ARTICLE 19



Serie de normas internacionales

Vigilar a los vigilantes

Moderación de
contenidos, gobernanza y
libertad de expresión



Serie de normas internacionales

Vigilar a los vigilantes

Moderación de contenidos, gobernanza y libertad de expresión



ARTICLE 19 trabaja por un mundo en el que todas las personas, en cualquier lugar, puedan expresarse libremente y participar activamente en la vida pública sin temor a la discriminación. Toda nuestra labor se basa en dos libertades interrelacionadas. La libertad de expresión se refiere al derecho de toda persona a expresar y difundir opiniones, ideas e información por cualquier medio, así como a discrepar y cuestionar a quienes tienen el poder. La libertad de saber se refiere al derecho a exigir y recibir información de quienes detentan el poder para garantizar la transparencia, la buena gobernanza y el desarrollo sostenible. Cuando estas libertades se ven amenazadas por la incapacidad de quienes ostentan el poder para protegerlas adecuadamente, ARTICLE 19 habla con una sola voz, a través de los tribunales de justicia, las organizaciones internacionales y regionales, y de la sociedad civil dondequiera que estemos presentes.

✉ info@article19.org

🌐 www.article19.org

🐦 [@article19org](https://twitter.com/article19org)

📧 [@article19](https://www.instagram.com/article19)

🌐 www.linkedin.com/company/article19/

📘 facebook.com/article19org

978-1-910793-49-7

© ARTICLE 19, 2023

Esta obra se publica bajo la licencia Creative Commons Attribution-Non-Commercial-ShareAlike 3.0.

Se permite su reproducción, distribución y muestra, así como su utilización como base para otras obras, siempre que:

1. se reconozca la autoría de ARTICLE 19
2. no se utilice con fines comerciales
3. cualquier obra derivada de esta publicación se distribuya con arreglo a una licencia idéntica a esta.

Para acceder al texto legal completo de la licencia, véase:
<http://creativecommons.org/licenses/by-nc-sa/3.0/legalcode>

ARTICLE 19 agradecería recibir una copia de cualquier material en el que se utilice la información de este informe. ARTICLE 19 es el único responsable del contenido de este documento.

Todas las ilustraciones:
Mariana Coan
www.marianacoan.com

RESUMEN EJECUTIVO

En el presente informe de política, ARTICLE 19 describe su posición sobre la regulación de la moderación de contenidos por parte de las plataformas de redes sociales, de manera que proteja el derecho a la libertad de expresión e información.

Este documento de política se basa en nuestro trabajo anterior en este ámbito, en particular nuestras políticas sobre la responsabilidad de los intermediarios y sobre las directrices comunitarias o condiciones de servicio de las empresas. Anteriormente, nuestras propuestas se basaban en gran medida en lo que podría entenderse como una regulación mínima. Argumentábamos que las plataformas de redes sociales deberían seguir beneficiándose en gran medida de la exención de responsabilidad por el contenido de sus usuarios y usuarias. Esta importante norma no impide que las empresas rindan cuentas por no eliminar contenidos ilícitos. Este modelo se basa, asimismo, en que las empresas implementen condiciones de servicio compatibles con los estándares internacionales de derechos humanos. Hemos defendido este modelo porque creemos que protege mejor los derechos de expresión de los usuarios y usuarias. Es importante no solo para las plataformas de redes sociales, sino también para muchos otros actores de Internet que prestan servicios a nivel de infraestructura, como servicios de red de distribución de contenidos y registradores de nombres de dominio.

No obstante, en la actualidad, dada la magnitud de las plataformas de redes sociales más grandes y su sistemática incapacidad para responder de forma adecuada a las críticas que han recibido —desde el tratamiento de Facebook de la crisis de los rohingya en Myanmar hasta la incomprensible posición de YouTube con respecto al «discurso de odio» y los implacables ataques contra periodistas en Twitter— este paradigma parece insostenible. Asimismo, consideramos que la transparencia debe ser un requisito básico que impregne todos los ámbitos de actuación de las empresas y debe ir acompañada de una mayor rendición de cuentas y un compromiso con la protección de los derechos humanos.

En la presente política, ARTICLE 19 analiza si el modelo que hemos defendido hasta la fecha da respuesta a las críticas contra las plataformas de redes sociales más grandes. Examinamos las ventajas y los inconvenientes de las distintas propuestas que se han realizado recientemente en todo el mundo y, finalmente, presentamos nuestra posición revisada sobre la regulación de las plataformas.

Recomendaciones clave

1. Los Estados deben evitar regular innecesariamente la moderación de contenidos en línea.
2. Los principios generales de cualquier marco legislativo deben ser la transparencia, la rendición de cuentas y la protección de los derechos humanos.
3. Debe mantenerse la exención condicional de responsabilidad por los contenidos de terceros, pero es preciso aclarar su ámbito de aplicación y los procedimientos de notificación y actuación.

4. Se debe mantener la prohibición de supervisión general del contenido.
5. Todo marco legislativo debe tener un alcance estrictamente limitado. La regulación debe centrarse en el contenido ilícito, en lugar de en el contenido «lícito pero nocivo». Los servicios de mensajería privada y las organizaciones de noticias deben quedar fuera de su ámbito de aplicación. Las medidas no deben aplicarse extraterritorialmente.
6. Las obligaciones con arreglo a cualquier régimen legislativo deben estar claramente definidas. En particular, las obligaciones de garantizar la transparencia y un debido proceso interno.
7. Todos los reguladores deben ser independientes tanto en la ley como en la práctica.
8. Todo marco legislativo debe ser proporcionado.
9. Todo marco legislativo debe facilitar acceso a recursos efectivos.
10. Se debe exigir a las grandes plataformas que separen sus funciones de alojamiento y curación de contenidos, y que garanticen su interoperabilidad con otros servicios.

ÍNDICE

Introducción	11
Conceptos clave	17
Estándares internacionales de derechos humanos aplicables	22
Garantías del derecho a la libertad de expresión	22
Límites al derecho a la libertad de expresión	23
Empresas de redes sociales y libertad de expresión	23
Protección del derecho a la privacidad y el anonimato en Internet	29
Disyuntivas sobre la regulación y moderación de contenidos	31
Argumentos a favor de la regulación de las plataformas	31
Posición de ARTICLE 19 sobre los argumentos a favor de la regulación de las plataformas	36
Conceptos clave y características de la regulación sobre moderación de contenidos	45
Aspectos positivos de las propuestas legislativas sobre la regulación de la moderación de contenidos	47
Deficiencias de las propuestas normativas sobre la moderación de contenidos	49

Recomendaciones de ARTICLE 19	57
Recomendación 1	57
Recomendación 2	58
Recomendación 3	59
Recomendación 4	65
Recomendación 5	66
Recomendación 6	68
Recomendación 7	75
Recomendación 8	77
Recomendación 9	79
Recomendación 10	80
Notas a pie de página	82



Introducción

Introducción

En 1996, John Perry Barlow, uno de los padres de Internet, declaró la independencia del ciberespacio.¹ Su visión era una en la que «cualquiera, en cualquier sitio, pueda expresar sus creencias, sin importar lo singulares que sean, sin miedo a ser coaccionado mediante el silencio o el conformismo». Los gobiernos no eran bienvenidos. La gobernanza se derivaría de la «moral, de un progresista interés propio, y del bien común». Casi 25 años después, Internet se ve más como una serie de plataformas de redes sociales (particularmente Facebook, Google y Twitter), la censura está generalizada y los gobiernos están decididos a regularlo. ¿Qué ha pasado?

La percepción de la gente sobre el papel de las plataformas de redes sociales también ha cambiado. En sus inicios, en general, estas plataformas se consideraban como una poderosa fuerza positiva, que liberaba la libre expresión, permitía la conexión entre personas y encabezaba una revolución democrática en todo el mundo. Sin embargo, con el paso de los años, han pasado a ser vistas como un semillero de «discursos de odio», acoso, intimidación, teorías de conspiración y propaganda. De limitarse a «alojar»² contenidos de una manera relativamente neutral, ahora promueven activamente contenido de determinados terceros, o incluso producen el suyo propio. En el banquillo de los acusados se sienta la reiterada incapacidad de las plataformas de redes sociales más grandes para comprender y abordar las preocupaciones de sus usuarios y usuarias, y de los gobiernos, desde el escándalo de Cambridge Analytica hasta el fracaso de Facebook para eliminar contenidos de incitación al genocidio contra los rohingyas en Myanmar y la lucha de YouTube para suprimir el vídeo del ataque terrorista de Christchurch en Nueva Zelanda.

El gran poder que estas empresas ejercen sobre las personas y su dominio en varios mercados³ se hizo más evidente en enero de 2021, cuando las principales empresas de redes sociales, incluidas Twitter, Facebook y muchas otras, suspendieron las cuentas del entonces presidente de los Estados Unidos, Donald Trump, por la posibilidad de violencia como consecuencia de sus tuits incitando al asalto al Capitolio de EE. UU. a principios de ese mes.⁴ En el mismo mes, Twitter bloqueó la cuenta oficial de la embajada china en EE. UU., a raíz de una publicación que defendía las políticas del gobierno chino en la región occidental de Xinjiang, donde los críticos afirman que China está implicada en la esterilización forzada de mujeres pertenecientes a la minoría uigur.⁵ Aunque no es la primera vez que esto sucede, fue una notable demostración de poder por parte de las plataformas de redes sociales. Esto ha reavivado uno de los debates más tensos en los círculos de moderación de contenido, a saber, quién decide qué está permitido o no en el discurso público en Internet.

Se trata de un debate importante en el que los Estados deben participar. En la práctica, se ha echado gran parte de la culpa a los marcos legislativos que regulan la responsabilidad de los intermediarios de Internet. En muchos países, las redes sociales y otras plataformas digitales se han beneficiado de un exención amplia o condicional de la responsabilidad por el alojamiento de contenidos ilícitos.⁶ En los últimos años, estos marcos legislativos han sido objeto de un ataque sostenido por dar carta blanca a estas empresas y ayudarlas a aumentar sus ganancias gracias a algoritmos que promueven una relación adictiva con contenidos «extremistas» y «nocivos». Esta situación ha planteado la cuestión de si es necesaria una mayor regulación para controlar el poder de las empresas de redes sociales dominantes, hacer frente al contenido ilícito y nocivo, y garantizar un mayor control democrático de sus decisiones por parte del público en general. Recientemente, los gobiernos también han respondido con propuestas que tratan de someter a las plataformas de redes sociales al control de reguladores

similares a los de radiodifusión, o proponen que tengan un «deber de diligencia» para con sus usuarios y usuarias a fin de evitar el «daño» que pudiera causar el discurso de otras personas que utilizan la plataforma.⁷

Para quienes defienden la libertad de expresión, estas propuestas suscitan preguntas difíciles acerca de a quién debemos confiar la vigilancia de la expresión de los usuarios y usuarias. Aunque se solía considerar que las plataformas de redes sociales ofrecían un alto nivel de protección de la libertad de expresión, han restringido cada vez más sus normas comunitarias y con frecuencia silencian las voces de las minorías.⁸ Hasta la fecha, la transparencia y los mecanismos de resolución de controversias sobre la eliminación de contenido han sido insuficientes para permitir un escrutinio adecuado de sus acciones y ofrecer una reparación significativa a sus usuarios y usuarias. Por último, es cuestionable que un pequeño número de plataformas dominantes pueda tener tanto poder sobre lo que la gente puede ver sin una responsabilidad pública más directa.

Al mismo tiempo, la perspectiva de una nueva regulación de las plataformas es sumamente problemática. Muchas propuestas actuales se centran en realidad en la regulación de contenidos en Internet, es decir, en la regulación del discurso de los usuarios y usuarias, ya que, en la práctica, los Estados exigen que las empresas vigilen las comunicaciones humanas y decidan qué discurso es «ilícito» o «nocivo».⁹ Se trata de una situación muy problemática ya que únicamente los tribunales pueden determinar la ilegalidad y diferentes tipos de contenido pueden requerir distintos tipos de regulación: las soluciones que se utilizan para abordar el material de abuso infantil pueden no ser apropiadas para hacer frente a la desinformación o responder a cuestiones relacionadas con los derechos de autor.

En cualquier caso, ARTICLE 19 considera que las nuevas propuestas de regulación de las plataformas requieren un examen minucioso desde la perspectiva de la libertad de expresión. Para ello, los Estados deben identificar el método necesario y menos restrictivo para lograr una protección efectiva de cada uno de los objetivos que tradicionalmente se asignan a la regulación de los medios (como el pluralismo y la diversidad de la libertad de expresión), teniendo en cuenta la evolución y el papel de las plataformas digitales en la promoción y protección de los derechos humanos en línea, incluida la libertad de expresión.

La presente política lleva a cabo este tipo de análisis teniendo en cuenta los estándares internacionales en materia de libertad de expresión. Se basa en nuestra posición actual sobre la responsabilidad de los intermediarios,¹⁰ nuestra política sobre las normas internas o directrices comunitarias de las plataformas digitales y la libertad de expresión,¹¹ nuestras propuestas de un nuevo modelo de reglamento multisectorial (consejos de redes sociales)¹² y nuestro trabajo en defensa de las plataformas digitales. Al tiempo que evaluamos si las posiciones que hemos mantenido durante mucho tiempo sobre estos temas siguen siendo sostenibles, formulamos recomendaciones sobre las garantías mínimas que debe incluir el marco legislativo que rija las actividades de las plataformas de redes sociales.

El ámbito de aplicación de la presente política se centra en las plataformas de redes sociales «dominantes», a las que a veces nos referimos como «plataformas digitales». No obstante, somos muy conscientes del contexto más amplio de nuestra investigación y nos referimos a un grupo más amplio de actores en el ecosistema de Internet que incluye entre otros a proveedores de infraestructura de Internet (como las redes de distribución de equipos informáticos) y proveedores de otros servicios digitales (como la mensajería privada o los servicios en la nube).

La presente política se divide en cuatro partes:

- En primer lugar, establece algunos términos clave, como plataformas de redes sociales «dominantes», «moderación de contenidos» y «autorregulación»;
- En segundo lugar, describe los estándares aplicables a la protección de la libertad de expresión en Internet que deben guiar cualquier esfuerzo legislativo y político en este ámbito;
- En tercer lugar, establece los principales argumentos a favor de una mayor regulación de las políticas de moderación de contenidos de las plataformas y nuestra respuesta general a tales argumentos;¹³ y
- Finalmente, ARTICLE 19 formula recomendaciones sobre las garantías básicas que debe incluir cualquier nuevo marco legislativo en este ámbito a fin de proteger los derechos humanos.

La presente política complementa otras políticas de ARTICLE 19 relacionadas con la regulación de las plataformas, en particular nuestra propuesta sobre la separación de la curación de contenidos,¹⁴ la diversidad de los medios de comunicación en el ecosistema digital¹⁵ y las obligaciones de transmisión.¹⁶ Además, tenemos previsto publicar más posiciones políticas sobre otras cuestiones relacionadas con este complejo asunto.¹⁷

Conceptos clave

Conceptos clave

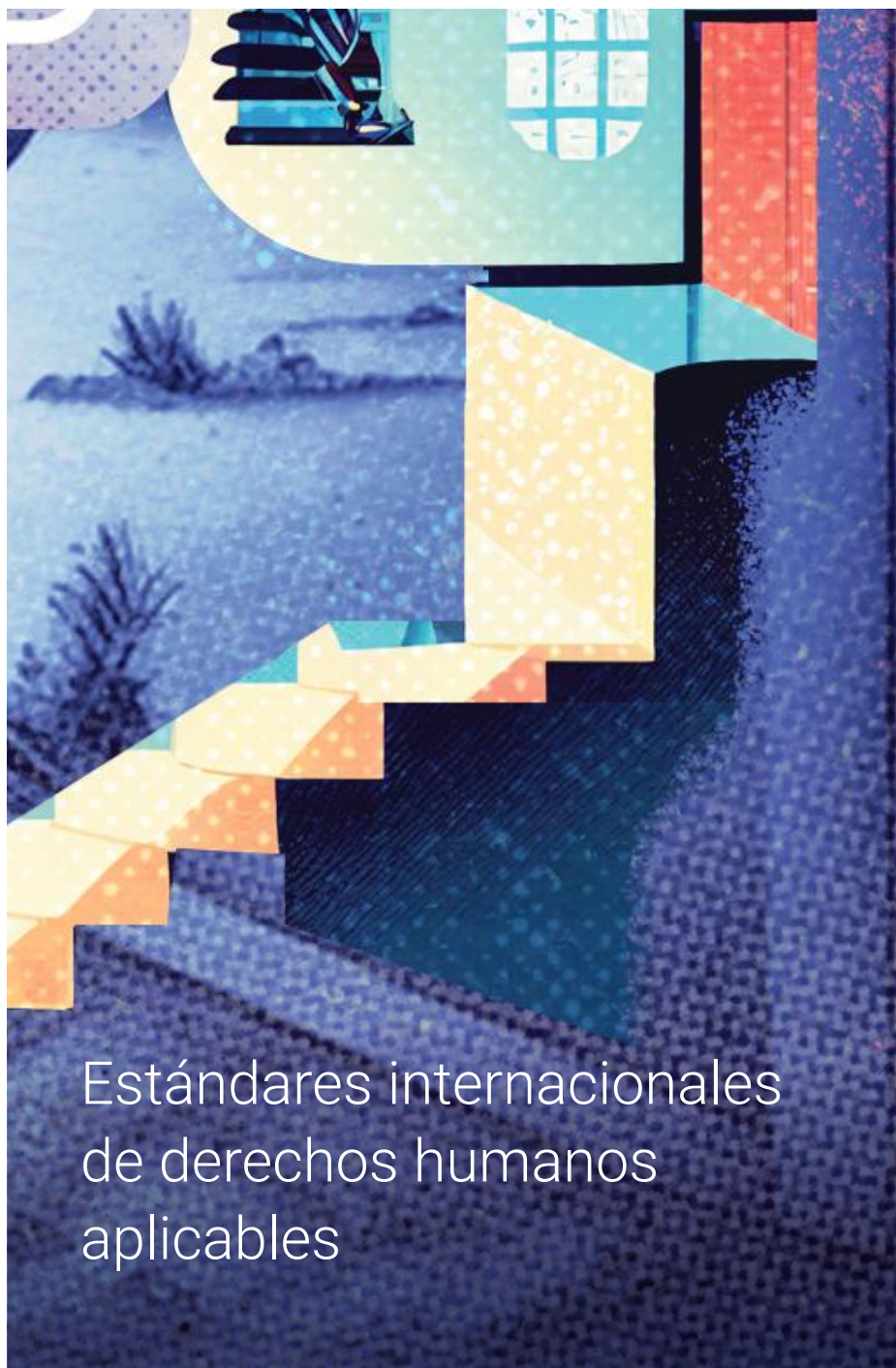
ARTICLE 19 ha propuesto previamente una tipología de intermediarios de Internet: proveedores de servicios de Internet, proveedores de alojamiento de páginas, plataformas de redes sociales y motores de búsqueda.¹⁸ Estos siguen siendo pertinentes hoy en día, pero dado que las prácticas de las plataformas de redes sociales han evolucionado y sus actividades han sido objeto de un mayor escrutinio, es importante definir algunos conceptos clave en este debate. A efectos de la presente política:

- Las **plataformas de redes sociales** son empresas que permiten a las personas y grupos conectarse e interactuar con otros usuarios y usuarias, y compartir contenidos utilizando redes de comunicación electrónica. Al hacerlo, prestan servicios de «alojamiento» a los usuarios y las usuarias, pero esto también suele ir acompañado de una variedad de otros servicios, como mensajería privada, «moderación de contenidos» y la selección de fuentes de noticias. Además, las plataformas de redes sociales suelen ofrecer espacios publicitarios y servicios conexos a los anunciantes, así como servicios de recomendación de noticias a los editores de noticias, conectando a estos últimos con los usuarios y usuarias. Además, pueden proporcionar interfaces de programación de aplicaciones (API) a los proveedores de aplicaciones, lo que les permite programar y desarrollar herramientas que se integran con la plataforma. Por esta razón, algunos académicos han identificado provisionalmente cuatro mercados pertinentes en los que suelen operar las plataformas de redes sociales: los servicios de redes sociales, la muestra de publicidad, la recomendación de noticias y las plataformas para aplicaciones.¹⁹

- El término **plataforma digital** se utiliza con frecuencia en la presente política en lugar de «plataformas de redes sociales», pero estas cubren un conjunto más amplio de actividades que las redes sociales. Pueden incluir mercados en línea, tiendas de aplicaciones, sistemas de pago, motores de búsqueda y plataformas para la economía colaborativa. Las plataformas digitales se caracterizan, entre otras cosas, por: 1) tener un modelo de negocio que se basa en gran medida en recopilar, procesar y editar grandes cantidades de datos; 2) operar en mercados multifacéticos; 3) beneficiarse de los efectos de red, que tienden a aumentar el valor del servicio con el número de usuarios y usuarias; y 4) depender de las tecnologías de la información y de la comunicación para llegar a estas personas de manera instantánea y sin coste alguno.²⁰
- La **moderación de contenidos** incluye los diferentes conjuntos de medidas y herramientas que las plataformas de redes sociales utilizan para gestionar el contenido ilícito y hacer cumplir sus normas comunitarias sobre los contenidos que los usuarios y usuarias generan en su servicio. Por lo general, implica la alerta por parte de los usuarios y usuarias, la utilización de alertadores fiables o «filtros», la eliminación, el etiquetado, la clasificación inferior o desmonetización del contenido, y la desactivación de ciertas funciones.
- La **curación de contenidos** es el uso que hacen las plataformas de redes sociales de sistemas automatizados para clasificar, promocionar o relegar contenidos en los canales de noticias, generalmente en función de los perfiles de sus usuarios y usuarias. También pueden promocionarse contenidos en las plataformas a cambio de un pago. Además, las plataformas pueden seleccionar contenidos utilizando anuncios intersticiales para advertir a los usuarios y usuarias sobre contenidos confidenciales o aplicar ciertas etiquetas para resaltar, por ejemplo, si el contenido proviene de una fuente fiable.

