



La protección del espacio cívico en línea

Un repaso a las amenazas actuales a la libertad de expresión en internet

Asociación por los Derechos Civiles



Con el apoyo de:



Julio 2021

adc.org.ar

Redacción: Eduardo Ferreyra
Diagramación: Matías Chamorro
Diseño de tapa: El Maizal



La protección del espacio cívico en línea. Un repaso a las amenazas actuales a la libertad de expresión en internet se publica bajo una licencia Creative Commons Atribución-NoComercialCompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0). Para ver una copia de la misma visite: creativecommons.org/licenses/by-nc-sa/4.0/deed.es

Contenidos

Resumen Ejecutivo | 4

Introducción | 5

La criminalización del discurso en línea | 7

El monitoreo de posteos en redes sociales (ciberpatrullaje) | 12

La influencia privada en la toma de decisiones sobre moderación de contenido | 16

Las expresiones de figuras políticas y la violencia | 19

Recomendaciones preliminares | 22

Notas | 27

Resumen ejecutivo

La Asociación por los Derechos Civiles (ADC) tiene la misión de defender los derechos fundamentales mediante la contribución al diseño de políticas por parte del sector público y privado. Las respuestas de los gobiernos deben promover un espacio público amplio y robusto en internet. Este deber implica asegurar la libertad de expresión de los usuarios, garantizar condiciones favorables para la creación de grupos online (libertad de asociación) y permitir el ejercicio del derecho de protesta en plataformas (libertad de reunión).

Al mismo tiempo, las empresas privadas también tienen la obligación de contribuir a dicha finalidad. Las grandes plataformas poseen un amplio poder para determinar las condiciones en que el discurso puede circular a través de internet. La influencia de estas entidades es incluso mayor que la de muchos países y, por lo tanto, su accionar puede impedir seriamente la circulación de ideas y opiniones. De este modo, el sector privado también se encuentra sujeto al deber de respetar los derechos humanos.

La crisis sanitaria producida por el Covid-19 y la intensificación de la polarización política alrededor del mundo -con la situación de Estados Unidos como ejemplo central- ha atraído la atención de la ciudadanía acerca del rol de los Estados y las plataformas para abordar fenómenos como la desinformación y el discurso de odio. Las acciones adoptadas por gobiernos y compañías deben ser sometidas a la más estricta evaluación, ya que ellas influenciarán la forma en que el debate público se llevará a cabo en el futuro.

Las recomendaciones que se brindarán tienen el objetivo de colaborar en la construcción de una respuesta que garantice el fortalecimiento del Estado de derecho, la vigencia de una cultura democrática y el disfrute de un espacio cívico online inclusivo.

Introducción

Las crisis suelen ser oportunidades para que surjan nuevas regulaciones por parte de Estados y grandes compañías. Esta vez no ha sido la excepción. Por un lado, la explosión de la pandemia del coronavirus generó preocupación acerca de los efectos que la desinformación puede ocasionar en la salud pública de la población. Por otro lado, la crisis institucional sufrida por los Estados Unidos puso la atención en los mensajes de odio, discriminación o violencia que provienen ya no solamente de personas comunes y corrientes sino de figuras políticas en todo el mundo. Ambos fenómenos dieron lugar a variadas respuestas por parte de autoridades públicas y las plataformas. Como hemos dicho, estas medidas nacen como respuesta a hechos extraordinarios pero es más que probable que continúen vigentes luego del retorno a la normalidad. Y si no, los precedentes que dejan servirán como anticipo de tendencias regulatorias en el futuro. En cualquier caso, el desafío de fondo persiste. Las grandes empresas han adquirido un poder tan grande que varios aspectos de su comportamiento entran en tensión con algunas de nuestras intuiciones sobre cómo debería funcionar una democracia. Al mismo tiempo, los Estados utilizan la coyuntura para avanzar en políticas que fomentan la vigilancia y la represión de voces.

El presente documento se enfocará en cuatro aspectos de este panorama:

- La criminalización del discurso online.
- El monitoreo de posteos en redes sociales (ciberpatrullaje).
- La toma de decisiones sobre moderación de contenido en las plataformas.
- Las expresiones de funcionarios políticos y la violencia.