- La **autorregulación** es un marco que se basa completamente en el cumplimiento voluntario y en el que la legislación no desempeña ningún papel en la aplicación de las normas pertinentes. Su razón de ser es hacer que sus miembros rindan cuentas ante el público en general, promover el conocimiento entre sus miembros, desarrollar normas éticas y garantizar su respeto. Los modelos de autorregulación se basan principalmente en el entendimiento común por parte de los miembros de los valores y la ética que sustentan su conducta profesional. En el contexto de las empresas de redes sociales, el término «autorregulación» se utiliza para referirse a la «regulación por cuenta propia» (la regulación del discurso a través de condiciones de servicio o normas comunitarias) y la regulación sectorial (la gama de iniciativas que las plataformas de redes sociales adoptan para abordar problemas particulares, por ejemplo, la «desinformación»). En la presente política a veces también se utiliza el concepto de «autorregulación» para describir iniciativas gubernamentales en las que se presiona a las empresas para que adopten un conjunto de medidas bajo la supervisión del gobierno, pero cuyo incumplimiento no puede dar lugar a sanciones legales.²¹
- La **responsabilidad de los intermediarios**²² se refiere a las leyes con arreglo a las cuales los intermediarios de Internet pueden ser legalmente responsables del contenido difundido o creado por sus usuarios y usuarias. En muchos países, los intermediarios de Internet están *exentos* de responsabilidad (es decir, no se les pueden llevar ante los tribunales) siempre que eliminen el contenido cuando tengan conocimiento *cierto* de la ilegalidad, o un tribunal les ordene que lo hagan. La exención de responsabilidad generalmente se considera esencial para la supervivencia de los intermediarios de Internet, dada la magnitud del contenido que producen los usuarios y usuarias.

- La **responsabilidad de los editores**²³ se refiere a la responsabilidad jurídica de cualquier persona por el contenido que publica. En la práctica, significa que los editores están sujetos a las leyes de aplicación general, por ejemplo, la legislación sobre difamación, las leyes que impiden la publicación de secretos oficiales, etc. Los periódicos son un buen ejemplo. Son legalmente responsables por el contenido que producen en su sitio web o en forma impresa, incluidas las cartas a los editores. En la práctica, el contenido está sujeto a una revisión minuciosa y jurídica antes de la publicación, ya sea en línea o fuera de línea. Hoy en día, gracias a Internet, cualquiera puede ser editor. Por lo tanto, los usuarios y usuarias de Internet también pueden ser considerados responsables del contenido que producen.



Estándares internacionales de derechos humanos aplicables

Estándares internacionales de derechos humanos aplicables

Garantías del derecho a la libertad de expresión

El artículo 19 de la Declaración Universal de los Derechos Humanos protege el derecho a la libertad de expresión,²⁴ al que se otorga carácter jurídico vinculante a través del artículo 19 del Pacto Internacional de Derechos Civiles y Políticos²⁵ y de los tratados regionales.²⁶

El ámbito de aplicación del derecho a la libertad de expresión es amplio e incluye todos los modos de expresión electrónicos y de Internet. Exige que los Estados garanticen a todas las personas la libertad de buscar, recibir o difundir información o ideas de cualquier tipo, sin consideración de fronteras, a través de cualquier medio de su elección. Según los estándares internacionales de derechos humanos, el marco jurídico que regula los medios de comunicación debe tener en cuenta las diferencias entre los medios de comunicación impresos y audiovisuales e Internet,²⁷ y los sectores de las telecomunicaciones y la radiodifusión no pueden transferirse simplemente a Internet.²⁸ Los Estados deben adoptar un enfoque personalizado para abordar el contenido ilícito en línea y promover la autorregulación como una herramienta eficaz para reparar el daño causado por el discurso nocivo en línea.²⁹

Límites al derecho a la libertad de expresión

Según los estándares internacionales de derechos humanos, los Estados pueden limitar excepcionalmente la libertad de expresión, siempre y cuando cumplan los estrictos requisitos del llamado «principio de tres partes» o «tripartito», que requiere que las limitaciones:

- **Estén previstas por la ley:** toda ley o reglamento debe estar formulado con la suficiente precisión para que los individuos puedan regular su conducta en consecuencia;
- **Persigan un fin legítimo:** este debe estar detallado exhaustivamente. Por ejemplo, el respeto de los derechos o la reputación de terceros, la protección de la seguridad nacional o del orden público, o la protección de la salud o la moral públicas;
- **Sean necesarias y proporcionales en una sociedad democrática:** se exige que se aplique la medida menos intrusiva capaz de lograr el mismo objetivo.³⁰

Además, el artículo 20, apartado 2, del Pacto Internacional de Derechos Civiles y Políticos establece que toda apología del odio nacional, racial o religioso que constituya incitación a la discriminación, la hostilidad o la violencia debe estar prohibida por la ley. Los mismos principios se aplican a las formas electrónicas de comunicación o expresión difundidas a través de Internet.³¹

Empresas de redes sociales y libertad de expresión

Los organismos internacionales de derechos humanos han hecho observaciones sobre la relación entre la libertad de expresión y las empresas de redes sociales en varios ámbitos.

Responsabilidad del intermediario

Los mandatos especiales sobre la libertad de expresión han sostenido durante mucho tiempo que la exención de responsabilidad es la forma más eficaz de proteger la libertad de expresión en línea. Han recomendado que los intermediarios que prestan servicios técnicos no deben ser considerados responsables del contenido producido por otros y que únicamente se les debe exigir dicha responsabilidad si han intervenido específicamente en el contenido que se publica en línea.³²

Los sucesivos Relatores Especiales de la ONU sobre la promoción y protección del derecho a la libertad de opinión y expresión (Relatores Especiales sobre la Libertad de Expresión) han declarado en varias ocasiones que la censura nunca debe delegarse en una entidad privada, y que los Estados no deben hacer uso de intermediarios ni ejercer presión sobre ellos para que censuren en su nombre.³³ Además, David Kaye planteó que «[u]na normativa inteligente, y no una excesivamente estricta con una única óptica, debería ser la norma, una normativa centrada en garantizar la transparencia de las empresas y en la reparación».³⁴ Si bien no descartó la posibilidad de regulación bajo ciertas condiciones, reiteró que los Estados solo deberían tratar de restringir los contenidos por medio de una orden dictada por una autoridad judicial imparcial e independiente, y de acuerdo con las debidas garantías procesales y las normas establecidas de legalidad, necesidad y legitimidad.³⁵ Asimismo, señaló que los Estados deben abstenerse de imponer sanciones desproporcionadas, ya se trate de cuantiosas multas o de penas de prisión, a los intermediarios de Internet, dado su importante efecto disuasorio sobre la libertad de expresión.³⁶ Su sucesora y actual Relatora Especial sobre Libertad de Expresión, Irene Khan, ha llegado a una conclusión similar.³⁷

Responsabilidades del sector privado en materia de derechos humanos

Las recomendaciones sobre las responsabilidades de las empresas de redes sociales de respetar los derechos humanos incluyen (pero no se limitan a) los siguientes instrumentos:

- **Los Principios Rectores sobre las Empresas y los Derechos Humanos:** Puesta en práctica del marco de las Naciones Unidas para «proteger, respetar y remediar» (los Principios Rectores) proporcionan un punto de partida para articular el papel del sector privado en la protección de los derechos humanos en Internet.³⁸ Reconocen la responsabilidad de las empresas de respetar los derechos humanos, independientemente de las obligaciones del Estado o de la implementación de esas obligaciones, y recomiendan varias medidas que deben adoptar.³⁹ Entre ellas se incluye la incorporación de salvaguardas de derechos humanos por diseño para mitigar los impactos adversos, fomentar la influencia y actuar colectivamente para fortalecer su poder frente a las autoridades gubernamentales, y facilitar recursos cuando se creen efectos adversos sobre los derechos humanos.
- **Los titulares de los mandatos sobre la libertad de expresión** han abordado el papel de las plataformas de redes sociales en la promoción de la libertad de expresión y han recomendado que respeten y promuevan los Principios Rectores a este respecto. Entre otras cosas, han recomendado, que las empresas establezcan condiciones de servicio claras e inequívocas de conformidad con las normas y principios internacionales de derechos humanos,⁴⁰ que elaboren informes sobre la transparencia⁴¹ y proporcionen recursos efectivos para los usuarios afectados en caso de violaciones.⁴² Asimismo, han pedido repetidamente a los Estados que no exijan al sector privado que tome medidas que interfieran innecesaria o desproporcionadamente con la libertad de

expresión, ya sea a través de leyes, políticas o medios extralegales, ya que generalmente no están preparadas para determinar si el contenido es ilícito.⁴³ Asimismo, el Relator Especial sobre la Libertad de Expresión ha dejado claro que las empresas deben adoptar enfoques en materia de transparencia radicalmente distintos en todas las etapas de sus operaciones y deben abrirse a la rendición de cuentas, sugiriendo que esto podría tomar la forma de consejos de redes sociales.⁴⁴

- A nivel regional:
 - El Comisario para los Derechos Humanos del Consejo de Europa ha recomendado que los Estados dejen de depender de las empresas de Internet para imponer restricciones que violen las obligaciones de los Estados en materia de derechos humanos,⁴⁵ y que se elaboren más orientaciones sobre las responsabilidades de las empresas en relación con sus actividades en Internet.⁴⁶ Asimismo, el Comité de ministros del Consejo de Europa ha recomendado que las empresas de redes sociales respeten los derechos humanos y el Estado de derecho, en particular, las garantías procesales,⁴⁷ y sean transparentes sobre el uso de técnicas de tratamiento automatizado de datos.⁴⁸
 - La Organización de Estados Americanos (OEA) ha elaborado varios informes sobre la libertad en Internet con recomendaciones similares.⁴⁹ Por ejemplo, la OEA ha enfatizado que los intermediarios siguen siendo entidades privadas, con intereses que difieren de los del Estado, y que exigirles que actúen como un tribunal va más allá de su ámbito de competencia y puede generar incentivos para el abuso.⁵⁰

- La Comisión Africana de Derechos Humanos y de los Pueblos ha adoptado recientemente la Declaración de Principios revisada sobre la Libertad de Expresión y el Acceso a la Información en África.⁵¹ El principio 39 de la Declaración establece que los Estados deben exigir a los intermediarios de Internet que garanticen que, al moderar y filtrar el contenido en línea, incorporen salvaguardias de derechos humanos en sus procesos, adopten estrategias de mitigación para abordar todas las restricciones a la libertad de expresión y el acceso a la información en línea, garanticen la transparencia de todas las solicitudes de eliminación de contenido, incorporen mecanismos de apelación y proporcionen recursos efectivos cuando se produzcan violaciones de derechos.⁵²
- La **sociedad civil** también ha formulado recomendaciones para que las empresas de redes sociales respeten las normas internacionales de derechos humanos. Por ejemplo, los Principios de Manila sobre la Responsabilidad de los Intermediarios detallan los tipos de medidas que las empresas deben tomar para respetar los derechos humanos.⁵³ En particular, las prácticas de restricción de contenido de las empresas deben cumplir los principios de necesidad y proporcionalidad en virtud de legislación sobre derechos humanos⁵⁴ y proporcionar a los usuarios y usuarias mecanismos de reclamaciones para cuestionar las decisiones de las empresas.⁵⁵ El proyecto «Ranking Digital Rights» evalúa el cumplimiento de los indicadores de derechos digitales de las principales empresas de Internet, que incluyen, entre ellos, la disponibilidad de las condiciones del servicio, los motivos para restringir el contenido, las cuentas o el servicio, la notificación a los usuarios y usuarias, un proceso para responder a solicitudes de terceros y diversas actividades relacionadas con la transparencia.⁵⁶

Principios sobre contenidos específicos

Los titulares de los mandatos de derechos humanos han publicado una serie de declaraciones conjuntas que destacan las responsabilidades de los Estados y las empresas en relación con contenidos específicos, como el «extremismo violento»,⁵⁷ las «noticias falsas»,⁵⁸ el acoso y abuso en línea por motivos de género⁵⁹ y el «discurso de odio».⁶⁰ En estas, han llamado a los Estados a no someter a las plataformas digitales a órdenes obligatorias para retirar o limitar de otra manera el contenido, excepto cuando el contenido se limite legalmente con arreglo a los estándares internacionales.⁶¹ Asimismo han pedido a las empresas que garanticen que sus usuarios y usuarias puedan acceder y comprender fácilmente sus políticas y prácticas, cómo se aplican y cómo respetan las directrices mínimas del debido proceso⁶² y que garanticen que estas políticas estén directamente vinculadas con el derecho internacional de los derechos humanos.⁶³

Papel de la inteligencia artificial en la moderación de contenidos

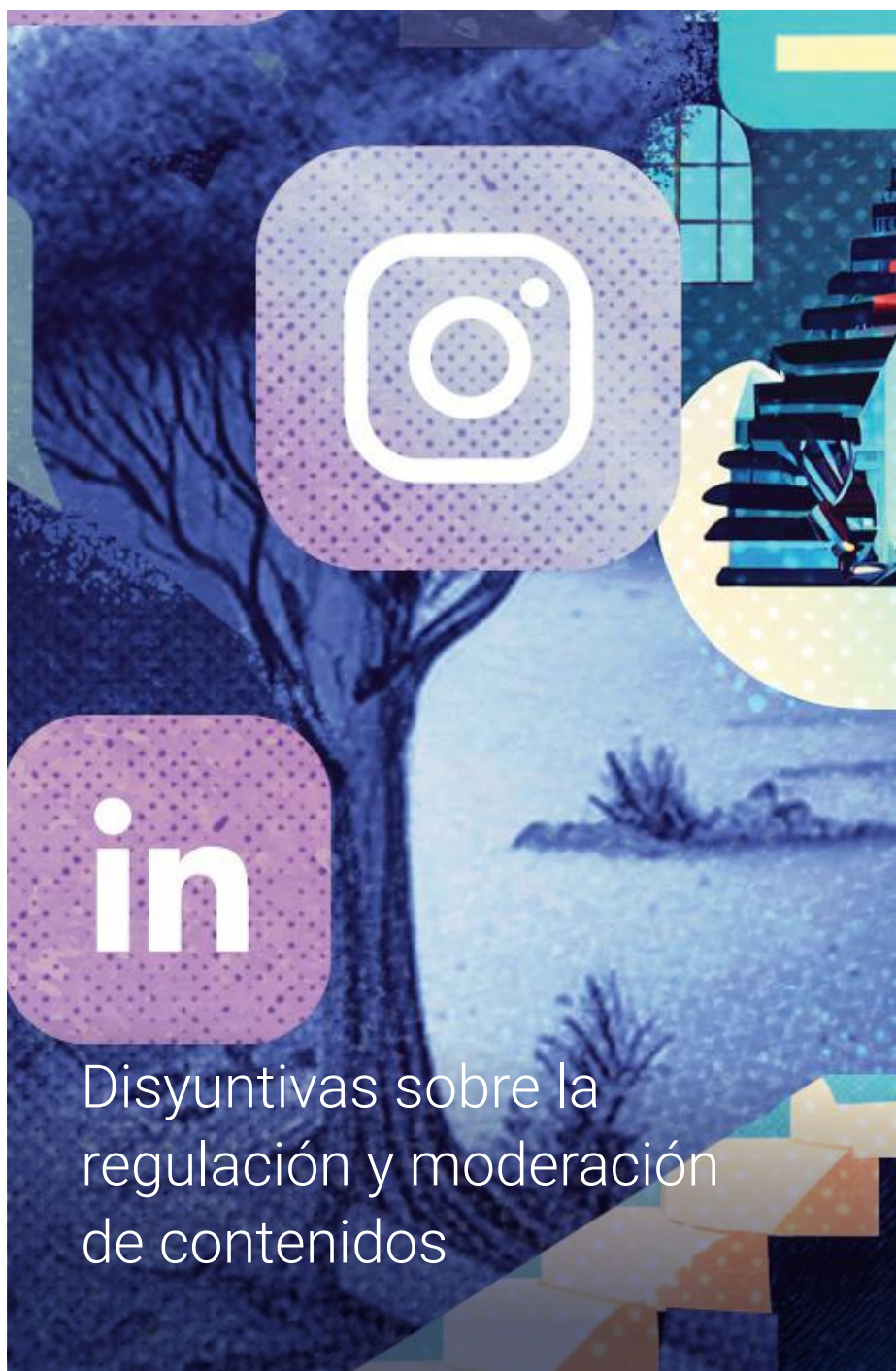
En agosto de 2018, en un informe a la Asamblea General, el Relator Especial sobre la Libertad de Expresión formuló una serie de recomendaciones sobre el papel de la inteligencia artificial (IA) en la moderación de contenidos y sus posibles repercusiones para los derechos humanos.⁶⁴ En particular, recomendó que los Estados crearan un marco normativo y legislativo propicio para un entorno de información diverso y pluralista en el campo de la IA.⁶⁵ Estas medidas podían incluir la regulación de los monopolios tecnológicos para prevenir la concentración de pericia y poder en materia de IA en manos de unas pocas empresas dominantes, la regulación encaminada a aumentar la interoperabilidad de servicios y tecnologías.⁶⁶

En el caso de las empresas, recomendó, entre otras cosas, que debía hacer explícitos dónde y cómo se usan en sus plataformas, servicios y aplicaciones las tecnologías de inteligencia artificial;⁶⁷ publicar datos sobre la supresión de contenido, estudios de casos y educación sobre el perfilado comercial y político;⁶⁸ y facilitar a los usuarios y usuarias medidas correctivas de los efectos adversos de los sistemas de IA en los derechos humanos.⁶⁹

Protección del derecho a la privacidad y el anonimato en Internet

Para que las personas tengan la confianza para ejercer libremente su derecho a la libertad de expresión, es fundamental garantizar el derecho a la privacidad en las comunicaciones en línea.⁷⁰ La incapacidad de comunicarse en privado afecta sustancialmente al derecho a la libertad de expresión de las personas.

David Kaye, el Relator Especial sobre la Libertad de Expresión, ha reconocido esta situación en varios informes, en los que ha expresado su preocupación con respecto a los Estados y actores privados que monitorean y recopilan información sobre las comunicaciones y actividades de las personas en Internet. Estas prácticas pueden constituir una violación del derecho a la privacidad de los usuarios y usuarias de Internet y, en última instancia, impedir el libre flujo de información e ideas en línea.⁷¹ El Relator también ha recomendado que los Estados velen por que las personas puedan expresarse anónimamente en línea y que se abstengan de adoptar sistemas de registro de los nombres reales.⁷² Además, recomendó que los Estados se abstengan de establecer la identificación de los usuarios como condición para acceder a las comunicaciones digitales y a los servicios en línea, y de obligar a los usuarios de teléfonos móviles a registrar su tarjeta SIM;⁷³ y que los actores empresariales consideren asimismo sus propias políticas que restringen el cifrado y el anonimato (entre otras cosas, mediante el uso de seudónimos).⁷⁴



Disyuntivas sobre la regulación y moderación de contenidos

Disyuntivas sobre la regulación y moderación de contenidos

Como se ha señalado anteriormente, en los últimos años ha habido un mayor impulso para que los Estados regulen las actividades de moderación de contenido de las plataformas digitales consideradas «dominantes» en el mercado de redes sociales. En la práctica, estos llamamientos a la regulación de las plataformas suelen traducirse en la regulación de contenidos en línea, aunque los límites a menudo se difuminan en el discurso público.

En la presente sección, ARTICLE 19 examina algunas de las principales razones por las que se percibe la necesidad de que los reguladores intervengan y, en particular, de que adopten un marco general para la regulación del contenido en línea. A continuación, procedemos a examinar si estas razones resisten un análisis riguroso.

Argumentos a favor de la regulación de las plataformas

Poder de la plataforma y poder de mercado

En general, el factor más importante que motiva el impulso actual en favor de la regulación es el poder de mercado de un pequeño número de plataformas de redes sociales, que hace que controlen cómo una gran cantidad de personas ejercen su derecho a la libertad de expresión en línea. El tamaño y los ingresos de estas plataformas suelen ser mayores que los de algunos países.⁷⁵ Desde el punto de vista de la libertad de expresión y la democracia, la mayor preocupación es que solo un puñado de empresas internacionales, principalmente con sede en EE. UU., deciden qué información pueden ver, acceder y compartir los usuarios y usuarias.

En la práctica, la preocupación por la concentración de poder se ha traducido en gran medida en una serie de propuestas que pondrían a las plataformas de redes sociales dominantes bajo el control de un regulador similar al de la radiodifusión. Por lo tanto, el enfoque de estas propuestas se centra principalmente en la regulación del contenido en línea y no abordan directamente el poder de mercado.⁷⁶ Algunos elementos de estas propuestas tienen también en cuenta el tamaño de las plataformas más grandes, por ejemplo, en referencia a su número de usuarios y usuarias, y el volumen de negocios.⁷⁷

Otras propuestas que implican la difusión de contenidos tienen el objetivo más explícito de garantizar la igualdad de condiciones entre los distintos tipos de industrias.⁷⁸ Además, los reguladores en materia de competencia y radiodifusión han comenzado a elaborar sus propias recomendaciones para hacer frente a las deficiencias del mercado en lo que respecta a la prestación de servicios en línea, por ejemplo, mediante el desarrollo de intervenciones reguladoras efectivas y la prevención de las consecuencias negativas no intencionadas.⁷⁹

Por último, cabe señalar que una serie de medidas relacionadas con la protección de datos, que abarcan desde la portabilidad de los datos hasta el consentimiento y la limitación de la finalidad, inciden en la competitividad de las plataformas más allá que la protección de los datos personales de los usuarios y usuarias. Por lo tanto, a fin de mantener bajo control el poder de estas gigantescas empresas y poner límites a sus modelos comerciales extractivos, es fundamental la adopción de marcos sólidos de protección de datos, como el Reglamento General de Protección de Datos.

Aunque la mayoría de las propuestas de un nuevo marco legislativo están orientadas directamente a limitar el poder excesivo de las mayores plataformas de redes sociales de EE. UU., también se han utilizado como una oportunidad para regular a los actores más pequeños y a otros servicios, como

la mensajería, el alojamiento en la nube y los servicios de intercambio de archivos. De hecho, algunos proveedores de infraestructura participan cada vez más en el debate sobre la moderación de contenidos, en particular acerca de la prestación de sus servicios a grupos neonazis y de extrema derecha.⁸⁰

Modelo de negocio de las plataformas

La segunda preocupación clave subyacente al actual impulso a favor de una mayor regulación del contenido en línea es el modelo de negocio de las plataformas de redes sociales más grandes, basado en la recopilación de grandes cantidades de datos sobre los usuarios y usuarias y sus hábitos en línea (datos sobre el comportamiento), y su monetización a través de la publicidad personalizada en línea.⁸¹ Los defensores de la privacidad y los derechos digitales, en particular, han alertado desde hace mucho tiempo de que el «capitalismo de la vigilancia» es manipulador y perjudicial los derechos humanos de los usuarios y usuarias.⁸² Se basa en técnicas que suponen una gran injerencia en la privacidad (por ejemplo, el seguimiento en línea, las ofertas en tiempo real y la personalización) y que son sumamente opacas.⁸³ Estas técnicas buscan mantener a los usuarios y usuarias interesados en las plataformas y fomentar la difusión generalizada de ciberanzuelos y contenidos sensacionalistas y «extremistas» en línea, que no permite a dichos usuarios y usuarias comprender por qué ven un tipo particular de contenido o anuncio en las redes sociales.⁸⁴ Por tanto, inciden en la autodeterminación informativa de los usuarios y usuarias. Además, tienden a promover una menor diversidad y calidad de contenidos,⁸⁵ y presentan riesgos significativos de discriminación en ámbitos como los anuncios de vivienda y empleo, entre otros.⁸⁶

Percepción de deficiencias en la autorregulación y la regulación por cuenta propia, y ausencia de un control democrático

Otro factor clave que ha impulsado las propuestas normativas actuales es la percepción de que las principales plataformas de redes sociales no han abordado una serie de aspectos «nocivos», que van desde la «desinformación» y el «discurso de odio»⁸⁷ hasta el terrorismo y las imágenes de abuso infantil. En particular, los críticos señalan la disponibilidad generalizada de material ilícito o «nocivo» en línea —incluido su impacto a veces trágico fuera de línea—⁸⁸ y la ausencia de un planteamiento coherente y transparente de las plataformas de redes sociales para la eliminación de contenidos.

Gran parte de la culpa se ha atribuido al modelo predominante de exención condicional de responsabilidad, que se considera que ha dado a las plataformas carta blanca. Se considera que se benefician de los males de la sociedad, que ellas mismas permiten, sin ninguna responsabilidad pública. Por consiguiente, las diversas iniciativas que las grandes plataformas de redes sociales han adoptado por propia voluntad propia⁸⁹ o en respuesta a códigos de práctica «autorreguladores»,⁹⁰ se consideran insuficientes tanto para abordar estas cuestiones como para resolver el problema de la falta de rendición de cuentas.⁹¹

Necesidad de garantizar la seguridad jurídica y una transparencia real

Ante los desafíos descritos anteriormente y la gran cantidad de leyes diferentes para hacerles frente, algunos gobiernos han argumentado que un marco legislativo general aportaría más claridad y seguridad jurídica.⁹²

Además, se considera que estos marcos son necesarios para obligar a las empresas de redes sociales a ser más transparentes sobre sus operaciones de moderación de contenidos y cómo funcionan sus algoritmos. Más allá de las estadísticas, sigue sin estar muy claro qué contenido se elimina,

por qué motivos y por qué los usuarios y usuarias pueden ver determinado tipo de contenidos. Con arreglo a estos nuevos marcos, las empresas también estarían obligadas a introducir mecanismos de recurso por la eliminación ilícita de contenidos. Un nuevo regulador, ajeno a la toma de decisiones sobre el contenido, supervisaría estas nuevas obligaciones.⁹³

Falta de diversidad de los medios de comunicación

Por último, otra posible razón para que los reguladores intervengan es la protección del pluralismo y la diversidad de los medios de comunicación en el ecosistema digital. Algunos legisladores ya han presentado propuestas para remediar lo que perciben como un fracaso de las plataformas de redes sociales en este ámbito.⁹⁴ Esto suscita preocupación por tres razones principales:

- **Concentración excesiva en el mercado de redes sociales y obstáculos:** al concentrar la mayor cantidad de usuarios y usuarias, un número muy limitado de grandes plataformas puede decidir cómo se distribuyen las noticias en su servicio y los criterios que se aplican para dicha distribución, y las empresas de medios de comunicación no tienen más remedio que colaborar con ellas, en gran medida en sus términos, para lograr la atención de estos usuarios y usuarias.⁹⁵
- **Contenido personalizado:** el modelo de las plataformas de redes sociales implica que sus usuarios y usuarias solo están expuestos a contenido altamente personalizado, en lugar de a una diversidad de puntos de vista,⁹⁶ lo que ha provocado temor por el uso de burbujas de filtro⁹⁷ y acusaciones de parcialidad por parte de algunos partidos políticos.⁹⁸
- **Dominio de la publicidad en línea:** como consecuencia de la publicidad en línea, los medios de comunicación tradicionales no solo han sufrido una pérdida de audiencia debido a las fuentes en línea y las redes sociales «gratuitas», sino también una pérdida de ingresos publicitarios. El problema de la sostenibilidad de los medios de comunicación también supone una amenaza para su pluralismo y diversidad.

Posición de ARTICLE 19 sobre los argumentos a favor de la regulación de las plataformas

ARTICLE 19 reconoce que las preocupaciones tras el impulso a favor de una mayor regulación de las prácticas de las plataformas son completamente válidas. De hecho, las compartimos. Sin embargo, creemos que es probable que muchas de las soluciones que se proponen actualmente no logren su objetivo, afiancen el «dominio» de los actores más importantes, y puedan ser objeto de abuso por parte de los gobiernos al darles más control sobre las plataformas y el contenido. A continuación, presentamos algunas soluciones que, creemos, garantizarían mejor la protección de la libertad de expresión.

La concentración de poder en manos de unas pocas grandes plataformas debe abordarse principalmente a través de instrumentos que favorezcan la competencia

En opinión de ARTICLE 19, los avances más recientes en materia de «regulación de las plataformas» se refieren a la complejidad de los desafíos que plantean las grandes plataformas digitales y el alcance de la interrelación entre varias disciplinas, desde la regulación de contenido en línea hasta la competencia, la protección del consumidor y la protección de datos. Esto demuestra, asimismo, que cualquier intervención normativa debe ser clara sobre sus objetivos y los «efectos nocivos» que pretende abordar. En particular, ARTICLE 19 cree que las preocupaciones sobre el tamaño, el poder o el «dominio» de las empresas de redes sociales deben abordarse principalmente mediante instrumentos de competencia y marcos económicos favorables a la competencia, en lugar de a través de la regulación del contenido en línea.

Al mismo tiempo, reconocemos que algunas empresas de redes sociales se han vuelto tan dominantes que puede ser apropiado adoptar enfoques escalonados. Mientras que los actores más grandes generalmente podrán cumplir con las nuevas obligaciones normativas (por ejemplo, sobre transparencia, debido proceso o incluso objetivos de eliminación de contenidos), los actores pequeños simplemente quebrarían tratando de cumplirlas. En otras palabras, es vital no cargar indebidamente a las empresas más pequeñas con pesadas obligaciones normativas, de lo contrario, no podrán competir con las empresas más grandes. Más adelante, formulamos recomendaciones sobre la forma de abordar esta cuestión.

No obstante, señalamos que los enfoques escalonados deben concebirse de manera que eviten el efecto no deseado de consolidar la posición de los actores establecidos, ya que estos pueden hacer frente a los nuevos requisitos normativos (por ejemplo, seguridad y fiabilidad) que harían que sus servicios sean más atractivos para los usuarios y usuarias.⁹⁹ Asimismo, observamos que los enfoques escalonados pueden tener el efecto no deseado de desplazar a los «malos actores» a las plataformas menos vigiladas.¹⁰⁰ Por tanto, seguirían operando al margen, con menos visibilidad para el público en general pero más visibilidad entre ellas.

No se ha demostrado la efectividad de enfoques reglamentarios más estrictos

ARTICLE 19 también considera que no está claro que los nuevos reglamentos estatales sobre la moderación de contenidos por parte de las plataformas sean estrictamente necesarios para abordar el contenido problemático en línea, como el «discurso de odio» o el terrorismo.

Nos preocupa que la nueva legislación no siempre respeta el derecho internacional de los derechos humanos.¹⁰¹ Además, a menudo no existen pruebas suficientes de la eficacia de las medidas propuestas, ya sea para combatir cuestiones específicas (como el «extremismo» o el «discurso de odio»)¹⁰² o de su daño colateral a la libertad de expresión.¹⁰³ Con demasiada frecuencia, los legisladores se esfuerzan por adoptar leyes para enviar un mensaje político al público de que «se está haciendo algo» para abordar un problema, en lugar de invertir recursos en soluciones a largo plazo que son menos visibles pero más efectivas.

Por todas estas razones, ARTICLE 19 sigue siendo muy escéptico en cuanto a que el tipo de nuevas leyes que se han presentado puedan ofrecer soluciones para contrarrestar el contenido problemático. Es más probable que resulten contraproducentes, estén orientadas al objetivo equivocado y sean perjudiciales para la libertad de expresión.¹⁰⁴

En cambio, deben establecerse modelos de gobernanza independientes o multisectoriales

ARTICLE 19 también reconoce que las empresas de redes sociales, en general, han tardado en reaccionar a una serie de preocupaciones legítimas —inquietudes sobre el «discurso de odio», la «desinformación» y la protección de la infancia— con respecto a sus plataformas. Con el tiempo, esta situación ha erosionado considerablemente la confianza pública en ellas y ha suscitado dudas legítimas sobre su falta de responsabilidad pública.

No obstante, en nuestra opinión, es demasiado pronto para descartar como ineficaces las medidas e iniciativas que se han adoptado para hacer frente a estos problemas. Aunque todavía les queda un largo camino por recorrer, las empresas de redes sociales han respondido adoptando una serie de medidas, que, con el tiempo, se han vuelto más sofisticadas. Entre ellas se incluye la actualización y aclaración periódica de sus políticas,¹⁰⁵

la inversión en la actualización de sus sistemas para ayudarles a detectar contenido problemático en una gama más amplia de idiomas,¹⁰⁶ la expansión de sus programas de verificación de hechos¹⁰⁷ o de alertadores fiables a un mayor número de países para comprender mejor las especificidades nacionales,¹⁰⁸ y la elaboración de respuestas más sofisticadas para abordar el contenido «problemático». Asimismo, tratan constantemente de innovar y periódicamente desarrollan nuevas funciones que ofrecen a los usuarios y usuarias más control sobre lo que ven,¹⁰⁹ y han introducido mecanismos para recurrir la eliminación indebida de contenidos según las normas de su comunidad.¹¹⁰ Algunas empresas han empezado a realizar evaluaciones de derechos humanos¹¹¹ y continúan aumentando la cantidad de moderadores humanos en sus plataformas,¹¹² a menudo en respuesta a crisis particulares o iniciativas reglamentarias previstas.

Además, muchas de las áreas de contenido problemático que impulsan la regulación de las plataformas son complejas y difíciles de definir. Por ejemplo, la «desinformación» es un problema complicado, sobre todo porque, en la mayoría de los casos, es difícil separar los hechos y de las opiniones. Por esta razón, la prohibición y otras restricciones legales sobre el intercambio de información falsa pueden ser objeto de abuso y tener efectos devastadores en el discurso político. ARTICLE 19 observa que el impacto de las políticas de las empresas de redes sociales en este ámbito es casi desconocido.¹¹³ En general, creemos que existen pocos datos que demuestren si han tenido éxito o no las iniciativas actuales para abordar las preocupaciones sobre la difusión de «desinformación». Además, algunos estudios académicos han señalado que es sumamente especulativo hasta qué punto la desinformación realmente es perjudicial para la democracia.¹¹⁴ Por todas estas razones, creemos que no queda probado el argumento a favor de una mayor regulación en este ámbito.¹¹⁵ Al mismo tiempo, reconocemos que podría ser necesaria una supervisión o auditoría pública independiente de las afirmaciones que hacen las plataformas de redes sociales sobre sus actuaciones para abordar cómo hacen frente a la «desinformación».¹¹⁶

Además, muchos tipos de contenido problemático son cuestiones difíciles y complejas que no tienen soluciones fáciles. Por ejemplo, varios de los problemas que plantea el «discurso de odio» sin duda se han exacerbado en línea, ya que este tipo de contenido es propenso a volverse viral. Se espera que los gobiernos y las empresas tomen medidas para abordar estos problemas y es preciso que lo hagan. La cuestión es si las medidas que las empresas de redes sociales han adoptado hasta la fecha son suficientes y, de no ser así, si la nueva legislación que aborda el «discurso de odio» en línea es una solución adecuada para el problema al que nos enfrentamos. Asimismo, no deben descartarse de más los pasos que las empresas de redes sociales han dado para abordar el «discurso de odio».¹¹⁷

Por lo tanto, es difícil concluir que las iniciativas de autorregulación hayan sido un fracaso total. Esto no significa que las empresas de redes sociales puedan permitirse ser complacientes; deben seguir haciendo más para abordar el «discurso de odio» y ser más transparentes al respecto. En particular, es importante que los distintos actores en este debate, incluidos los gobiernos, las empresas y la sociedad civil, continúen dialogando para garantizar un equilibrio adecuado entre la protección de los derechos a la igualdad y la libertad de expresión.

A este respecto, ARTICLE 19 considera que los nuevos modelos de autorregulación independientes o multisectoriales, como los consejos de redes sociales, podrían ser parte de la solución.¹¹⁸ Esto es especialmente cierto en algunos ámbitos, particularmente cuando el contenido en cuestión no es ilícito. Por lo tanto, siguen siendo necesarias algunas formas de autorregulación independiente por las siguientes razones fundamentales:

Los argumentos a favor de un marco legislativo general son problemáticos

ARTICLE 19 observa que los otros argumentos que se han presentado a favor de un marco legislativo general sobre el contenido en línea son muy discutibles. En particular, los siguientes:

- **La necesidad de garantizar la seguridad jurídica:** se trata de un poderoso argumento a favor de un marco legislativo más unificado. Este es el caso, por ejemplo, en la UE.¹¹⁹ Al mismo tiempo, ARTICLE 19 observa que, por lo que respecta a la UE, este argumento parece difícil de reconciliar con su enfoque hasta el momento, es decir, la adopción de instrumentos legales separados para abordar los derechos de autor o el terrorismo en línea, y la revisión de la Directiva de servicios de comunicación audiovisual para responder al «discurso de odio» y el contenido en línea nocivo para la infancia.¹²⁰ En muchos casos, las medidas que la UE ha adoptado en los últimos años todavía no se han puesto en práctica. Por lo tanto, no queda claro si son efectivas. En general, creemos que la consiguiente adopción de una Ley de Servicios Digitales no necesariamente aportará mayor claridad en este ámbito del derecho. Además, es muy cuestionable que un regulador único, ya sea a nivel nacional o europeo, esté en condiciones de abordar preocupaciones tan diversas como la «desinformación», el «discurso de odio», las imágenes de abuso infantil y el contenido terrorista.
- **Regulación de una amplia variedad de servicios:** además, como explicamos en detalle más adelante, nos preocupa que la reforma en este ámbito pueda utilizarse para regular servicios o actores como las aplicaciones de mensajería privada o la prensa, que deberían estar fuera de su ámbito de aplicación debido a la amenaza que dicha regulación representaría para la protección de los derechos a la libertad de expresión y la privacidad.

- **Establecimiento de un organismo regulador de ventanilla única:** por último, aunque reconocemos que un regulador puede contribuir al cumplimiento de las obligaciones de transparencia y debido proceso, nos preocupa que las cuestiones sobre el procedimiento a menudo se mezclan con las relacionadas con el contenido, como la adopción de filtros para detectar contenidos nocivos o ilícitos. Los reguladores también siguen siendo una importante vía para que los gobiernos ejerzan control sobre los flujos de información, particularmente en países donde estos reguladores no son independientes. Conceder a los reguladores poder sobre la prestación de servicios de redes sociales podría, en última instancia, tener un efecto disuasorio sobre la libertad de expresión de los usuarios y usuarias.

Por lo tanto, ARTICLE 19 considera que no se han demostrado los argumentos a favor de que un regulador supervise el contenido en línea y de la eliminación de la exención de responsabilidad, especialmente cuando se trata de contenido que entra en la categoría de «nocivo pero lícito». Creemos que otros mecanismos, como los organismos independientes multisectoriales (consejos de redes sociales), podrían ser más adecuados para supervisar el cumplimiento por parte de las empresas de redes sociales de un conjunto de principios derivados de los estándares de derechos humanos.

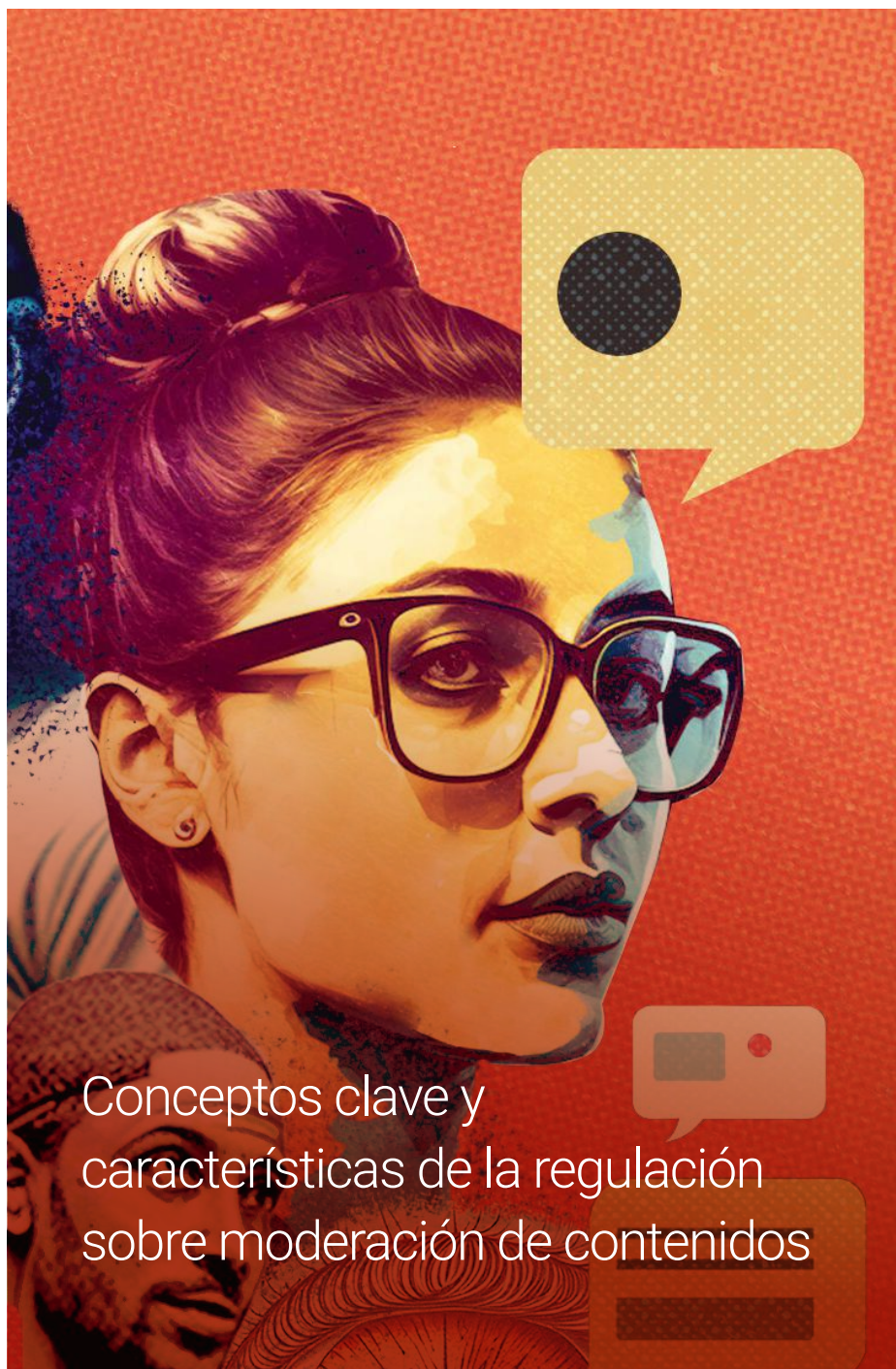
Es preciso abordar la falta de diversidad de los medios de comunicación

Como organización de defensa de la libertad de expresión, ARTICLE 19 se toma muy en serio las preocupaciones sobre el pluralismo y la diversidad de los medios de comunicación. Aunque las preocupaciones sobre las burbujas de filtro pueden haberse exagerado,¹²¹ está igualmente claro que los algoritmos que se utilizan para la curación de contenidos no son neutrales respecto a la diversidad, y que es muy probable que los cambios que se hagan en ellos tengan un impacto en el consumo de

noticias. Además, existen pruebas que demuestran que las audiencias en línea están más polarizadas, lo que podría, a su vez, incentivar la producción de contenidos más sesgados.¹²² En definitiva, la preocupación general sigue siendo que ciertas plataformas o servicios de búsqueda, debido a su tamaño, cuota de mercado y base de usuarios, tienen demasiado poder sobre las condiciones de distribución de noticias y, en última instancia, sobre la supervivencia del periodismo. ARTICLE 19 considera que algunas de estas preocupaciones podrían abordarse mediante una serie de medidas, en particular, la creación de consejos de redes sociales (que fomentarían el diálogo entre las plataformas y las organizaciones de noticias) y la separación de los servicios de alojamiento y curación de contenidos (que favorecería la competencia en el ámbito de la curación de contenidos y, por tanto, la diversidad y la capacidad de elección de los usuarios y usuarias). Más abajo, describimos estas propuestas de forma más detallada.