Analizaremos estas tendencias en primer lugar y luego ofreceremos algunas recomendaciones preliminares para colaborar en el proceso de toma de decisiones por parte de autoridades públicas y empresas. El objetivo es contribuir a que ambos actores puedan abordar los anteriores problemas de una forma respetuosa de los derechos humanos.

La criminalización del discurso en línea

El uso del derecho penal para castigar conductas que valoramos negativamente es el último recurso al que un Estado democrático puede recurrir (principio de última ratio del derecho penal). En el caso de la libertad de expresión, este principio se refuerza. Someter a una persona a un juicio penal por algo que dijo, comentó o posteo es propio de un gobierno autoritario que no soporta la discusión política. Por eso, se ha afirmado en varias oportunidades que el “Derecho Penal es el medio más restrictivo y severo para establecer responsabilidades respecto de una conducta ilícita”.¹

En Argentina -y otros países del continente- la pandemia dio lugar al inicio de diversos procesos criminales a personas por comentarios hechos en redes sociales. El motivo fue que estas personas habrían incurrido en el delito de “intimidación pública”² al realizar posteos con la finalidad de crear un temor público o provocar incidentes.

El caso más famoso en nuestro país fue el de Kevin Guerra. El 7 de abril de 2020 el joven de 20 años publicó en su cuenta de Twitter el siguiente comentario : “Che, qué onda los que no cobramos el bono de 10.000 pesos. Sigue en pie lo del saqueo, no?”,³ en referencia al no cobro de una asistencia económica de emergencia.⁴ El comentario fue detectado por la Subdirección de Delitos Tecnológicos de la Gendarmería Nacional junto al Área de Ciberpatrullaje de la Policía Bonaerense.⁵ Las fuerzas de seguridad buscaban posteos en Twitter bajo el parámetro “saquear/cuarentena/argentina”⁶ y cuando se encontraron con el tuit de Guerra, emitieron un “alerta temprana” porque entendían que el posteo constituía un delito. De esta manera, Guerra⁷ fue notificado de que se le abría una causa judicial por un delito que tiene una pena de dos a seis años de prisión. Finalmente, el fiscal federal que entendió en la causa consideró que no había comisión de delito alguno. El Juzgado Federal N° 3 de Mar del Plata compartió

la apreciación y a fines de 2020 sobreseyó a Guerra. Al mismo tiempo, varios procesos judiciales por intimidación pública vinieron por el lado de la difusión de información sobre la pandemia que se consideró falsa. En diferentes ciudades de Argentina, personas fueron imputadas por enviar audios de Whatsapp⁸ o hacer publicaciones en sitios web con noticias que presuntamente no eran ciertas.⁹ Incluso hubo periodistas que sufrieron sanciones administrativas de este tenor.¹⁰

La situación provocó que la Comisión Interamericana de Derechos Humanos (CIDH) y la Relatoría Especial para la Libertad de Expresión (RELE) expresaran su preocupación. Ambos organismos reiteraron que “la introducción de tipos penales podría retrotraer a la región a una lógica de criminalizar expresiones sobre funcionarios o asuntos de interés público y establecer una herramienta con un fuerte efecto inhibitorio de la difusión de ideas, críticas e información”.¹¹ Efectivamente, la utilización de figuras como “intimidación pública”, “temor público” o similares siempre han generado sospechas debido a su vaguedad. Pero su aplicación al discurso en línea genera numerosos inconvenientes adicionales. Algunos de ellos son:

- **Mayor restricción de la protesta.** El delito de “intimidación pública” ha sido utilizado para restringir el derecho de la ciudadanía a manifestarse en contra de alguna medida de gobierno,¹² solicitar puestos de trabajo¹³ o reclamar al Poder Judicial.¹⁴ De esta manera, el uso desproporcionado de esta figura penal ha llevado a numerosas instancias de criminalización de la protesta social. No existen motivos para pensar que su aplicación al ámbito online será distinto. **Las limitaciones al espacio público producto de la pandemia han incentivado a que la gente transcurra más tiempo en internet, realizando actividades que antes efectuaba de manera presencial. Dentro de estas acciones también está incluida la protesta.** Precisamente, el caso de Kevin Guerra puede

ser visto como un ejemplo de una protesta sarcástica que buscaba expresar el malestar por la ausencia de ayuda social por parte del gobierno. El accionar subsiguiente de las fuerzas de seguridad demuestra que el abuso continúa e incluso puede agrandarse debido a la masividad del contenido que circula en redes sociales.