Es preciso reformar el modelo de negocio de las plataformas para que cumpla con lo dispuesto en la legislación en materia de protección de datos

ARTICLE 19 cree que el modelo de negocio de las grandes plataformas y otras empresas de redes sociales interfiere de forma significativa en el derecho a la privacidad y puede tener un efecto disuasorio sobre la libertad de expresión. Como mínimo, es evidente que las empresas de redes sociales deben cumplir la legislación en materia de protección de datos a fin de restringir la cantidad de datos personales que recopilan sobre los usuarios y usuarias. Asimismo, ARTICLE 19 considera que, a la luz de los recientes acontecimientos (particularmente en torno a las elecciones), deben revisarse las prácticas políticas y publicitarias actuales por su impacto en la libertad de expresión, y podría ser necesaria regulación en este ámbito. ARTICLE 19 aborda esta cuestión en otra política.¹²³



Conceptos clave y
características de la regulación
sobre moderación de contenidos

Conceptos clave y características de la regulación sobre moderación de contenidos

En el espacio de unos pocos años, se ha cuestionado cada vez más el amplio consenso, en los países occidentales, de que la exención condicional de responsabilidad es un requisito previo necesario para la innovación técnica y la libertad en Internet. Asimismo, se ha socavado de manera continuada la prohibición de la supervisión general del contenido en las plataformas de redes sociales. Como se ha señalado previamente, varios Estados han propuesto o adoptado gradualmente una serie de medidas que erosionan estos importantes cimientos de la protección de la libertad de expresión en Internet.¹²⁴ En la presente sección, revisamos algunos de los conceptos clave que se presentan en las propuestas legislativas en este ámbito.

En primer lugar, ARTICLE 19 señala una serie de conceptos erróneos recurrentes en los debates sobre la regulación de contenidos en línea:

- **Responsabilidad frente a regulación:** las normas sobre responsabilidad permiten a los particulares demandar a personas físicas o jurídicas por causarles un perjuicio, sobre la base de fundamentos existentes. Por el contrario, la exención de responsabilidad en el sector digital significa que las empresas digitales están protegidas contra demandas salvo que no adopten medidas cuando tienen conocimiento cierto de la ilegalidad. Teniendo en cuenta los millones de contenidos que se publican cada minuto y la posibilidad de que este contenido sea ilícito, la exención de responsabilidad

es una protección muy importante tanto para las empresas como para la libertad de expresión en general. Si se eliminara esta protección, estas empresas probablemente dejarían de existir por completo, debido al riesgo legal que correrían al permitir que se publiquen contenidos en sus servicios, algo que también iría en detrimento de la libertad de expresión. La regulación, por el contrario, significa que las empresas tienen que cumplir con un conjunto definido de obligaciones establecidas por ley y, si no lo hacen, pueden enfrentarse a sanciones como multas. La exención de responsabilidad y la regulación pueden coexistir.

- **Regulación del contenido frente a regulación de las plataformas:** el debate actual sobre cómo se regulan las empresas digitales a menudo incluye referencias a la «regulación de contenido» o a la «regulación de las plataformas». Es importante recordar que el contenido está regulado principalmente por la ley que se aplica directamente a las personas físicas, por ejemplo, la legislación sobre difamación o la tipificación como delito de la incitación a la violencia. La situación de las empresas es diferente, ya que permiten que terceros (es decir, los usuarios y usuarias) publiquen contenidos sin verificación previa (aunque esto está cambiando cada vez más con el uso de tecnología de filtrado). Aunque las propuestas legislativas actuales se presentan como una regulación de las plataformas, en gran medida delegan en estas la responsabilidad de regular el discurso de los usuarios y usuarias, a menudo más allá de lo que exige la ley (véase más abajo). Además, es importante tener en cuenta que, en la práctica, la regulación está dirigida a las *actividades*, los *comportamientos* y los *servicios* de las empresas, más que a las empresas en sí.

- **La responsabilidad del editor frente a la responsabilidad del intermediario:** las empresas digitales se etiquetan cada vez más como «empresas de medios de comunicación». Algunos argumentan que la misma responsabilidad aplicable a los editores también debería aplicarse a las plataformas de redes sociales dominantes como Facebook, Twitter y YouTube. Aunque podría afirmarse que algunas plataformas de redes sociales se comportan cada vez más como las empresas de medios de comunicación tradicionales, siguen existiendo algunas diferencias clave. En particular, el hecho de que tienen millones de usuarios y usuarias que publican contenidos en su plataforma. Estos usuarios y usuarias son los principales editores de ese contenido. Las plataformas de redes sociales lo seleccionan y organizan y, por lo tanto, ejercen una forma de responsabilidad editorial, un proceso que a veces se denomina «curación de contenidos», pero no son responsables del contenido en sí mismo, salvo que lo modifiquen lo suficiente como para que pueda considerarse propio. La cuestión clave es si el mero uso de algoritmos y filtros es suficiente para que las empresas tecnológicas tengan conocimiento cierto de la ilegalidad en sus plataformas y, por lo tanto, pierdan la exención de responsabilidad. En la siguiente sección, respondemos a esta pregunta.

Aspectos positivos de las propuestas legislativas sobre la regulación de la moderación de contenidos

ARTICLE 19 constata que varias propuestas normativas en esta materia incluyen algunos aspectos positivos para la protección del derecho a la libertad de expresión, entre ellos los siguientes:

- **Requisitos más estrictos de transparencia** sobre cómo las redes sociales y otras empresas gestionan la moderación de contenidos, que incluyen la transparencia sobre la eliminación de los contenidos de los usuarios y usuarias o sobre la tramitación de sus quejas.¹²⁵

- **Requisitos para establecer mecanismos de recurso y varias otras garantías procesales**, como la notificación de eliminación de contenido a los usuarios y usuarias cuyos contenidos se han suprimido, y la posibilidad de que los usuarios y usuarias pueden recurrir las decisiones de las empresas ante los tribunales.¹²⁶
- **Exención de responsabilidad**: la mayoría de las propuestas¹²⁷ no rechazan explícitamente la exención de responsabilidad, pero la ponen en riesgo al exigir (más o menos expresamente) la adopción de «medidas proactivas» para detectar los contenidos ilícitos o nocivos. La pregunta es: ¿siguen teniendo las empresas un incentivo suficiente para proteger la libertad de expresión si se les garantiza la exención de responsabilidad a través de una «cláusula del buen samaritano»,¹²⁸ incluso cuando aplican filtros y otras medidas? En términos generales, por las razones que explicamos más adelante, creemos que la respuesta a esta pregunta es «sí».
- **Enfoques sistémicos**: si bien las personas conservan la capacidad de demandar a las empresas por no eliminar el contenido «ilícito», las propuestas normativas actuales no prevén sancionar los casos individuales en los que no se han eliminado contenidos, más bien, adoptan un enfoque más «macro» de la moderación de contenidos y se centran en los sistemas y procesos de las empresas.¹²⁹
- **Enfoques escalonados**: en la medida en que la regulación se considere necesaria —en nuestra opinión, no lo es— una característica positiva de muchas propuestas es que tienden a plantear enfoques escalonados según el tamaño de la empresa en cuestión, al menos en referencia al número de usuarios y usuarias.¹³⁰

Deficiencias de las propuestas normativas sobre la moderación de contenidos

A pesar de tener varias características positivas, la mayoría de las propuestas actuales sobre la regulación de la moderación de contenidos crean problemas significativos desde el punto de vista de los estándares legales internacionales en materia de protección de la libertad de expresión. Como se ha señalado anteriormente, para que una interferencia en el derecho a la libertad de expresión esté justificada con arreglo al Derecho internacional, debe estar prevista en la ley, perseguir un fin legítimo, como se expone exhaustivamente en los tratados de derechos humanos, y ser necesaria y proporcionada para tal fin. La gran mayoría de las propuestas no satisfacen los principios de legalidad y proporcionalidad.

Focalización en el contenido «lícito pero nocivo»

La mayoría de las propuestas pendientes van más allá del asunto en cuestión, pues no solo cubren contenidos *ilícitos*, sino también contenidos meramente *nocivos*.¹³¹ Esta cuestión es sumamente preocupante pues significa que probablemente las empresas estén jurídicamente obligadas a tomar medidas contra el contenido que está efectivamente permitido por la ley. Las personas podrían decir cosas en la calle que no están permitidas en las plataformas de redes sociales. Además, la mayor parte del contenido «nocivo» es intrínsecamente difícil de definir y, por lo tanto, es probable que dé lugar a censura de contenido que, aunque ofensivo para algunos, debería mantenerse. Este es el caso, por ejemplo, del material que tiene el efecto «probable» de humillar a una persona,¹³² o del material «extremista» «con una definición menos clara».¹³³ Si bien algunas de las propuestas proporcionan mayor seguridad jurídica (ya que se refiere al material ilícito), las leyes a las que se hace referencia suscitan preocupación desde la perspectiva de la libertad de expresión.¹³⁴

Variedad excesiva de empresas consideradas

Otra preocupación importante es que una amplia gama de servicios, como la mensajería privada o los sitios de noticias, probablemente entren en el ámbito de aplicación de las propuestas existentes¹³⁵ cuando en realidad deberían permanecer fuera. A ARTICLE 19 le preocupa especialmente que la regulación en esta área pueda exigir que las empresas que entran en su ámbito de aplicación filtren el contenido, lo que debilitaría significativamente el cifrado y, por lo tanto, la privacidad de las comunicaciones de los usuarios y usuarias. En el caso de los sitios de noticias, nuestra preocupación es que puedan ser objeto de una regulación que podría utilizarse para silenciar su información, especialmente cuando son críticos con los responsables políticos. Además, en las propuestas existentes, parece que se ha prestado poca atención al grado en que los diferentes proveedores de servicios pueden verse afectados de diferentes maneras. Por ejemplo, un proveedor de infraestructura como Cloudflare no debería estar sujeto a las mismas obligaciones que una plataforma de redes sociales como Facebook.

Si bien algunas propuestas tienden a adoptar un enfoque escalonado de las obligaciones que se imponen a las empresas que entran en su ámbito de aplicación, ya sea tomando como referencia el número de usuarios y usuarias o el volumen de negocios, no siempre es así.¹³⁶ Además, hay poca consistencia en los criterios utilizados, no siempre se prevén excepciones para las organizaciones sin fines de lucro (por ejemplo, Wikipedia), y los proveedores más pequeños con una gran base de usuarios pero con un volumen de negocios bajo aún pueden estar sujetos a obligaciones onerosas.¹³⁷

Obligaciones vagas y problemáticas

Si bien algunas obligaciones propuestas son, en general positivas (por ejemplo, en lo que respecta a la transparencia y el debido proceso), otras son demasiado vagas o totalmente incompatibles con los derechos a la libertad de expresión y la privacidad:

- **Deber de diligencia:** algunas propuestas se refieren de forma general al «deber de diligencia», sin definir claramente qué tipo de obligaciones conlleva¹³⁸ ni su relación con la responsabilidad del intermediario. En general, el enfoque de «deber de diligencia» se basa en una evaluación del riesgo de «daño» para los usuarios y usuarias. Incluso si el «deber de diligencia» implicara una evaluación del riesgo de «daño» derivado de los sistemas de las empresas,¹³⁹ el término «daño» es difícil de definir, y determinar lo que constituye un riesgo «inaceptable» se dejaría a discreción del regulador.¹⁴⁰ Sin más claridad, el «deber de diligencia» podría parecerse a cualquiera de las obligaciones que se detallan más adelante, es decir, la eliminación de contenidos dentro de plazos excesivamente cortos, las medidas proactivas o los requisitos de eliminación del cifrado.¹⁴¹

ARTICLE 19 observa que quienes proponen una «debe de diligencia» obligatorio parecen sugerir que sus propuestas son consistentes con el tipo de evaluaciones de impacto en los derechos humanos que promueven los Principios Rectores.¹⁴² En este análisis, parece sugerirse que las evaluaciones de riesgo sobre las posibles repercusiones en los derechos humanos son lo mismo que las evaluaciones de riesgo de «daño» potencial. ARTICLE 19 considera que esto es fundamentalmente erróneo. Los derechos humanos son conceptos jurídicos incluidos en leyes nacionales, constituciones e instrumentos internacionales. Aunque por lo general se expresan en términos amplios, se definen además en base a la vasta jurisprudencia de los tribunales. En consecuencia, a pesar de que en apariencia son muy amplios, en realidad están cuidadosamente definidos. El concepto de «daño», por el contrario, no tiene una definición legal general. En la medida en que se usa en entornos jurídicos, tiende a definirse cuidadosamente y en un contexto muy específico, por ejemplo, en el caso del delito de causar daños corporales.

Los Principios Rectores están concebidos para prevenir violaciones de derechos humanos, no para evitar «daños». Además, a diferencia del «deber de prevenir daños», los Principios Rectores no exigen que las empresas prevengan las violaciones de derechos humanos *causadas por otros*.¹⁴³ A lo sumo, exigen que las empresas adopten procedimientos que tengan por objeto evitar o minimizar su participación en violaciones de derechos humanos.

- **Obligación de eliminar determinados contenidos:** la gran mayoría de las propuestas delegan efectivamente competencias de censura en empresas privadas para que eliminen contenidos ilícitos o nocivos. Si no lo hacen, el regulador puede ordenar el bloqueo de estos servicios.¹⁴⁴ Aunque algunas propuestas incluyen el requisito de que el organismo de supervisión se asegure de que las empresas no retiren contenidos *de forma excesiva*,¹⁴⁵ en general, no existe un requisito correspondiente de transparencia sobre el contenido eliminado *ilícitamente*. En cualquier caso, salvo que se presente una reclamación, es posible que nunca se sepa la cantidad de contenidos que se han suprimido ilícitamente. Además, depende en parte de la eficacia de los mecanismos internos de recurso y de otras formas de reparación judicial. No obstante, la mayoría de las propuestas no hacen ninguna referencia a cómo deberían financiarse estas últimas.
- **Obligación de eliminar contenidos a corto plazo:** en general, la mayoría de las propuestas tienden a exigir la eliminación, en 24 horas, de contenidos que constituyen «discurso de odio» o son «manifiestamente ilícitos» y, en una hora, los contenidos terroristas.¹⁴⁶ Estos plazos son claramente demasiado cortos para hacer una evaluación adecuada, y mucho menos legal, de las afirmaciones que se hacen.

- **Obligación de tomar medidas proactivas:** algunas propuestas incluyen el requisito de que las empresas adopten medidas «proactivas» o que hagan «todos los esfuerzos» para abordar el contenido ilícito o nocivo. A veces, esta obligación se presenta como «recomendaciones» formuladas por el regulador, aunque la ley deja claro que se espera su cumplimiento¹⁴⁷. En la práctica, las «medidas proactivas» y otras condiciones relacionadas implican la adopción de filtros u otra tecnología para identificar o evitar la carga de contenido «problemático». Los filtros, por supuesto, no pueden evaluar la legalidad del contenido y se ha demostrado repetidamente que son propensos a errores, particularmente cuando es necesario un análisis del contexto, como en el caso de contenido «terrorista» o del «discurso de odio». Por consiguiente, pueden eliminarse ilícitamente contenidos que son legítimos.

El regulador

La mayoría de las propuestas para regular las plataformas de redes sociales las colocaría bajo el control de un regulador de «radiodifusión», lo que ampliaría considerablemente su mandato y facultades. Para ARTICLE 19, esto suscita varias preocupaciones:

- **Necesidad y legitimidad:** en primer lugar, una de las principales preocupaciones con respecto a estas propuestas es que, en última instancia, pondrían el discurso de los usuarios y usuarias bajo el control de una autoridad pública a la que se le podrían otorgar amplias facultades discrecionales con respecto a lo que constituye contenido «nocivo». Todavía no está claro si esta medida es necesaria ni su impacto probable en el derecho de los usuarios y usuarias de Internet a la libertad de expresión. En cualquier caso, es claramente inapropiado que una empresa privada o un regulador determinen la legalidad del contenido. Este es un asunto para los tribunales.

- **Independencia:** otra preocupación fundamental para una organización internacional como ARTICLE 19 es que el regulador puede no ser independiente, ya sea en la ley o en la práctica. No se trata de una preocupación teórica ni en democracias establecida¹⁴⁸ ni en muchos países del mundo donde el Estado de derecho es débil o está amenazado
- **Ventanilla única:** por último, nos preocupa que las propuestas actuales dejen una amplia variedad de diferentes tipos de contenido bajo el control de un regulador que puede no ser el más capacitado para abordar estos problemas, y menos aún para ofrecer el tipo de soluciones que ayudarían a resolverlos. Por ejemplo, la eliminación de material de abuso infantil es un área especializada que puede requerir procedimientos y consideraciones especiales, que una institución dedicada abordaría mejor. En nuestra opinión, es poco probable que un regulador de «ventanilla única» para todo el contenido ilícito o «nocivo» sea efectivo.

Sanciones desproporcionadas

ARTICLE 19 observa que las sanciones que se consideran actualmente en varios países, como Francia o Reino Unido, son especialmente altas, por ejemplo, hasta el 4 % de la facturación total de las empresas. Nos preocupa que esto, unido a unas obligaciones demasiado vagas, incentive a las empresas a eliminar más contenido en detrimento de la protección de la libertad de expresión en línea. Además, constatamos que con frecuencia no está claro el umbral al que es probable que se apliquen las sanciones¹⁴⁹ y estas medidas son desproporcionadas.

Otras preocupaciones

A ARTICLE 19 le preocupa que las leyes o proyectos de ley destinados a regular las plataformas cada vez incluyen más requisitos para que las empresas designen puntos de contacto o representantes «en el país». En la práctica, esto permite a las autoridades ejercer una presión significativa sobre las empresas para que cumplan con sus demandas, incluso si tienen poca o ninguna base legal y claramente buscan silenciar a las personas críticas con el gobierno.¹⁵⁰

Asimismo, cuestionamos hasta qué punto las empresas con operaciones internacionales pueden cumplir regímenes diferentes y potencialmente contradictorios en todo el mundo. En nuestra opinión, esto podría dar lugar a una Internet más fragmentada y, potencialmente, a una competición a la baja, en la que el contenido controvertido se elimine cada vez más por defecto con arreglo a las condiciones de servicio de las empresas.

Por último, constatamos que las leyes en este ámbito con frecuencia se adoptan sin un proceso participativo adecuado en el que se implique a todos los sectores, incluidas las plataformas digitales pequeñas y locales.



Recomendaciones de ARTICLE 19

Recomendaciones de ARTICLE 19

Recomendación 1:

Los Estados deben evitar regular innecesariamente la moderación de contenidos en línea

ARTICLE 19 deja claro que no apoyamos los modelos de regulación que hemos visto surgir en Europa Occidental o América Latina. Nos sigue preocupando que las normas propuestas suelen ser demasiado vagas y se basan en la independencia de un regulador de buena fe, lo cual no es el caso en muchas partes del mundo. El contexto más amplio en el que se hacen las propuestas legislativas es especialmente importante.¹⁵¹ Además, según el principio de tres partes para limitar la libertad de expresión, los Estados deben demostrar plenamente que existen pruebas suficientes para concluir que la regulación es necesaria. Si bien es entendible el impulso de regular las plataformas muy grandes, estas podrán adaptarse y cumplir los requisitos normativos, pero es poco probable que este sea el caso de las plataformas pequeñas o medianas. Una mayor regulación en línea, como mínimo, probablemente beneficiaría a las empresas muy grandes y afianzaría su posición dominante.

Por consiguiente, instamos a los legisladores a resistir la tentación de adoptar reglamentos innecesarios. Si las principales preocupaciones de nuestras sociedades se derivan del poder y el dominio de un pequeño número de plataformas, en nuestra opinión, los legisladores deberían abordar esas preocupaciones utilizando los recursos disponibles en virtud de la legislación en materia de competencia y la regulación económica favorable a la competencia. También por este motivo, consideramos que nuestras posiciones y recomendaciones sobre la responsabilidad de los intermediarios y las normas de moderación de contenidos para las plataformas de redes sociales siguen siendo pertinentes.¹⁵²

Además, ARTICLE 19 aboga por la supervisión de las empresas de redes sociales por parte de una institución independiente multisectorial, como los consejos de redes sociales. Seguimos convencidos de que la creación de una institución integral comportaría más beneficios para la libertad de expresión al evitar sanciones demasiado severas. Asimismo, fomentaría una rendición de cuentas más eficaz, ya que sus partes interesadas asumirían como propio el éxito de la organización. Los consejos de medios sociales y la regulación de plataformas muy grandes a través de la ley en materia de competencia desempeñan funciones complementarias.

Recomendación 2:

Los principios generales de cualquier marco legislativo deben ser la transparencia, la rendición de cuentas y la protección de los derechos humanos

Sin perjuicio de lo anterior, ARTICLE 19 reconoce que una mayor regulación del sector digital y, en particular, de las empresas de redes sociales parece casi inevitable.¹⁵³ Por esta razón, creemos que los objetivos de cualquier nuevo marco legislativo no pueden limitarse únicamente a «luchar contra la ilegalidad». Asimismo, la «prevención del daño» es un concepto demasiado amplio para ser un objetivo significativo, y mucho menos legítimo, de cualquier marco de este tipo.

Creemos que los principios generales de cualquier marco legislativo deben ser la transparencia, la rendición de cuentas y la protección de los derechos humanos. Esto último significa que los principios de legalidad y proporcionalidad deben respetarse en todo momento. Además, cualquier marco de este tipo debe basarse en pruebas sólidas a fin de adoptar las soluciones más adecuadas.

Recomendación 3:

Debe mantenerse la exención condicional de responsabilidad por los contenidos de terceros, pero es preciso aclarar su ámbito de aplicación y sus procedimientos de notificación y actuación

ARTICLE 19 señala que eliminar, o limitar indebidamente, la exención de responsabilidad incentivaría a las empresas digitales a filtrar y eliminar la mayor cantidad posible de contenido de los usuarios y usuarias o a ser completamente neutrales y no eliminar ningún contenido. Es decir, daría lugar a una mayor censura o eliminaría cualquier incentivo para que las empresas participen en la moderación de contenidos. En nuestra opinión, ambos resultados serían lamentables para la protección de la libertad de expresión en línea.

Por esta razón, consideramos que las empresas digitales y de redes sociales deben seguir beneficiándose de una exención de responsabilidad amplia, o al menos condicional. A este respecto, señalamos que, por definición, la responsabilidad solo puede vincularse al contenido ilícito, no al contenido «nocivo». En la práctica, esto significa que no se deben exigir responsabilidades a las empresas por no tomar «medidas razonables» indefinidas para abordar el contenido «nocivo».

La responsabilidad debe diferenciarse en función de los distintos tipos de actividades y servicios de la siguiente manera:

Se debe mantener una amplia exención de responsabilidad para aquellos que prestan servicios de infraestructura esenciales, incluida la «mera transmisión» y el «alojamiento» neutral

ARTICLE 19 considera que las empresas que prestan servicios de infraestructura esenciales para el funcionamiento de Internet, como las redes de distribución de contenidos, deben beneficiarse de una exención de responsabilidad más amplia

que los servicios que se dedican a la moderación de contenidos en la capa de aplicación. Solo se les debe exigir que eliminen el contenido por orden de un tribunal. En la práctica, esto significa que los proveedores de infraestructura no deben ser sancionados por alojar ciertos sitios web, a menos que no hayan cumplido con una orden judicial válida que les exija interrumpir sus servicios en dicho sitio web porque es ilegal.¹⁵⁴ Del mismo modo, no se les debe exigir que alojen dicho sitio web si no desean hacerlo. Es decir, los proveedores de infraestructura esencial no deberían estar obligados a transportar contenido.

Al mismo tiempo, los proveedores que prestan un servicio esencial deben exponer claramente las razones por las que pueden decidir interrumpir la prestación de servicios a determinados actores.

Procedimientos de notificación y actuación para aquellos proveedores que prestan servicios de alojamiento y moderación de contenidos

ARTICLE 19 considera, asimismo, que se debe mantener el estándar actual de conocimiento exigido para beneficiarse de la exención de responsabilidad, es decir, debe seguir siendo un conocimiento «cierto» en lugar de «implícito». ARTICLE 19 sigue convencido de que únicamente puede tenerse conocimiento cierto de la *ilegalidad* mediante una orden judicial. Sostener lo contrario sería aceptar que el contenido es ilícito simplemente porque un tercero, como el titular de los derechos de autor, así lo afirma.

Al mismo tiempo, reconocemos que un marco legislativo podría aclarar los diferentes tipos de procedimientos de notificación y actuación aplicables a distintos tipos de contenido, sin perjuicio de la determinación de la legalidad. Es decir, la ley debe aclarar cómo las empresas pueden tener «conocimiento cierto» de la presunta ilegalidad y qué deben hacer al respecto una vez que lo tienen.

ARTICLE 19 ha establecido previamente cómo podría funcionar en la práctica.¹⁵⁵ Consideramos que los procedimientos que hemos sugerido siguen siendo válidos y constituyen la mejor manera de proteger el derecho a la libertad de expresión. En resumen, incluyen lo siguiente:

Procedimientos de «notificación y notificación» para disputas privadas, como derechos de autor o difamación

Con arreglo a este procedimiento, el denunciante o el «alertador fiable» tendrían que dar su nombre y establecer en una notificación por qué cree que se han vulnerado sus derechos, el fundamento jurídico de su reclamación, la ubicación del material presuntamente infractor y la hora y fecha de la presunta infracción. Se exigiría al proveedor de alojamiento transmitir la notificación al presunto infractor (es decir, el proveedor de contenido) tan pronto como sea posible, pero dentro de un plazo máximo (por ejemplo, 72 horas). El proveedor de contenido tendría la opción de eliminar el contenido o presentar una contra notificación dentro de un plazo razonable (por ejemplo, 14 días).

De nuevo, se exigiría al proveedor de alojamiento transmitir la contra notificación tan pronto como sea posible, pero dentro de un plazo máximo (por ejemplo, 72 horas). A continuación, se daría al denunciante un plazo (por ejemplo, 14 días) para decidir si quiere llevar el asunto a los tribunales. El contenido se eliminaría previa orden judicial. El proveedor de alojamiento podría ser considerado responsable de la indemnización que marca la ley si no cumple con sus obligaciones emanadas del principio de «notificación-notificación» o si no elimina el contenido tras la orden judicial. Por el contrario, si el proveedor de contenido no responde o no presenta una contra notificación dentro de un plazo determinado, el proveedor de alojamiento pierde la exención de responsabilidad. Puede o bien eliminar el contenido presuntamente ilícito o si el denunciante decide llevar el asunto ante los tribunales u otro órgano judicial independiente,

puede ser considerado responsables del contenido en cuestión. A fin de proteger la libertad de expresión, cualquier nuevo marco de «notificación y notificación» también debe contemplar sanciones por notificaciones abusivas.

Procedimiento «notificación y actuación» para denuncias de delitos graves

Con arreglo a este procedimiento, el proveedor de alojamiento estaría *obligado* a eliminar contenido cuando reciba una orden judicial a tal efecto. Es decir, sería responsable por el incumplimiento de dicha orden. En la práctica, esto significa que, si las autoridades competentes creen que debe eliminarse una parte del contenido y el asunto no es urgente, deben solicitar una orden judicial, si es necesario, a petición de la parte interesada. Sin embargo, si la situación reviste urgencia, por ejemplo, la vida de alguien está en riesgo, la ley debería delegar a esas autoridades competencia para disponer el retiro inmediato o bloquear el acceso del contenido en cuestión. No obstante, una disposición semejante precisaría confirmación ante un tribunal dentro de un plazo determinado (por ejemplo, 48 horas). No debe permitirse el uso de mecanismos informales, por ejemplo, llamadas de teléfono o correos electrónicos para solicitar al host que retire el contenido.

Por el contrario, si los proveedores de alojamiento reciben una notificación de un usuario o usuaria sobre un presunto contenido delictivo, el host o la plataforma deben notificar, a su vez, a las autoridades pertinentes si tienen motivos para creer que la queja está bien fundada y amerita una mayor investigación. El host o la plataforma también pueden decidir eliminar el contenido en cuestión, como medida provisional, de conformidad con las condiciones de servicio. No obstante, no estarían obligados a hacerlo, y no retirarlo no generaría responsabilidad.

El mismo procedimiento se aplicaría a los organismos privados que colaboran con las fuerzas del orden y operan líneas directas a las que los usuarios y usuarias de Internet pueden llamar si sospechan que se han publicado contenidos delictivos en línea.¹⁵⁶ Es decir, la línea directa informaría tanto al host como a las autoridades competentes sobre el contenido en cuestión. El host aplicaría el mismo procedimiento que se utiliza para tramitar las quejas de los usuarios y usuarias, es decir, podría decidir libremente si elimina el contenido, basándose en sus condiciones de servicio. El mismo modelo podría aplicarse a otros organismos, ya sean públicos o privados, que reciban denuncias del público sobre contenidos potencialmente delictivos en línea, o para detectar problemas mediante «alertadores fiables» (más adelante se incluye información detallada sobre los programas de alertadores fiables). Cualquiera que sea la opción que se elija, es importante que se notifique a las autoridades cualquier denuncia de conducta delictiva grave para que pueda ser debidamente investigada y sancionada de acuerdo con el procedimiento establecido por el sistema de justicia penal.

Los Principios de Manila sobre la responsabilidad de los intermediarios ofrecen otras orientaciones útiles sobre cómo deberían funcionar los procedimientos de «notificación y actuación».¹⁵⁷ Consideramos que esta es la forma más proporcionada y respetuosa con los derechos en la que se pueden aplicar los procedimientos de «notificación y actuación», en particular en el caso de la pequeña empresa.

Protección frente a la responsabilidad en caso de medidas de moderación de contenidos que las empresas por iniciativa propia

ARTICLE 19 cree que las plataformas de redes sociales y otras empresas digitales no deben ser consideradas responsables simplemente porque han adoptado normas comunitarias y utilizan moderadores humanos u otras herramientas para aplicarlas.¹⁵⁸ En este sentido, apoyamos la adopción de una

«cláusula del buen samaritano» que fomente los esfuerzos de moderación de contenidos realizados de buena fe. En nuestra opinión, no hacerlo impediría la adopción de soluciones y herramientas técnicas innovadoras, como la desmonetización o la eliminación de ciertas funciones de la plataforma, con las que se podría conseguir un equilibrio más proporcionado entre la protección de la libertad de expresión y la lucha contra los contenidos ilícitos o incluso «nocivos». Al mismo tiempo, las empresas que utilizan estas herramientas deben estar sujetas a estrictos requisitos de transparencia y debido proceso sobre la forma en que los utilizan.

Del mismo modo, las empresas deberían beneficiarse de una exención de responsabilidad amplia por las recomendaciones que hacen sus algoritmos, en los casos en que dichos algoritmos recomienden contenido ilícito en respuesta al contenido que los usuarios y usuarias han visto. Si bien los desarrolladores y codificadores de sistemas definen los parámetros en los que operan los algoritmos, no controlan ni determinan el resultado de estos procesos automatizados. Los algoritmos producen resultados a partir de conjuntos de datos de forma compleja e impredecible. Por lo general, también son propensos a cometer errores y son incapaces de distinguir entre contenido legal e ilegal. Por lo tanto, sería inviable y desproporcionado responsabilizar a las empresas por todos los posibles «errores» que cometen sus sistemas. En la medida en que la responsabilidad se ocupa de casos específicos de ilegalidad, es también un instrumento deficiente para abordar los desafíos sistémicos que plantean los algoritmos.

En cambio, las empresas, en particular aquellas con un poder de mercado significativo, deberían estar sujetas a mayores obligaciones de transparencia y se les debe exigir que realicen evaluaciones de impacto en los derechos humanos, como se describe a continuación. En nuestra opinión, el mismo razonamiento debería aplicarse a los servicios de navegación

o de «localización», es decir, no se les deberían penalizar si el algoritmo de su motor de búsqueda devuelve contenidos ilícitos, pero deberían ser transparentes y explicar al público de qué forma opera su algoritmo para devolver los resultados de la búsqueda.

Por el contrario, entendemos que las empresas deben perder la exención de responsabilidad cuando «promuevan» u «optimicen» la presentación de contenido ilícito en la sección de publicidad de su plataforma como consecuencia de acuerdos comerciales.¹⁵⁹

Recomendación 4:

Se debe mantener la prohibición de supervisión general del contenido

ARTICLE 19 cree que los gobiernos deben abstenerse de imponer a las empresas una obligación de supervisión general del contenido.

Aunque se puede argumentar que la supervisión simplemente permite a las empresas detectar contenido potencialmente ilícito o problemático, en la práctica, la mera detección casi siempre va acompañada de la eliminación u otro tipo de acciones que reducen la disponibilidad de dicho contenido. Esto es sumamente problemático, dado que la tecnología de supervisión de contenidos no es tan avanzada como a veces se sugiere. En particular, los algoritmos que comparan hashes y las herramientas de procesamiento del lenguaje natural actualmente son incapaces de distinguir el contenido cuya legalidad puede variar según el contexto, como las noticias o las parodias.¹⁶⁰ Por consiguiente, se pueden eliminar grandes cantidades de contenido legítimo. Además, estas tecnologías interfieren en el derecho a la privacidad de los usuarios y usuarias, ya que requieren el análisis de las comunicaciones de las personas.

Además, si la ley condicionara la exención de responsabilidad a la «supervisión general», la adopción de «medidas proactivas» o a que las empresas hagan «todos los esfuerzos» para abordar el contenido ilícito,¹⁶¹ inevitablemente estas pecarían de exceso de cautela y eliminarían el contenido por defecto para evitar riesgos legales y costes de ejecución. Esto podría dar lugar a que las plataformas solo autorizaran oradores preseleccionados, o utilizaran sus condiciones de servicio para prohibir contenidos controvertidos,¹⁶² y podría, asimismo, disuadir a los nuevos participantes en el mercado de desafiar a los operadores históricos.¹⁶³

Al mismo tiempo, ARTICLE 19 reconoce que la supervisión y la eliminación «específicas» de vídeos u otras imágenes de abuso sexual infantil indiscutiblemente ilegales, es decir, la representación de actividad sexual (por ejemplo, penetración) entre un niño y un adulto, puede ser compatible con los derechos a la libertad de expresión y la privacidad.¹⁶⁴ Dada la gravedad de la conducta en cuestión y el hecho de que este tipo de contenido se puede identificar de forma fiable como ilícito independientemente del contexto. Sin embargo, no estamos de acuerdo en que dichas obligaciones específicas de supervisión deban aplicarse a ningún otro tipo de contenido.¹⁶⁵

Recomendación 5:

Todo marco legislativo debe tener un ámbito de aplicación estrictamente limitado

Como se ha señalado anteriormente, ARTICLE 19 considera que cualquier marco que tenga como objeto regular las actividades de moderación de contenidos de las plataformas debe tener un ámbito de aplicación limitado, en particular en lo que respecta al tema, las entidades a las que afecta y su aplicación geográfica. En concreto, formulamos las recomendaciones siguientes:

La regulación debe centrarse en el contenido ilícito, en lugar de en el contenido «lícito pero nocivo»

ARTICLE 19 considera que cualquier marco de este tipo debe limitarse al contenido «ilícito» en lugar del «nocivo», por la sencilla razón de que contenido «nocivo» es un concepto intrínsecamente vago, lo que hace que sea difícil de aplicar, sea susceptible al abuso e impugnabile por motivos legales. En nuestra opinión, el contenido ilícito que, no obstante, está prohibido con arreglo a las normas comunitarias de las empresas, debe estar sujeto a la supervisión de entidades independientes multisectoriales, como los consejos de redes sociales que ha propuesto ARTICLE 19.

Si el contenido «legal pero nocivo» se incluye dentro del ámbito de aplicación de la legislación, en contra de nuestras recomendaciones, solo deben imponerse requisitos de transparencia y debido proceso con objeto de hacer cumplir las normas comunitarias de la empresa. Por consiguiente, el papel del regulador se limitaría a garantizar que los sistemas de moderación de contenidos de las empresas sean lo suficientemente transparentes y que los usuarios y usuarias tengan a su disposición mecanismos de recurso claros y efectivos.

Los servicios de mensajería privada y las organizaciones de noticias deben quedar fuera del ámbito de aplicación de cualquier marco legislativo

Asimismo, creemos que el ámbito de aplicación de cualquier marco legislativo debe limitarse para que queden excluidos los comentarios que los lectores pueden publicar debajo de los artículos en los sitios web y blogs de los periódicos. Igualmente, el contenido periodístico debe, en principio, quedar excluido del ámbito de aplicación, incluso si se aloja en plataformas de redes sociales u otros servicios (por ejemplo, motores de búsqueda). En nuestra opinión, sería totalmente inapropiado que las medidas destinadas a regular las prácticas de las plataformas de redes sociales muy grandes se utilizaran como una puerta trasera para regular el contenido periodístico.

Las aplicaciones de mensajería y otros canales privados de comunicación también deben quedar fuera del ámbito de aplicación. En particular, los reguladores no deben tener facultades para imponer obligaciones a los proveedores cuando tales obligaciones impliquen una interferencia injustificable en el derecho a la privacidad de los usuarios y usuarias, como en el caso del debilitamiento del cifrado de extremo a extremo o los filtros obligatorios.

Las medidas no deben tener aplicación extraterritorial

Por último, consideramos que la implementación de medidas con arreglo a un nuevo marco legislativo debe limitarse geográficamente al país que ordena dichas medidas, de conformidad con los principios internacionales de cortesía y el principio de proporcionalidad en virtud del derecho internacional de los derechos humanos. Es decir, ningún país debe poder emitir órdenes para eliminar o restringir contenidos que puedan ser legales fuera de sus fronteras.

Recomendación 6:

Las obligaciones con arreglo a cualquier régimen legislativo deben estar claramente definidas

ARTICLE 19 considera que cualquier obligación con arreglo a un nuevo régimen legislativo que rija las actividades de las plataformas y otras empresas tecnológicas debe estar claramente definida. A continuación, detallamos los tipos de medidas que podrían incluirse como parte de dicho marco y las que deberían excluirse. En particular, creemos que un nuevo marco legislativo podría exigir lo siguiente:

Obligaciones de transparencia

En nuestra opinión, la transparencia debe ser un requisito básico que impregne todo lo que hacen las empresas. En particular, debe aplicarse:

- **Distribución de contenidos:** las plataformas de redes sociales y las empresas digitales deben proporcionar información esencial y explicar al público cómo se utilizan sus algoritmos para presentar, clasificar, promocionar o relegar los contenidos. Los contenidos que la empresa o un tercero promocionan a cambio de una remuneración deben estar claramente marcados.¹⁶⁶ Las empresas también deben explicar cómo orientan el contenido promocionado (no solicitado) a los usuarios y usuarias, ya sea por iniciativa propia o en nombre de terceros como un servicio pagado.¹⁶⁷
- **Condiciones de servicio y normas comunitarias de las empresas:** las empresas deben publicar normas comunitarias o condiciones de servicio que sean fáciles de entender y facilitar ejemplos de «jurisprudencia» que explique cómo se aplican. Deben publicar información sobre los métodos y procesos internos para la elaboración de normas comunitarias, que deben seguir incluyendo consultas con una amplia gama de actores, incluida la sociedad civil.¹⁶⁸
- **Recursos humanos y tecnológicos que ayudan a garantizar el cumplimiento:** las empresas deben incluir información detallada sobre los regímenes de alertadores fiables, en particular quién está en la lista de alertadores fiables, cómo han sido seleccionados y cualquier «privilegio» asociado a ese estado. Asimismo, deben publicar información sobre cómo funcionan sus algoritmos para detectar contenido ilícito o presuntamente «nocivo» con arreglo a las normas de su comunidad. Esto debe incluir, en particular, información sobre el porcentaje de falsos negativos y falsos positivos y los indicadores, si los hay, para evaluar el contenido que probablemente se vuelva viral, por ejemplo, la exposición a un público más amplio.¹⁶⁹

- **Proceso de toma de decisiones:** las empresas deben notificar sus decisiones a las partes afectadas y dar razones suficientemente detalladas de las medidas que toman contra contenidos o cuentas particulares. Además, deben proporcionar información clara sobre cualquier mecanismo interno de reclamación.
- **Informes sobre transparencia:** las empresas deben publicar información detallada de conformidad con los Principios de Santa Clara.¹⁷⁰ Cabe señalar que es especialmente importante no limitar la información estadística a la eliminación de contenidos. Se deben incluir también datos sobre el número de apelaciones procesadas y su resultado. Además, los informes relativos a la transparencia deben distinguir entre los contenidos señalados por terceros (incluso si se trata de organismos públicos o entidades privadas), alertadores fiables (ya sean organismos públicos o entidades privadas) y algoritmos. Asimismo, se debe facilitar más información sobre los diferentes tipos de restricciones aplicadas al contenido en el marco de procesos de moderación de contenidos, como la desmonetización o la relegación. Para toda restricción, las empresas deben facilitar información sobre las normas en las que se ha basado la decisión y, cuando esté disponible, el resultado de cualquier apelación.

En general, consideramos que cualquier requisito de información sobre la transparencia debe tener como objetivo proporcionar un análisis mucho más cualitativo de las decisiones de moderación de contenidos. La métrica para evaluar el éxito a la hora de abordar el contenido ilícito no debe vincularse al índice de eliminación de contenido, ya que esto fomenta la supresión excesiva. Asimismo, los informes sobre la transparencia no deben limitarse a la información que presentan las empresas, también deben incluir la que proporcionan las agencias gubernamentales pertinentes. Sin perjuicio de las medidas aplicables en virtud de la legislación en materia de consumo.¹⁷¹

- **Auditorías de transparencia:** las empresas deben facilitar, a los reguladores e investigadores independientes previamente escrutados —ya sean académicos, periodistas, etc.— un mayor acceso a los conjuntos de datos a fin de permitirles verificar si los sistemas y algoritmos funcionan como la empresa afirma. En particular, los auditores deben tener acceso a datos sobre: 1) los programas de moderación de contenidos de las empresas; 2) la forma en que las empresas ordenan, clasifican, priorizan, recomiendan o personalizan el contenido; y 3) cómo se aplica esto a la publicidad política.¹⁷² Los reguladores podrían tener acceso a datos confidenciales y comerciales, y los terceros previamente escrutados, a conjuntos de datos anónimos. Estas auditorías de las operaciones de las plataformas deben realizarse periódicamente.
- **Archivos de la publicidad política y digital:** las nuevas normas deben prohibir la «opacidad de datos» relativos a los anuncios políticos y solicitar que las plataformas utilicen mecanismos de cumplimiento claros en caso de violación de sus políticas. Las plataformas no deben desempeñar la función de arbitrar o mitigar el discurso político agresivo y la «desinformación». Además, instamos a garantizar una mayor transparencia con respecto a todos los gastos de publicidad política de las partes interesadas pertinentes, incluidos los partidos políticos, las empresas tecnológicas y los anunciantes externos. Los anuncios políticos deben distinguirse claramente del contenido editorial, incluidas las noticias, cualquiera que sea su forma, también en línea, y deben estar claramente etiquetados con información que indique quién ha pagado por ellos. Además, apoyamos la utilización de bases de datos de anuncios digitales para mantener un registro de todos los anuncios regulados, la cantidad de dinero gastado en publicidad y el nombre de la persona que autorizó el anuncio, y hacer esta información pública. Los datos deben ser accesible en un formato que permita a los investigadores y responsables políticos su recuperación en bloque.

Obligaciones internas de debido proceso

Sin perjuicio de lo dispuesto en la recomendación 9, ARTICLE 19 considera que cualquier marco legislativo que regule las actividades de las plataformas dominantes debe incluir la obligación de establecer lo siguiente:

- **Normas claras de notificación y actuación**, en consonancia con los Principios de Manila.
- **Mecanismos internos de recurso** para tramitar reclamaciones sobre restricciones al ejercicio del derecho a la libertad de expresión, como la eliminación ilícita de contenidos o la aplicación indebida de etiquetas que sugieran que una fuente de noticias no es fiable. Por otra parte, los mecanismos de apelación también deben poder abordar la negativa de una empresa a eliminar contenidos que presumiblemente viola sus normas comunitarias. En cualquier caso, los mecanismos internos de recurso deben respetar las garantías del debido proceso.¹⁷³
- **Obligación de promover la diversidad de los medios de comunicación**: dado el riesgo de que el contenido en las plataformas de redes sociales esté excesivamente personalizado, se debe exigir a las empresas de redes sociales muy grandes que tomen medidas para garantizar que su servicio incluya un nivel suficiente de diversidad de medios de comunicación,¹⁷⁴ así como una cobertura equilibrada durante las elecciones.¹⁷⁵ En particular, deben proporcionar información suficiente que explique cómo se seleccionan los canales de noticias y el material que promocionan. Asimismo, deben facilitar a los usuarios y usuarias alternativas de elección viables, que no se basen en la elaboración de perfiles. En determinadas circunstancias limitadas, como durante las elecciones, se les podría exigir que transmitan mensajes de medios de comunicación públicos a fin garantizar que el segmento más amplio posible de la población tenga la información básica que necesita para participar de manera significativa en el proceso electoral.

Abstenerse de imponer determinadas obligaciones

Por otra parte, ARTICLE 19 considera que ningún régimen legislativo de este tipo debe incluir la siguiente lista (no exhaustiva) de obligaciones:

- **Un «deber de diligencia» amplio e indefinido para prevenir una noción igualmente indefinida de «daño»:** en nuestra opinión, es poco probable que tales nociones superaran la prueba de legalidad conforme al derecho internacional de los derechos humanos. En la práctica, crearían inseguridad jurídica y otorgarían, a las autoridades reguladoras, poderes, en gran medida ilimitados, lo que sería sumamente problemático para la libertad de expresión.
- **Una obligación general de supervisar el contenido:** como se ha señalado anteriormente, no debe haber ninguna obligación general de supervisar el contenido, ni medidas que sean sustancialmente equivalentes, como exigir que se lleven a cabo «todos los esfuerzos» o se adopten «medidas proactivas» para abordar el contenido ilícito. Del mismo modo, dicho marco debería abstenerse de ejercer presión para que las empresas adopten tales medidas, formulándolas como puramente voluntarias o simplemente «recomendadas», cuando, en realidad, la falta de adopción podría dar lugar a fuertes sanciones.
- **Plazos indebidamente cortos:** no debe exigirse a las empresas que eliminen contenidos en plazos indebidamente cortos, en particular cuando el contenido en cuestión puede dar lugar a cuestiones de interpretación difíciles de resolver, como en el caso del contenido de «discurso de odio» o «terrorista». Los plazos de eliminación breves no facilitan que las empresas revisen las notificaciones con suficiente cuidado. Por consiguiente, promueven la eliminación ilícita de contenidos y no protegen la libertad de expresión. Además, la obligación de eliminar contenidos en un plazo breve puede

llevar a que las empresas destinen recursos a suprimir notificaciones, independientemente de su gravedad, o a que se centren en los contenidos que se han publicado en las últimas 24 horas ignorando otros más antiguos que podrán merecer más atención.¹⁷⁶

- **Establecimiento de objetivos de cumplimiento:** del mismo modo, la legislación o los reguladores no deben imponer objetivos de cumplimiento numéricos que puedan tener el efecto de alentar a las empresas a ampliar la definición del contenido que no permiten en su plataforma a fin de aumentar su tasa de cumplimiento.¹⁷⁷ Es decir, los objetivos numéricos alentarían la eliminación de cantidades cada vez mayores de contenido legítimo. Además, creemos que, en la medida en que los legisladores consideren varias métricas y umbrales para garantizar el cumplimiento, deben tener en cuenta hasta qué punto puede esperarse que la sociedad tolere un grado de riesgo de daños en línea, como lo hace en el mundo fuera de Internet.¹⁷⁸
- **Obligación de cooperar o denunciar contenidos ilícitos:** las obligaciones de cooperar imprecisas son problemáticas porque pueden implicar graves interferencias en los derechos de los usuarios y usuarias, como el acceso a sus datos por parte de las fuerzas del orden sin garantías suficientes. Al mismo tiempo, esta imprecisión permite a las empresas afirmar que han cooperado de otras maneras menos intrusivas. En resumen, es probable que sea difícil exigir el cumplimiento de tales obligaciones y, por lo tanto, no queda claro que sean necesarias. Las obligaciones de denunciar contenidos ilícitos probablemente llevarían a las empresas a prestar mucha atención a las notificaciones que reciben de contenido presuntamente ilegal, independientemente de su gravedad, y denunciarlo a la policía. Además, podrían disuadirlas de invertir en herramientas automatizadas para detectar contenido ilegal, si se les sancionara por tener conocimiento de la ilegalidad, o se considerara

que han incumplido sus obligaciones por no denunciar todo el contenido potencialmente ilegal que identifican automáticamente en sus redes. Ambos resultados serían lamentables y probablemente tendrían un impacto negativo en la libertad de expresión, ya que se denunciarían grandes cantidades de contenido legítimo.

- **Obligaciones de transmisión:** si bien creemos que las «obligaciones de transmisión», es decir, el requisito de que las plataformas publiquen contenidos que son legales pero que violan sus condiciones de servicio, pueden imponerse en circunstancias muy limitadas, en nuestra opinión, sería inapropiado imponerlas de forma general. En particular, las obligaciones de transmisión interferirían tanto en el derecho a la libertad de expresión como en el derecho a la propiedad de las plataformas de redes sociales.¹⁷⁹ El derecho a la libre expresión de las plataformas debe implicar que son libres de, por ejemplo, permitir o rechazar la publicación, en su sitio web, de teorías de conspiración de extrema derecha, aunque puedan ser legales en algunos países. Desde el punto de vista de la proporcionalidad, los usuarios y usuarias que deseen promover tales teorías tendrían otros canales disponibles para expresar esas opiniones. En la práctica, las obligaciones generales de transmisión también socavarían casi con certeza la moderación de contenidos, lo que sería negativo.

Recomendación 7:

Todos los reguladores deben ser independientes tanto en la ley como en la práctica

ARTICLE 19 considera que, para que cualquier régimen de regulación de contenidos en línea tenga algún tipo de legitimidad, debe estar supervisado por un regulador independiente, es decir, un regulador que no sea objeto de interferencias políticas o comerciales.

La independencia y la autonomía institucional de dicho regulador deben estar garantizadas y protegidas por ley, en particular mediante:

- una declaración clara de la política general de la plataforma general y de regulación de contenidos en línea;
- una definición clara de las competencia y responsabilidades del regulador;
- normas para los miembros; y
- acuerdos de financiación y rendición de cuentas al público a través de un organismo multilateral. El gobierno debe mantenerse al margen y no participar en ninguno de esos acuerdos de financiación ni organismos de rendición de cuentas.

En concreto, para que el regulador pueda proteger la libertad de expresión, la ley que lo crea debe incluir lo siguiente:

- disposiciones generales que destaquen la importancia de proteger la libertad de expresión, incluida la expresión que pueda escandalizar, ofender o perturbar;
- disposiciones que dejen en claro que la misión de cualquier regulador en este ámbito es proteger los derechos humanos, incluida la libertad de expresión;
- disposiciones que exijan que los reguladores auditen las decisiones de eliminación de contenidos y consideren hasta qué punto las empresas eliminan o restringen en exceso dichos contenidos, ya sea previa solicitud o por iniciativa propia; y
- disposiciones que dejen claro que las empresas no deben ser sancionadas por no eliminar contenido legal.

Los reguladores encargados de supervisar las operaciones de una amplia gama de proveedores de servicios digitales deben garantizar la cooperación con otros reguladores pertinentes, como las autoridades de protección de datos, protección del consumidor y competencia.

Por último, los reguladores deben rendir cuentas al público, también a través de obligaciones de transparencia sobre sus actividades e informes anuales presentados ante el Parlamento.

Recomendación 8:

Todo marco legislativo debe ser proporcionado

ARTICLE 19 considera que, para que cualquier marco legislativo cumpla con los estándares internacionales en materia de libertad de expresión, debe ser estrictamente proporcional al fin perseguido:

- **Enfoque escalonado:** los gobiernos deben ser extremadamente cautelosos a la hora de adoptar medidas destinadas a hacer que las grandes empresas de redes sociales rindan cuentas, pero que, en última instancia, impondrían una carga indebida a otros proveedores de servicios más pequeños. Por consiguiente, creemos que es necesario aplicar un enfoque escalonado en este ámbito, es decir, las grandes plataformas de redes sociales podrían estar sujetas a obligaciones más estrictas que los actores más pequeños. Para evaluar el dominio de una plataforma, se pueden tener en cuenta los siguientes factores: 1) su número de usuarios y usuarias; 2) su facturación total anual; y 3) su poder de mercado. Las organizaciones sin fines de lucro, como Wikipedia, deberían estar exentas y continuar operando con arreglo a un amplio marco de exención de responsabilidad.

Al mismo tiempo, recomendamos encarecidamente que cualquier medida propuesta sea objeto de evaluaciones de impacto rigurosas, que analicen también los posibles efectos que puedan restringir la competencia. Las grandes empresas de redes sociales probablemente se adapten a cualquier demanda que se les imponga y estas demandas, en última instancia, podrían dar lugar a una percepción de que son «más seguras», algo que con casi total certeza les daría una ventaja sobre los participantes más pequeños, que no podrían participar en el mismo tipo de actividades de moderación de contenidos que los operadores históricos.

- **Evaluación de las deficiencias sistémicas:** para que una legislación en este ámbito sea sostenible y proporcionada, no se debe evaluar a las empresas por no haber eliminaron un contenido o por haber publicado solo un conjunto de datos. En cambio, el regulador debe evaluar si han incumplido con sus obligaciones en virtud de la ley de manera sistémica. El umbral para determinar las deficiencias sistémicas debe definirse por ley con referencia a criterios claros que permitan una evaluación holística en lugar de una puramente numérica. Por ejemplo, la ley no debería sancionar a las empresas por no haber eliminado una determinada cantidad o porcentaje de contenido marcado como ilícito o nocivo. Más bien, debe prever una evaluación general de las medidas que han adoptado para mitigar los riesgos para los derechos humanos.
- **Sanciones proporcionadas:** el incumplimiento de las obligaciones descritas anteriormente debe penalizarse con sanciones proporcionadas. Aunque estas sanciones pueden incluir multas significativas, no deben ser tan altas como para desincentivar la protección de la libertad de expresión. En nuestra opinión, el 4 % de la facturación total probablemente sea una sanción demasiado alta para proteger la libertad de expresión. Asimismo, las sanciones penales impuestas a los directores ejecutivos

por incumplimiento de estas obligaciones podrían tener un efecto disuasorio sobre la libertad de expresión. Ante la perspectiva de varios años de prisión, los directivos de las empresas adoptarían, casi con seguridad, políticas que favorecerían una mayor eliminación de contenidos u otro tipo de restricciones. Por consiguiente, los gobiernos deben abstenerse de adoptar estas sanciones penales.

Recomendación 9:

Todo marco legislativo debe facilitar acceso a recursos efectivos

A parte de los mecanismos internos de reclamaciones, ARTICLE 19 considera que los gobiernos deben garantizar que las personas tengan acceso a recursos judiciales para impugnar la eliminación ilícita de sus contenidos por parte de las plataformas de redes sociales sobre la base de sus condiciones de servicio. Dichos recursos deben incluir no solo el acceso a los tribunales, sino también mecanismos alternativos de resolución de controversias, como tribunales electrónicos o el defensor del pueblo.¹⁸⁰ En la práctica, los gobiernos deben elaborar propuestas para financiar dichos mecanismos, por ejemplo, a través de un impuesto a las plataformas de redes sociales.

Sin perjuicio de los regímenes de autorregulación, como los consejos de redes sociales, que, entre otras cosas, permitirían a los usuarios y usuarias cuestionar las decisiones de moderación de contenidos de las plataformas de redes sociales en referencia a un conjunto acordado de principios, como una «Carta de derechos de los usuarios y usuarias».

Recomendación 10:

Se debe exigir a las grandes plataformas que separen sus funciones de alojamiento y selección de contenidos, y que garanticen su interoperabilidad con otros servicios

ARTICLE 19 considera que, para abordar el excesivo poder de mercado de un número reducido de plataformas de redes sociales, debe descentralizarse la curación de contenidos.

Los reguladores podrían obligar a las grandes plataformas a separar sus funciones de alojamiento y curación de contenidos, a fin de permitir que terceros accedan a su plataforma (en la práctica, su API) para prestar servicios de curación de contenidos a los usuarios y usuarias.¹⁸¹

Esta forma de separación funcional no impediría que las grandes plataformas de redes sociales ofrezcan servicios de curación de contenidos a sus usuarios y usuarias, pero serían ellos quienes decidirían si participar. Es decir, al crear, por ejemplo, un perfil en Facebook se debe solicitar al usuario o usuaria que seleccione un proveedor de servicios de curación de contenidos, y Facebook seguiría siendo una de las opciones que se pueden seleccionar. Lo ideal es que para evitar que los usuarios y usuarias permanezcan «atados» a un contrato, tengan la libertad de cambiar su elección en cualquier momento a través de la configuración de la plataforma. En nuestra opinión, se deben explorar más a fondo este tipo de soluciones a fin de permitir a los usuarios y usuarias recuperar el control, garantizar una competencia sana y la innovación en los mercados de las redes sociales, y retomarla promesa de un Internet diverso y descentralizado.¹⁸²

Como mínimo, creemos que las plataformas de redes sociales deben ofrecer a los usuarios y usuarias más opciones de selección de contenidos, en términos del tipo de contenido que les gustaría ver y en base a qué criterios, por ejemplo, en orden cronológico.



Notas a pie
de página

Notas a pie de página

(Haga clic en el número de nota final para leer en contexto)

- 1  John Perry Barlow, Declaración de Independencia del Ciberespacio, Electronic Frontier Foundation, 8 de febrero de 1996.
- 2  Los «proveedores de alojamiento» o «hosts» suelen ser quienes alquilan espacio en un servidor web para que sus clientes puedan configurar sus propios sitios web. Sin embargo, el término «host» también ha adquirido un significado más general, es decir, cualquier persona o empresa que controla un sitio o una página web que permite que terceros carguen o publiquen material. Por esta razón, las plataformas de redes sociales y los servicios que permiten compartir vídeos y fotos generalmente se denominan «hosts». Por tanto, la función de los proveedores de alojamiento o «hosting» se caracteriza por la ausencia de intervención editorial en el contenido; *c.f.* ARTICLE 19, Intermediarios de Internet: Disyuntiva de la aplicación de la responsabilidad civil y penal, 2013, p. 7.
- 3  Según estimaciones recientes, en la actualidad hay 2 500 millones de usuarios activos de Facebook y 330 millones de usuarios de Twitter cada mes, y cada día se ven 1 000 millones de horas de vídeo en YouTube. Véase, por ejemplo, Portal Statistics, Leading countries based on Facebook audience size (Países líderes según el tamaño de la audiencia de Facebook, información en inglés); Portal Statistics, Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2019 (Número de usuarios de Twitter activos mensuales en todo el mundo desde el primer trimestre de 2010 hasta el primer trimestre de 2019, información en inglés); y YouTube for Press, YouTube in numbers (YouTube en números, información en inglés).
- 4  Véase, por ejemplo, Twitter, Suspensión permanente de @realDonaldTrump, 8 de enero de 2021; y Guy Rosen Monika Bickert, Our response to the violence in Washington (Nuestra respuesta a la violencia en Washington, artículo en inglés), Facebook, 6 de enero de 2021.
- 5  K. Wagner y P. Martin, Twitter locks out Chinese Embassy in U.S. over post on Uighurs (Twitter bloquea a la embajada china en EE. UU. por una publicación sobre los uigures, artículo en inglés), Bloomberg, 20 de enero de 2021.
- 6  Véase, por ejemplo, la sección 230 de la Ley de decencia de las comunicaciones, de 1996 (US Code 47, sección 230) de EE. UU. o la Directiva 2000/31/CE del Parlamento Europeo y del Consejo, de 8 de junio de 2000, relativa a determinados aspectos jurídicos de los servicios de la sociedad de la información, en particular el comercio electrónico en el mercado interior (Directiva sobre el comercio electrónico).

- 7  Véase, por ejemplo, el Libro Blanco sobre la información nociva en línea, 2019, y el proyecto de Ley sobre la seguridad en línea, 2021, de Reino Unido; la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a un mercado único de servicios digitales (Ley de servicios digitales) y por el que se modifica la Directiva 2000/31/CE, 2020; la Ley de regulación de los medios de comunicación y la seguridad en línea de Irlanda, 2020; la Ley australiana sobre la seguridad en línea, 2021; y el proyecto de Directrices (modificadas) sobre Intermediarios, de la India, 2018.
- 8  Véase ARTICLE 19, #MissingVoices campaign.
- 9  See e.g. Online Harms White Paper: Full government response to the consultation (Libro Blanco sobre la información nociva en línea: Respuesta completa del gobierno sobre la consulta, documento en inglés), 15 de diciembre de 2020; o el proyecto de Ley de lucha contra la incitación al odio en línea (la denominada *Loi Avia* o *Ley Avia*), que posteriormente declaró inconstitucional el Tribunal Constitucional (Conseil d'État) de Francia.
- 10  Disyuntiva de la aplicación de la responsabilidad civil y penal, *op.cit.*
- 11  ARTICLE 19, *Sidestepping Rights: Regulating Speech by Contract* (Eludir los derechos: Regular la libertad de expresión mediante contrato), 2018.
- 12  *Ibid.*
- 13  La política también hace referencia a varias propuestas legislativas sobre la gobernanza de las plataformas que se han presentado recientemente en algunos países, ya que podrían utilizarse en otros lugares como inspiración para la adopción de legislación similar, en particular el proyecto de Ley sobre seguridad en línea, 2021, y la propuesta, de la UE, de Ley de Servicios Digitales [véase Comisión Europea, *The Digital Services Act package* (Paquete de la Ley de Servicios Digitales, información en inglés)]. Entendemos que estas propuestas han podido enmendarse o modificarse, y la presente política las cita en su forma en el momento de la publicación (diciembre de 2021).
- 14  ARTICLE 19, Cómo limitar el poder de las grandes tecnológicas: Proteger la libertad de expresión mediante la separación de servicios, el libre mercado, la competencia y los derechos de los usuarios y usuarias.
- 15  ARTICLE 19, informe de política sobre la diversidad de los medios de comunicación en línea, de próxima publicación.
- 16  ARTICLE 19, informe de política sobre la obligación de transmisión, de próxima publicación.
- 17  Entre ellas se incluyen, por ejemplo, las cuestiones relativas a la publicidad, los anuncios personalizados, y los modelos de negocio de las empresas y la libertad de información y de expresión.
- 18  Disyuntiva de la aplicación de la responsabilidad civil y penal, *op.cit.*

- 19  Véase, por ejemplo, L. Lovdahl Gormsen y J.T. Llanos, *Facebook's anticompetitive lean in strategies* (La inclinación anticompetitiva de Facebook en las estrategias, documento en inglés), SSRN, 6 de junio de 2019.
- 20  C.f., por ejemplo, Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, *Las plataformas en línea y el mercado único digital, Retos y oportunidades para Europa*, COM(2016)288 final, 25 de mayo 2016.
- 21  Véase también *Sidestepping rights* (Eludir los derechos), *op.cit.*, p. 36.
- 22  Véase, por ejemplo, *Disyuntiva por la atribución de responsabilidad civil y penal*, *op.cit.*
- 23  En el ámbito del derecho consuetudinario anglosajón, véase, por ejemplo, D. Rolph, *Liability in the Internet age* (La responsabilidad en la era de Internet), Blog de International Forum for Responsible Media (IFFRM), 17 de febrero de 2011.
- 24  Mediante su adopción en una resolución de la Asamblea General de la ONU, la DUDH no es estrictamente vinculante para los Estados. Sin embargo, se considera que muchas de sus disposiciones han adquirido fuerza legal como derecho internacional consuetudinario desde su adopción en 1948. Véase *Filartiga c. Pena-Irala*, 630 F. 2d 876 (1980) (Tribunal de Apelaciones del Circuito de EE. UU., segundo circuito).
- 25  Asamblea General de la ONU, Pacto Internacional de Derechos Civiles y Políticos, 16 de diciembre de 1966, Serie de los Tratados de las Naciones Unidas, vol. 999, p. 171 (en el documento en inglés).
- 26  Artículo 10 del Convenio Europeo para la Protección de los Derechos Humanos y las Libertades Fundamentales, 4 de septiembre de 1950; artículo 9 de la Carta Africana sobre los Derechos Humanos y de los Pueblos (Carta de Banjul), 27 de junio de 1981; y artículo 13 de la Convención Americana sobre Derechos Humanos, 22 de noviembre de 1969.
- 27  Comité de Derechos Humanos de la ONU, Observación general N.º 34 Artículo 19: *Libertad de opinión y libertad de expresión*, CCPR/C/GC/34, 12 de septiembre de 2011, párrafos 12, 17 y 39.
- 28  La Declaración Conjunta sobre la Libertad de Expresión e Internet de 2011, adoptada por el Relator Especial de la 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), junio de 2011.
- 29  *Ibid.* Véase también el Informe del Relator Especial sobre la Libertad de Expresión, A/66/290, 10 de agosto de 2011, párrafo 16.

- 30  Comité de Derechos Humanos, *Belichkin c. Belarus*, Comunicación N.º 1022/2001, UN Doc. CCPR/C/85/D/1022/2001 (2005).
- 31  Observación general N.º 34, *op.cit.*, párrafo 43.
- 32  Declaración Conjunta de 2011, *op. cit.*
- 33  Relator Especial sobre la Libertad de Expresión, Informe de 16 de mayo de 2011, A/HRC/17/27, párrafo 43.
- 34  Relator Especial sobre la Libertad de Expresión, Informe de 6 de abril de 2018, A/HRC/38/35, párrafo 66.
- 35  *Ibid.*
- 36  *Ibid.*
- 37  Relatora Especial sobre la Libertad de Expresión, Informe de 13 de abril de 2021, A/HRC/47/25, párrafo 91.
- 38  Principios Rectores sobre las Empresas y los Derechos Humanos: Puesta en práctica del marco de las Naciones Unidas para «proteger, respetar y remediar» (los Principios Rectores), elaborados por el Representante Especial del Secretario General para la cuestión de los derechos humanos y las empresas transnacionales y otras empresas, Informe del Representante Especial del Secretario General para la cuestión de los derechos humanos y las empresas transnacionales y otras empresas, John Ruggie, 7 de abril de 2008, A/HRC/8/5A/HRC/17/31. El Consejo de Derechos Humanos adoptó los Principios Rectores en su resolución 17/4 de 16 de junio de 2011.
- 39  *Ibid.*, principio 15.
- 40  *Ibid.*, Informe del Relator sobre la libertad de expresión, 6 de abril de 2018, A/HRC/38/35, párrafos 48 y 7; Relator Especial para la Libertad de Expresión de la OEA, Estándares para una Internet libre, abierta e incluyente, 15 de marzo de 2017, párrafos 111-112.
- 41  *Ibid.*, Informe del Relator Especial sobre la Libertad de Expresión de 2018, párrafos 45–46; Informe del Relator para la Libertad de Expresión de la OEA, de 2013, párrafo 113. Véase asimismo, Relator Especial para la Libertad de Expresión de la OEA, Estándares para una Internet libre, abierta e incluyente, 15 de marzo de 2017, párrafo 99.
- 42  *Ibid.*, Relator Especial sobre la Libertad de Expresión, párrafos 47 y 76; Informe de 2013 del Relator Especial para la Libertad de Expresión de la OEA, párrafo 115.
- 43  Informe de abril de 2018 del Relator Especial sobre la Libertad de Expresión, *op. cit.*, párrafos 40-44.
- 44  *Ibid.*, párrafos 70–72.

- 45  Comisario para los Derechos Humanos del Consejo de Europa, *The rule of law on the Internet and in the wider digital world* (El Estado de Derecho en Internet y en el mundo digital más amplio, CommDH/IssuePaper (2014) 1, 8 de diciembre de 2014.
- 46  *Ibid.*, p. 24.
- 47  Comité de Ministros del Consejo de Europa, *Recommendation CM/Rec(2012)4 of the Committee of Ministers to Member States on the protection of human rights with regard to social networking services* (Recomendación CM/Rec(2012)4 del Comité de Ministros de los Estados miembros sobre la protección de los derechos humanos en relación con los servicios de redes sociales, documento en inglés), *adoptada por el Consejo de Ministros el 4 de abril de 2012*. Estas recomendaciones se repitieron también en el Comité de Ministros, *Guide to human rights for Internet users, Recommendation CM/Rec(2014)6 and explanatory memorandum* (Guía de derechos humanos para usuarios de Internet, Recomendación CM/Rec(2014)6 y exposición de motivos, documento en inglés), p.4.
- 48  *Recommendation CM/Rec(2018)2 of the Committee of Ministers to Member States on the roles and responsibilities of Internet intermediaries* (Recomendación CM/Rec(2018)2 del Comité de Ministros de los Estados miembros sobre las funciones y responsabilidades de los intermediarios de Internet, documento en inglés), *adoptada por el Consejo de Ministros el 7 de marzo de 2018*.
- 49  Véase Informe de la OEA de 2013, *op. cit.* e Informe de la OEA de 2016, *op.cit.*
- 50  *Ibid.*
- 51  Véase la Declaration of Principles on Freedom of Expression and Access to Information in Africa (Declaración Africana de Principios sobre la Libertad de Expresión y el Acceso a la Información, documento disponible en inglés, francés, portugués y árabe), 2019, que la Comisión Africana de Derechos Humanos y de los Pueblos adoptó en su 65.º período ordinario de sesiones en Banjul, Gambia.
- 52  Declaration of Principles on Freedom of Expression and Access to Information in Africa (Declaración de Principios revisada sobre la Libertad de Expresión y el Acceso a la Información en África, documento disponible en inglés, francés, portugués y árabe). Véase, asimismo, la African Declaration of Internet Rights and Freedoms (Declaración sobre los Derechos y las Libertades en Internet, documento disponible en inglés, francés, portugués y árabe), 2014, elaborada por iniciativa de la sociedad civil.
- 53  Los principios de Manila sobre la Responsabilidad de los intermediarios (Los Principios de Manila), marzo de 2015.
- 54  *Ibid.*, principio 4.
- 55  *Ibid.*, principio 5(c).

- 56  Ranking Digital Rights, Corporate Accountability Index: 2015 Research Indicators (Índice de la Responsabilidad Corporativa: Indicadores de investigación de 2015, información en inglés), junio de 2015.
- 57  La Declaración conjunta sobre la libertad de expresión y el combate al extremismo violento, de 2016, adoptada por el Relator Especial de la ONU para la Libertad de Expresión, la Representante para la Libertad de los Medios de Comunicación de la OSCE, el Relator Especial de la OEA para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la CADHP, 4 de mayo de 2016.
- 58  La Declaración Conjunta sobre Libertad de Expresión y «Noticias Falsas» («Fake News»), Desinformación y Propaganda, de 2017, Relator Especial sobre la Libertad de Expresión, Representante para la Libertad de los Medios de Comunicación de la OSCE, Relator Especial para la Libertad de Expresión de la OEA y Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la CADHP, 3 de mayo de 2017.
- 59  Relatores Especiales de la ONU sobre la Libertad de Expresión y la Violencia contra las Mujeres, UN experts urge States and companies to address online gender-based abuse but warn against censorship (Expertos de la ONU instan a los Estados y las empresas a abordar el abuso de género en línea, pero advierten contra la censura, documento en inglés), comunicado de prensa conjunto, 8 de marzo de 2017.
- 60  Véase también el Informe del Relator Especial sobre la Libertad de Expresión, A/74/486, 9 de agosto de 2019, párrafo 57 d).
- 61  Declaración Conjunta de 2016, *op.cit.*, apartado 2 e).
- 62  Declaración Conjunta de 2017, *op.cit.*, apartado 4 b).
- 63  *Ibid.*, párrafo 58 b).
- 64  Informe del Relator Especial sobre la Libertad de Expresión, A/73/348, 29 de agosto de 2018.
- 65  *Ibid.*, párrafo 64.
- 66  *Ibid.*
- 67  *Ibid.*, párrafo 66.
- 68  *Ibid.*
- 69  *Ibid.*, párrafo 70.

- 70  Véase el artículo 17 del Pacto Internacional de Derechos Civiles y Políticos, *op.cit.* El Relator Especial de la ONU sobre la promoción y protección de los derechos humanos y las libertades fundamentales en la lucha contra el terrorismo ha argumentado que, al igual que las restricciones al derecho a la libertad de expresión en virtud del artículo 19, debe interpretarse que las restricciones del derecho a la privacidad en virtud del artículo 17 del Pacto Internacional de Derechos Civiles y Políticos están sujetas a la prueba de la admisibilidad de las limitaciones, que consta de tres elementos. Véase el Informe del Relator Especial sobre la promoción y protección de los derechos humanos y las libertades fundamentales en la lucha contra el terrorismo, Martin Scheinin, A/HRC/13/37, 28 de diciembre de 2009.
- 71  Informe del Relator Especial sobre la Libertad de Expresión de mayo de 2011, *op.cit.*, párrafo 53.
- 72  *Ibid.*, párrafo 84.
- 73  Informe del Relator Especial al Consejo de Derechos Humanos sobre la utilización del cifrado y el anonimato, y el ejercicio de la libertad de opinión y expresión en la era digital, A/HRC/29/32, 22 de mayo 2015, párrafo 60.
- 74  *Ibid.* Véase, asimismo, Relator Especial sobre la Libertad de Expresión, Encryption and anonymity follow-up report (Informe de seguimiento sobre el cifrado y el anonimato, informe en inglés), Research Paper 1/2018, junio de 2018.
- 75  Por ejemplo, en el tercer trimestre de 2020, Facebook registró 2 500 millones de usuarios y usuarias activos al mes en todo el mundo, casi el tamaño de la población total de China e India. En 2017, los ingresos anuales de Facebook fueron superiores al PIB de Serbia. Véase F. Belinchón y Q. Moynihan, 25 giant companies that are bigger than entire countries (25 empresas gigantes que son más grandes que países enteros, documento en inglés), Business Insider España, 25 de julio de 2018. Asimismo, en diciembre de 2019, Google tenía el 90 % de la cuota de mercado de los servicios de búsqueda en todo el mundo. Véase Statcounter, Search engine market share worldwide (Cuota de mercado de los motores de búsqueda en todo el mundo, información en inglés), abril de 2019–abril de 2020. Twitter, aunque es una empresa más pequeña, tiene más de 125 millones de usuarios y usuarias activos al día, 2,5 veces la población de Kenia.
- 76  Véanse, por ejemplo, las propuestas del Reino Unido sobre la información nociva en línea y Ley de seguridad en línea, *op. cit.*, el ahora suspendido proyecto de ley francés Loi Avia, *op. cit.*, o el proyecto de Ley de plataformas de comunicación de Australia, octubre de 2020.

- 77  Por ejemplo, la Ley francesa Loi Avia, que el Tribunal Constitucional declaró inconstitucional en junio de 2020, preveía que el gobierno estableciera umbrales por encima de los cuales se esperaba que las empresas cumplieran ciertas obligaciones, como la eliminación, en un plazo de 24 horas, del contenido ilícito que incitara al odio (véase la nota a pie de página 7). Asimismo, en Alemania, un proyecto de ley sobre la diversidad de los medios de comunicación establece una serie de obligaciones para las «plataformas de vídeo» que tengan más de un millón de usuarios y usuarias al mes en Alemania (propuesta sobre cuestiones relacionadas con las interfaces de radio, el control de las plataformas y los medios intermedios, los reglamentos estatales específicos relativos a la aprobación, la regulación de las plataformas y los intermediarios (los llamados *Medienstaatsvertrag*), 2018
- 78  Por ejemplo, la Directiva de la UE sobre derechos de autor, de 2019, establece que los prestadores de servicios para compartir contenidos en línea perderán la exención de responsabilidad si no cumplen una serie de condiciones estrictas. No obstante, el cumplimiento de estas nuevas obligaciones de la Directiva debe tener en cuenta «el tipo, la audiencia y la magnitud del servicio», de conformidad con el principio de proporcionalidad. Las empresas que han prestado servicios en la Unión Europea durante menos de tres años o cuyo volumen de negocios sea inferior a 10 millones EUR están sujetas a condiciones menos estrictas para beneficiarse de la exención de responsabilidad. Véase la [Directiva \(UE\) 2019/790](#) del Parlamento Europeo y del Consejo, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital y por la que se modifican las Directivas 96/9/CE y 2001/29/CE. En cambio, en el caso de los prestadores de servicios para compartir contenidos en línea que tienen un promedio de visitantes únicos mensuales que supera los cinco millones, la Directiva de la UE sobre derechos de autor, de 2019, prevé que estén sujetos a requisitos adicionales.
- 79  Ofcom, *Market Failure and Harm. An economic perspective on the challenges and opportunities in regulating online services* (Deficiencias y perjuicio del mercado. Una perspectiva económica sobre los desafíos y oportunidades de la regulación de los servicios en línea, documento en inglés), 28 de octubre de 2019; Australian Competition and Consumer Commission, *Digital Platforms Inquiry: Final report* (Investigación sobre las plataformas digitales; Informe final, documento en inglés), 26 de julio de 2019; y autoridades de competencia de Francia y Alemania, *Algorithms and competition* (Algoritmos y competencia, documento en inglés), informe conjunto, noviembre de 2019.
- 80  Por ejemplo, en 2017, Cloudflare decidió suspender la prestación de sus servicios a Daily Stormer, un foro de Internet de extrema derecha, neonazi y supremacista blanco, después de que un grupo de manifestantes contra la extrema derecha fuera arrollado por un coche, en Charlottesville, y una persona resultara muerta. En 2019, de nuevo Cloudflare suspendió la prestación de sus servicios a 8chan, un foro de Internet de extrema derecha, después de que el perpetrador del tiroteo en El Paso supuestamente publicara su manifiesto en ese sitio web. Véase M. Prince, *Why we terminated Daily Stormer* (Por qué hemos suspendido el servicio a Daily Stormer, blog en inglés), Cloudflare, 16 de agosto de 2017; M. Prince, *Terminating service for 8Chan* (Suspensión del servicio a 8Chan, blog en inglés), Cloudflare, 5 de agosto de 2019.

- 81  Véase, por ejemplo, Amnistía Internacional, Gigantes de la Vigilancia: La amenaza que el modelo de negocios de Google y Facebook representa para los derechos humanos, POL 30/1404/2019, noviembre de 2019; Privacy International, Privacy International's response to the open consultation on the Online Harms White Paper (Respuesta de Privacy International a la consulta relativa al Libro Blanco sobre la información nociva en línea, documento en inglés), julio de 2019.
- 82  *Ibid.*
- 83  Véase, por ejemplo, Iwańska, To track or not to track: Towards privacy-friendly and sustainable online advertising (Rastrear o no rastrear: hacia una publicidad en línea sostenible y respetuosa de la privacidad, documento en inglés), Panoptikon Foundation, noviembre de 2020.
- 84  *Ibid.*
- 85  *Ibid.*
- 86  Véase, por ejemplo, Access Now, Raising the alarm: Online tracking harms human rights (Dar la señal de alarma: el rastreo en línea vulnera los derechos humanos, documento en inglés) 14 de diciembre de 2020
- 87  Aunque no existe una definición de «discurso de odio» en el Derecho internacional de los derechos humanos, la expresión de odio hacia una persona o grupo sobre la base de una característica protegida se puede dividir en tres categorías, que se distinguen por la respuesta que el derecho internacional de los derechos humanos exige de los Estados: a) formas severas de «discurso de odio» que el derecho internacional requiere que los Estados prohíban; b) otras formas de «discurso de odio» que los Estados pueden prohibir; y c) el «discurso de odio» que es lícito pero que, no obstante, suscita preocupación en términos de intolerancia y discriminación, y amerita una respuesta crítica por parte del Estado, pero debe protegerse. Dada la complejidad de este término, ARTICLE 19 se refiere a él como «discurso de odio» en este y en otros documentos de política.
- 88  Por ejemplo, en Reino Unido, hubo una gran indignación pública cuando Molly Russell se quitó la vida y se encontraron, en su cuenta de Instagram, publicaciones gráficas sobre el suicidio y las autolesiones. Véase, por ejemplo, BBC News, Molly Russell: Did her death change social media? (*Molly Russell: ¿Ha cambiado su muerte las redes sociales?*, artículo en inglés), 27 de octubre de 2019.

89 

Por ejemplo, en el ámbito de la desinformación, la política de Facebook ha cambiado significativamente con el tiempo, de centrarse en el «comportamiento no auténtico» (es decir, el uso de cuentas falsas) a eliminar o reducir explícitamente la difusión de «noticias falsas» o «medios de comunicación manipulados» (por ejemplo, ultrafalsificaciones). Véase, por ejemplo, ARTÍCULO 19, *Facebook Community Standards: Analysis against international standards on freedom of expression* (Estándares comunitarios de Facebook: análisis con respecto a los estándares internacionales sobre libertad de expresión), 30 de julio de 2018. Twitter, por su parte, solo elimina una parte de la información engañosa en el contexto de las elecciones, pero continúa abordando la «manipulación de la plataforma» y el spam. Véase Política de integridad física, 2021, y Twitter, Política relativa al spam y la manipulación de la plataforma, septiembre de 2020. Por lo que respecta al «discurso de odio», las empresas de redes sociales han tenido que tomar medidas en respuesta a las graves críticas que han recibido por parte de varios grupos de derechos humanos, periodistas, legisladores e instituciones como la ONU. Por ejemplo, Amnistía Internacional ha denunciado que Twitter no protege los derechos en línea de las mujeres y crea un ambiente tóxico que las aleja del discurso público. Véase Amnistía Internacional, *Toxic Twitter: A toxic place for women* (Toxic Twitter: Un espacio tóxico para las mujeres), 15 de marzo de 2018. YouTube ha sido objeto de críticas por no tomar medidas contra el presentador conservador de YouTube, Steven Crowder, después de que este acosara durante dos años al periodista de Vox, Carlos Maza, incluso con insultos racistas y homófobos en vídeos de YouTube sobre él. Véase E. Stuart, *"We don't want to be knee-jerk": YouTube responds to Vox on its harassment policies* («No queremos ser viscerales»: La respuesta de YouTube a Vox sobre sus políticas de acoso, artículo en inglés), Vox, 10 de junio de 2019. Los investigadores de la ONU señalaron el «papel» de Facebook en la difusión del «discurso de odio», y, por lo tanto, en facilitar del posible genocidio del pueblo Rohingya en Myanmar. Véase el Informe al Consejo de Derechos Humanos de la misión internacional independiente de investigación sobre Myanmar, 12 de septiembre de 2018, [A/HRC/39/64](#), párrafo 74.

90 

Por ejemplo, en el marco de la UE, se incluyen el Código de buenas prácticas de la Unión en materia de desinformación, 17 de junio de 2019, y el Código de Conducta para la lucha contra la incitación ilegal al odio en Internet, 2018. A escala internacional, el Global Internet Counter-Terrorism Forum (GIFCT, Foro Global de Internet contra el Terrorismo), una iniciativa que pusieron en marcha Facebook, Microsoft, Twitter y YouTube, en 2017, permite a sus miembros acceder a una base de datos de «hashes» que los miembros pueden utilizar para decidir si eliminan contenidos en su propio servicio y en base a sus propias normas comunitarias. En el marco de *Christchurch Call*, una iniciativa global que el gobierno de Nueva Zelanda puso en marcha tras la transmisión en directo de los ataques terroristas en Christchurch, los proveedores de servicios en línea se comprometen, entre otras cosas, a «tomar medidas transparentes y específicas para evitar la descarga de contenido terrorista y extremista violento, así como su difusión en las redes sociales y servicios similares de intercambio de contenidos, incluida su eliminación inmediata y permanente».

91 

Algunos Estados, como Francia y Singapur, ya han considerado que estas medidas son insuficientes y han adoptado legislación que prohíbe la desinformación.

- 92  En la UE, también se afirma que la Ley de Servicios Digitales es necesaria para garantizar la eficiencia del mercado único; *op.cit.* En el Reino Unido, disponer de un solo regulador para hacer frente a los diversos problemas que plantea el contenido en línea tiene el atractivo de la aparente simplicidad. Véanse las propuestas sobre la información nociva en línea y la Ley de seguridad en línea, *op.cit.*
- 93  Véase Secretario de Estado francés para Asuntos Digitales, *Creating a French framework to make social media platforms more accountable: Acting in France with a European vision* (Crear un marco francés para que las plataformas de redes sociales rindan cuentas: Actuar en Francia con una visión europea, documento en inglés y francés), Informe de la misión, versión 1.1, mayo de 2019, p. 11; ARTICLE 19, *France: ARTICLE 19 comments on French interim report for social media regulation* (Observaciones de ARTICLE 19 sobre el informe provisional relativo a la regulación en materia de redes sociales, documento en inglés), 19 de junio de 2019.
- 94  Por ejemplo, en EE. UU., el Senador Hawley ha tratado de introducir legislación para modificar la sección 230 de la Ley de decencia en las comunicaciones con objeto de que no se aplique la exención de responsabilidad a las plataformas a menos que una auditoría externa demuestre de manera convincente que sus algoritmos y prácticas de eliminación de contenido son políticamente neutrales. Véase *Senator Hawley introduces legislation to amend Section 230 immunity for Big Tech companies* (El Senador Hawley presenta legislación para modificar la exención de responsabilidad de las grandes empresas tecnológicas prevista en la sección 230, artículo en inglés), 19 de junio de 2019. Aunque el proyecto de ley no se ha aprobado, el debate en torno a la sección 230 todavía sigue candente en EE. UU. Por ejemplo, en enero de 2020, Joe Biden afirmó que quería derogarla, Véase, por ejemplo, Makena Kelly, *Joe Biden wants to revoke Section 230* (Joe Biden quiere derogar la sección 230, artículo en inglés), The Verge, 17 de enero de 2020. En Alemania, la Autoridad de Radiodifusión ha presentado una propuesta que impondría, por primera vez, obligaciones en materia de diversidad a las plataformas de vídeo y redes sociales. Véase N. Helberger, P. Leerssen y M. Van Drunen, *Germany proposes Europe's first diversity rules for social media platforms* (Alemania propone las primeras normas de Europa en materia de diversidad para las plataformas de redes sociales, artículo en inglés), Blog de LSE, 29 de mayo de 2019.
- 95  Véase, por ejemplo, E. M. Mazzoli y D. Tambini, *Prioritisation uncovered: The discoverability of public interest content online* (La priorización al descubierto: El descubrimiento del contenido de interés público en línea, documento en inglés), Estudio del Consejo de Europa DGI (2020)19, p. 40ff.
- 96  Para más información sobre los retos que plantea la distribución de noticias en las grandes plataformas de redes sociales, véase P.F. Docquir y M.L. Stasi, *The Decline of Media Diversity – and How we can Save it* (El declive de la diversidad de medios de comunicación y qué se puede hacer para preservarla, documento en inglés), 21 de enero de 2020.
- 97  The Guardian, *How social media filter bubbles and elections and algorithms influence the election* (Cómo las burbujas de filtro, las decisiones y los algoritmos de las redes sociales influyen en las elecciones, artículo en inglés), 22 de mayo 2017.

- 98  The New York Times, *Conservatives accuse Facebook of Political Bias* (Los conservadores acusan a Facebook de parcialidad política, artículo en inglés), 10 de mayo de 2016.
- 99  Véase, por ejemplo, The Guardian, *Watchdog cracks down on tech firms that fail to protect children* (El organismo de control adopta medidas enérgicas contra las empresas tecnológica que no protegen a la infancia, artículo en inglés), 22 de enero 2020.
- 100  Véase, por ejemplo, D. Keller y J. von Hoboken, *Design Principles for Intermediary Liability Laws* (Principios para la elaboración de leyes sobre la responsabilidad de los intermediarios, informe en inglés), Grupo de trabajo transatlántico de alto nivel sobre la moderación de contenidos, octubre de 2019.
- 101  Véase ARTICLE 19, *Germany: The Act to Improve Enforcement of the Law in Social Networks* (Alemania: La Ley para mejorar la aplicación de la legislación sobre redes sociales, documento en inglés), agosto de 2017; y ARTICLE 19, *France: Analysis of draft hate speech bill* (Francia: Análisis del proyecto de ley sobre discurso de odio, documento en inglés), 3 de julio de 2019.
- 102  Por ejemplo, iniciativas como el Foro Global de Internet contra el Terrorismo (GIFCT, por sus siglas en inglés) se han mantenido sumamente opacas. Hay pocos datos disponibles sobre el tipo de contenido que se elimina mediante la base de datos de «hashes» del GIFCT. Véase la página de transparencia en el sitio web de GIFCT. En Alemania, las opiniones siguen estando divididas sobre la eficacia de la Ley NetzDG, cuyo objetivo principal es abordar el «discurso de odio» en línea. Los expertos han señalado las dificultades para llevar a cabo una evaluación objetiva de la ley dada la falta de uniformidad de los informes sobre transparencia de las empresas de redes sociales. Sin embargo, existe cierto grado de consenso con respecto a que la nueva ley ha contribuido a una eliminación más rápida del contenido, aunque sigue sin estar claro su impacto real en el «discurso de odio» y es difícil probarlo empíricamente. Véase, por ejemplo, A. Held, *Germany is amending its online speech act NetzDG...but not only that* (Alemania está modificando su Ley NetzDG sobre el discurso en línea... pero no solo eso, documento en inglés), 6 de abril de 2020.
- 103  Por ejemplo, en Francia, un grupo de activistas LGBTI escribió recientemente una carta abierta en el periódico Libération en la que afirmaba que el proyecto de ley francés sobre el discurso de odio en línea (Ley Avia) sería contraproducente y daría lugar a una mayor eliminación de contenidos producidos por las comunidades LGBTI. El grupo exigía más recursos para que el sistema de justicia persiga a los autores reales de delitos de odio, como los insultos racistas. Véase Libération, *Feministas, LGBTI y antirracistas: no queremos la ley contra el discurso de odio en línea*, Tribune, 21 de enero de 2020.
- 104  Entre ellas se incluyen, por ejemplo, la Ley NetzDG de Alemania y la Loi Avia.
- 105  Véanse, por ejemplo, las actualizaciones recientes de Facebook (en inglés) y YouTube, *Our ongoing work to tackle hate* (Trabajo en curso para hacer frente al odio, artículo en inglés), 5 de junio de 2019.

- 106  *Ibid.* Véase también Wired, [Twitter and Instagram unveil new ways to combat hate – again](#) (Twitter e Instagram revelan, otra vez, nuevas formas de combatir el odio, artículo en inglés), 7 de noviembre de 2019.
- 107  Véase, por ejemplo, Google, informe elaborado por la propia empresa [Google News Initiative, one year in](#) (La iniciativa Google News después de un año, artículo en inglés), 20 de marzo de 2019.
- 108  Véase, por ejemplo, la Ficha del Código de Conducta de la UE para la lucha contra la incitación ilegal al odio en Internet (en inglés), febrero de 2019.
- 109  The Verge, [Twitter rolls out ‘hide replies’ to let you tame toxic discussions](#) (Twitter introduce «ocultar respuestas», que permitirá a los usuarios y usuarias controlar los debates tóxicos, artículo en inglés), 19 de septiembre de 2019.
- 110  Para más información sobre las deficiencias de estos mecanismos de apelación, véase ARTICLE 19, [Sidestepping Rights](#) (Eludir los derechos), *op. cit.* y [#Missing Voices Campaign](#).
- 111  Véase, por ejemplo, Facebook, [An Independent Assessment of the Human Rights Impact of Facebook in Myanmar](#) (Una evaluación independiente del impacto de Facebook en los derechos humanos, en Myanmar, artículo en inglés), 5 de noviembre de 2018; y Facebook, [A Second Update on Our Civil Rights Audit](#) (Segunda actualización de nuestra auditoría de derechos humanos, artículo en inglés), 30 de junio de 2019.
- 112  Véase, por ejemplo, Forbes, [Facebook is Hiring 3,000 Moderators In Push To Curb Violent Videos](#) (Facebook contrata a 3 000 moderadores en un intento por frenar los vídeos violentos, artículo en inglés), 3 de mayo de 2017; Financial Times, [Facebook steps up efforts to combat hate speech in Myanmar](#) (Facebook intensifica sus esfuerzos para combatir el discurso de odio en Myanmar, artículo en inglés), 16 de agosto 2018; o Telegraph, [Facebook to Hire 1000 new workers in the UK to help fight toxic content](#) (Facebook contratará a 1 000 nuevos empleados en Reino Unido para ayudar a luchar contra el contenido tóxico, artículo en inglés), 21 de enero de 2020.
- 113  Por ejemplo, en la UE, la Comisión Europea señaló en su primer informe de seguimiento de enero de 2019 que supervisaría la puesta en práctica por parte de las empresas de redes sociales de sus compromisos en virtud del Código de buenas prácticas de la Unión en materia de desinformación, que alentaron las políticas elaboradas por las empresas para hacer frente a la desinformación, pero señaló estar profundamente preocupada por la falta de valores de referencia para medir el progreso. Véase Comisión Europea, [First monthly intermediate results of the EU Code of Practice against disinformation](#) (Primeros resultados intermedios mensuales del Código de buenas prácticas de la UE en materia la desinformación, documento en inglés), 28 de febrero de 2019. En su último informe de seguimiento de mayo de 2019, la Comisión Europea señaló el progreso logrado en la mejora de la **transparencia de la publicidad política** y la divulgación pública de dichos anuncios, pero agregó que se debe hacer más para hacer frente a la desinformación de manera efectiva. Véase Comisión Europea, [Last intermediate results of the EU Code of Practice against disinformation](#) (Últimos resultados intermedios del Código de buenas prácticas de la Unión en materia de desinformación, documento en inglés), 14 de junio de 2019.

- 114  Véase, por ejemplo, King's College London Centre for the Study of Media, Communication and Power, Submission to the Culture Media and Sport Select Committee's Inquiry into Fake News (Informe para la investigación del Comité Selecto de Medios de Comunicación, Cultura y Deportes sobre las noticias falsas, documento en inglés), 2017.
- 115  Sin perjuicio de la posición más detallada de ARTICLE 19 sobre la publicidad política y la desinformación, de próxima publicación.
- 116  Véase, por ejemplo, B. Wagner y L. Kuklis, Disinformation, data verification and social media (Desinformación, verificación de datos y redes sociales, documento en inglés), 16 de enero de 2020.
- 117  Por ejemplo, en su cuarta evaluación del Código de buenas prácticas de la Unión en materia de desinformación, la Comisión Europea señaló que esta iniciativa había dado resultados positivos. En particular, destacó que ahora las empresas informáticas pueden evaluar el 89 % de los contenidos señalados en un plazo de 24 horas y retiran el 72 % de los contenidos que se consideran constitutivos de incitación ilegal al odio, en comparación con el 40 % y el 28 %, respectivamente, cuando el Código se puso en marcha en 2016. Véase Comisión Europea, op. cit. Véase, asimismo, Comisión Europea, Lucha contra la incitación ilegal al odio en línea: el Código de conducta de la UE garantiza una rápida respuesta, 4 de febrero de 2019. Asimismo, en Francia se constató que «la velocidad de implementación y el progreso realizado durante los últimos 12 meses por un operador como Facebook muestran los beneficios de capitalizar este enfoque de autorregulación que ya utilizan las plataformas, ampliándolo y legitimándolo» y que «la capacidad de autorregulación observada en estos operadores de redes sociales permite posicionarlos como elementos clave en la solución de los problemas de cohesión social que plantea la presencia de determinados contenidos en estas plataformas». Véase Secretario de Estado francés para Asuntos Digitales, Creating a French framework to make social media platforms more accountable: Acting in France with a European vision (Crear un marco francés para hacer que las plataformas de redes sociales rindan cuentas: Actuar en Francia con una visión europea, documento en inglés y francés), mayo de 2019, p. 11.
- 118  Para más información sobre los consejo de redes sociales, véase ARTICLE 19, Self-regulation and 'hate speech' on social media platforms (Autorregulación y «discurso de odio» en las plataformas de redes sociales, documento en inglés), 2 de marzo de 2018; ARTICLE 19, Social Media Councils: Consultation (Consulta sobre los consejos de redes sociales, documento en inglés), 11 de junio de 2019.
- 119  Véase la exposición de motivos de la propuesta de la EU de una Ley de servicios digitales, op. cit., p. 1.
- 120  Véase Comisión Europea, nota filtrada sobre la Ley de Servicios Digitales, Netzpolitik, junio de 2019.
- 121  Observación general N.º 34, op.cit., párrafo 43.

- 122  Véase, por ejemplo, R. Fletcher, *Polarisation in the news media* (La polarización en los medios de comunicación, artículo en inglés), Digital News Report, Instituto Reuters y Universidad de Oxford, 2017.
- 123  Véase la política de ARTICLE 19 sobre la publicidad política, de próxima publicación.
- 124  Para tener una visión general sobre varias propuestas legislativas, véase ARTICLE 19, *Kyrgyzstan: Platform regulation laws and freedom of expression* (Kirguistán: Las empresas de redes sociales y la libertad de expresión, documento en inglés), enero de 2021.
- 125  Véase, por ejemplo, las propuestas de la UE sobre el proyecto de Ley de Servicios Digitales, *op. cit.*; el proyecto de Ley de plataformas de comunicación de Australia, *op. cit.*; o la ahora suspendida Loi Avia, *op. cit.* Este es también uno de los aspectos positivos de la Ley alemana NetzDG, Véase, por ejemplo, B. Wagner, K. Rozgonyi, M.T. Sekwenz, J. Cobbe y J. Singh, *Regulating Transparency?: Facebook, Twitter and the German Network Enforcement Act*, (*¿Regular la transparencia?: Facebook, Twitter y la Ley alemana para mejorar la aplicación de la legislación sobre redes sociales*, documento en inglés), enero de 2020.
- 126  Véanse, por ejemplo, las propuestas de la UE sobre la Ley de Servicios Digitales, *op. cit.* o el proyecto de Ley de plataformas de comunicación de Australia, de 2020, *op. cit.*
- 127  *Ibid.*
- 128  La cláusula de «buen samaritano» es una disposición que garantiza que no se penalice con la pérdida de exención de responsabilidad a las plataformas u otros proveedores de alojamiento de contenidos si toman medidas para moderar el contenido por su iniciativa propia. Véase, por ejemplo, J. Barata, *Positive Intent Protections: Incorporating a Good Samaritan principle in the EU Digital Services Act* (Protección de la intención positiva: la incorporación del principio de buen samaritano en la Ley de Servicios Digitales de la UE, documento en inglés), The Center for Democracy and Technology, julio de 2020.
- 129  Véase, por ejemplo, la propuesta de la UE sobre la Ley de Servicios Digitales, *op. cit.* o la ahora suspendida Loi Avia, *op. cit.* También es una característica importante de la Ley alemana NetzDG. Véase, por ejemplo, *ARTICLE 19's analysis of the Draft NetzDG law* (Análisis de ARTICLE 19 sobre el proyecto de Ley NetzDG, documento en inglés), 2017.
- 130  Véase, por ejemplo, la propuesta de la UE sobre la Ley de Servicios Digitales, las propuestas de Reino Unido sobre la información nociva en línea y la Ley de seguridad en línea, la Directiva (UE) 2019/790 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital y por la que se modifican las Directivas 96/9/CE y 2001/29/CE (Directiva de la UE sobre derechos de autor, de 2019), la ahora suspendida Loi Avia sobre el discurso de odio en línea, *op. cit.* o la más reciente Ley de plataformas de comunicación de Australia, de 2020, *op. cit.* También es una característica importante de la Ley alemana NetzDG, *op. cit.*

- 131  Véanse las propuestas de Reino Unido, Irlanda y Australia y la nota filtrada sobre la Ley de Servicios Digitales, *op.cit.*
- 132  *Ibid.*, la propuesta de Irlanda.
- 133  Véanse las propuestas del Reino Unido sobre la información nociva en línea y la Ley de seguridad en línea, *op. cit.*
- 134  Véase, por ejemplo, la Ley NetzDG de Alemania o la Loi Avia de Francia, *op.cit.*
- 135  Véanse, por ejemplo, las propuestas del Reino Unido sobre la información nociva en línea y Ley de seguridad en línea, *op. cit.*
- 136  Por ejemplo, con arreglo a la propuesta de Irlanda, se deja a discreción del Comisionado de Seguridad en Línea la designación de «cualquier servicio en línea que permita a los usuarios compartir, difundir o acceder a contenido que otros usuarios han puesto a disposición», *op.cit.*
- 137  Véase la propuesta de informe de la misión sobre Facebook, de Francia, *op.cit.*
- 138  Véanse las propuestas de Reino Unido sobre la información nociva en línea y la Ley de seguridad en línea, *op. cit.*
- 139  Véase Gobierno de Reino Unido, Online Harms White Paper - Initial consultation response (Libro Blanco sobre la información nociva en línea: Respuesta a la consulta inicial, documento en inglés), febrero de 2020.
- 140  Véase ARTICLE 19, Response to the Consultations on the White Paper on Online Harms (Respuesta a las consultas relativas al Libro Blanco sobre la información nociva en línea, documento en inglés), junio de 2019.
- 141  C.f. propuesta de la India, *op.cit.*
- 142  Los Principios Rectores, *op.cit.*
- 143  *Ibid.*, principio IV.
- 144  Véanse, por ejemplo, las propuestas de Francia, Irlanda, Reino Unido y Australia, *op.cit.*
- 145  Véanse, por ejemplo, las propuestas de Francia y Australia, *op.cit.*
- 146  Véase, por ejemplo, la ahora suspendida Loi Avia de Francia o la Ley NetzDG de Alemania, *op.cit.*
- 147  Véase, por ejemplo, la ahora suspendida Loi Avia de Francia, *op.cit.*

- 148  Por ejemplo, ARTICLE 19 formó parte de una misión internacional conjunta sobre la libertad de prensa en Hungría que concluyó que el Consejo de Medios de Comunicación no es independiente y que todos sus miembros son nombrados por el partido Fidesz. Véase ARTICLE 19, *Hungary: Conclusions of the Joint International Press Freedom Mission* (Hungría: Conclusiones de la Misión Internacional Conjunta sobre la Libertad de Expresión, documento en inglés), 3 de diciembre de 2019. Véase, asimismo, ARTICLE 19 y Polish Helsinki Foundation for Human Rights, *Poland: independence of public service media* (Polonia: La independencia de los medios de comunicación públicos, documento en inglés), 31 de enero de 2017.
- 149  Por ejemplo, la Ley NetzDG no define qué constituyen incumplimientos «sistémicos» de las obligaciones con arreglo a la ley. Nos preocupan además las facultades que se propone conceder a los reguladores para ordenar el bloqueo de sitios web y la criminalización de los altos directivos por no cumplir con un vago «deber de diligencia», *op.cit.*
- 150  Por ejemplo, la policía india allanó las oficinas de Twitter en Delhi, India, después de que la empresa etiquetara los tuits de miembros del partido gobernante como medios «manipulados». Véase BuzzFeed, *Police In Delhi Have Descended On Twitter Headquarters In The Country* (La policía de Delhi irrumpe en la sede de Twitter en el país, artículo en inglés), 24 de mayo de 2021.
- 151  Por ejemplo, los enfoques reglamentarios propuestos en Europa pueden no ser fácilmente transferibles a otros lugares ni adecuados en ellos.
- 152  Véase, Disyuntiva de la aplicación de la responsabilidad civil y penal, *op.cit.*
- 153  C.f. los debates sobre la Ley de Servicios Digitales en la UE, *op.cit.*
- 154  Esto significa que no se debe sancionar a empresas como Cloudflare por alojar sitios como 8chan o DailyStormer; *op.cit.*
- 155  Disyuntiva de la aplicación de la responsabilidad civil y penal, *op.cit.*
- 156  Véase, por ejemplo, la Internet Watch Foundation en el Reino Unido o SaferNet en Brasil.
- 157  Los Principios de Manila, *op.cit.*
- 158  Esto además concuerda con la jurisprudencia de la mayoría de los Estados miembros de la UE. Véase, por ejemplo, Comisión Europea, *Overview of the notice and action procedures in EU member states* (Resumen de los procedimientos de notificación y actuación en los Estados miembros de la UE, documento en inglés), SMART 2016/0039, julio de 2018.
- 159  Véase *L'Oreal c. eBay*, C-324/09, 12 de julio de 2011.

- 160  Para un debate detallado sobre las tecnologías de moderación de contenidos, véase Center for Democracy and Technology, *Mixed Messages: the limits of automated social media content analysis* (Mensajes contradictorios: los límites del análisis automatizado de los contenidos en las redes sociales, documento en inglés), noviembre de 2017.
- 161  C.f. la Directiva de la UE sobre derechos de autor en el mercado único digital.
- 162  D. Keller y J. von Hoboken, *op.cit.*
- 163  *Ibid.*
- 164  Para un análisis más detallado de un enfoque de supervisión y eliminación de imágenes de abuso infantil respetuoso de los derechos humanos, véase Prostasia, *Child protection and the limits of censorship* (La protección de la infancia y los límites de la censura, documento en inglés), julio de 2018.
- 165  Para un análisis de las implicaciones del asunto *Eva Glawischnig-Piesczek c. Facebook Ireland Limited*, C-18/18, véase G. Smith, *Notice and stay down order and impact on online platform* (Orden de notificación y bloqueo y su impacto en las plataformas en línea, artículo en inglés) Bird & Bird, octubre de 2019. Véase también D. Keller, *The CJEU's new filtering case, the Terrorist Content Regulation, and the Future of Filtering Mandates in the EU* El nuevo asunto sobre filtrado del TJUE, la regulación de contenidos terroristas y el futuro de los mandatos de filtrado en la UE, artículo en inglés), blog del Center for Internet and Society, 2 de diciembre de 2019.
- 166  En la propuesta de ley alemana sobre las normas en materia de diversidad para las plataformas de redes sociales, se puede encontrar un ejemplo de obligaciones de transparencia e inspiración sobre este tipo de medidas. Véase *Germany proposes Europe's first diversity rules for social media platforms* (Alemania propone las primeras normas de Europa en materia de diversidad para las plataformas de redes sociales, artículo en inglés). *op. cit.*
- 167  M. MacCarthy, *Transparency requirements for digital social media platforms: Recommendations for policy makers and industry* (Requisitos de transparencia para las plataformas de redes sociales digitales: Recomendaciones para los responsables políticos y el sector, documento en inglés), Documento de trabajo, Grupo de trabajo transatlántico de alto nivel sobre moderación de contenidos y libertad de expresión, 12 de febrero de 2020.
- 168  C.f. Informe de la misión sobre Facebook, de Francia, *op.cit.*
- 169  Las empresas tienden a describirlo como la «prevalencia» del contenido, es decir, el número de veces que se ha visto.
- 170  Los Principios de Santa Clara sobre la transparencia y la responsabilidad en la moderación de contenidos, 2017.

- 171  Véase, por ejemplo, La Comisión Europea y las autoridades de protección de los consumidores de los Estados miembros solicitan a las empresas del sector de los medios de comunicación social que cumplan la normativa de protección de los consumidores de la UE, 17 de marzo de 2017.
- 172  M. MacCarthy, *op.cit.*
- 173  C.f. *Sidestepping Rights* (Eludir los derechos), *op.cit.*
- 174  Véase la política de ARTICLE 19 sobre la diversidad de los medios de comunicación en el ecosistema digital, de próxima publicación.
- 175  Véase la política de ARTICLE 19 sobre la libertad de expresión y las elecciones, de próxima publicación.
- 176  Para obtener más información sobre el debate acerca del impacto de los objetivos de cumplimiento, véase Facebook, *Charting a way forward, Online content regulation* (Trazar el camino a seguir, La regulación del contenido en línea, documento en inglés), febrero de 2020.
- 177  *Ibid.*
- 178  Véase, por ejemplo, V. Baines, *On online harms and folk devils: Careful now* (Cuidado con los daños en línea y los demonios populares, artículo en inglés), 24 de junio de 2019.
- 179  C.f. Tribunal Europeo de Derechos Humanos, *Appleby y otros c. Reino Unido*, App. N.º 44306/98, 6 de mayo de 2003.
- 180  Véase, en particular, H. Tworek, R. Ó Fathaigh, L. Bruggeman y C. Tenove, *Dispute resolution and content moderation: Fair, accountable, independent, transparent, and effective* (Resolución de controversias y moderación de contenidos justas, responsables, independientes, transparentes y efectivas, documento en inglés), Documento de trabajo, Grupo de trabajo transatlántico de alto nivel sobre la moderación de contenidos en línea y la libertad de expresión, enero de 2020.
- 181  Véase, por ejemplo, H. Feld, *The Case for the Digital Platform Act: Market structure and regulation of digital platforms* (Los argumentos a favor de una Ley de plataformas digitales: La estructura de mercado y la regulación de las plataformas digitales, documento en inglés), mayo de 2019.
- 182  Para más información, véase M. Masnick, *Protocols, not platforms: A technological approach to Free speech* (Protocolos, no plataformas: Un enfoque tecnológico de la libertad de expresión, documento en inglés), 21 de agosto de 2019; P.F. Docquir y M.L. Stasi, *The Decline of Media Diversity – and How we can Save it* (El declive de la diversidad en los medios de comunicación y qué se puede hacer para preservarla, documento en inglés), 21 de enero de 2020; M.L. Stasi, *Ensuring pluralism in social media markets: Some suggestions* (Garantizar el pluralismo en los mercados de redes sociales: Algunas sugerencias, documento en inglés), EUI Working Papers, RSCAS 2020/05, febrero de 2020.



**Defender la libertad de expresión
e información**

✉ info@article19.org

🌐 www.article19.org

🐦 [@article19org](https://twitter.com/article19org)

💬 [@article19](https://messenger.com/article19)

🌐 www.linkedin.com/company/article19/

📘 facebook.com/article19org

978-1-910793-49-7

© ARTICLE 19, 2023