• **Mayor riesgo de una errónea aplicación del concepto de intimidación pública.** A lo largo de la historia, la idea de peligro o riesgo de violencia fue utilizada por los gobiernos para reprimir discursos críticos y disidentes.¹⁵ Por esta razón, el estándar para calificar una expresión como ilegal debe ser muy estricto. La doctrina judicial ha afirmado que el discurso debe ser capaz de producir una “acción ilegal inminente”.¹⁶ Fuera de estos casos excepcionales, el Estado no puede asumir de manera legítima la función de decidir qué discurso es aceptable y cuál no. Bajo estos estándares es probable que la mayoría de los intentos por aplicar el delito de “intimidación pública” a posteos en redes sociales sean jurídicamente erróneos. Aquellos discursos suelen ser invocaciones abstractas de violencia que tienen mínimas chances de volverse realidad. En Argentina, varias personas fueron investigadas por tuitear amenazas contra el expresidente Mauricio Macri y la exministra Patricia Bullrich. Sin embargo, ellas fueron luego sobreseídas.¹⁷ En otros casos, incluso suelen ser comentarios en burla que -aunque puedan ser desagradables- no tienen la seriedad requerida para generar violencia. Esto no quiere decir que los discursos en redes sociales nunca pueden ser disparadores de situaciones de violencia. Mucha gente afirmaría que los posteos de Donald Trump resultan un claro ejemplo. Pero en este caso se trata de una figura pública de gran relevancia, con millones de personas que lo siguen y en el marco de una crisis institucional sin precedentes. A modo de justificación para su decisión de suspender a Trump, Facebook afirmó que el expresidente aprobó enfáticamente a la gente que estaba irrumpiendo en el Capitolio

cuando escribió “Los amamos. Ustedes son muy especiales”. En un sentido similar, la plataforma sostuvo que cuando Trump calificó a sus partidarios como “grandes patriotas” y alentó a “recordar este día para siempre” también estaba demostrando un apoyo a la intrusión del Congreso.¹⁸ **El vínculo entre líder político y seguidores amerita una consideración más estricta debido al poder que el primero ejerce sobre los segundos. Pero esta condición está ausente cuando es un ciudadano común el que hace comentarios semejantes en internet.** Por lo tanto, las razones que nos inclinan a considerar con preocupación los comentarios del ex presidente de Estados Unidos nos deben impulsar a rechazar todos los intentos estatales por criminalizar a ciudadanos comunes en base a que sus comentarios pueden incitar a disturbios o acciones semejantes.

• **Riesgo de que estándares privados influyan en la interpretación de figuras criminales.** Todas las grandes plataformas cuentan con políticas de moderación de contenido sobre discurso violento o de incitación al odio.¹⁹ Estas reglas determinan qué se entenderá por tal y cuáles serán las sanciones a aplicar. Además, las compañías tienen una manera propia de aplicar esas directrices. Por cierto, estas políticas no necesariamente están alineadas a los estándares internacionales de derechos humanos. **Las plataformas suelen utilizar definiciones y juicios más expansivos que le permiten remover contenido cuya supresión por parte del Estado nos resultaría cuestionable.** Por ejemplo, la política de discurso de odio de Facebook abarca a expresiones que contienen “estereotipos dañinos, afirmaciones de inferioridad, expresiones de desprecio, repulsión o rechazo o insultos” hacia una característica protegida.²⁰ A su vez, Twitter prohíbe “glorificar, celebrar, elogiar o condonar delitos violentos, eventos violentos dirigidos a ciertas personas por su pertenencia a un grupo protegido, o a los autores de dichos actos”.²¹ De este modo, la definición de glorificación de Twitter “incluye elogiar,

celebrar o condonar declaraciones tales como “Me alegro de que esto haya sucedido”, “Esta persona es mi héroe”, “Me gustaría que más personas hicieran cosas como esta”, o “Espero que esto inspire a otros a actuar”.²² Si gobiernos o el Poder Judicial comienzan a aplicar intensivamente figuras como “intimidación pública” al discurso online, existe la posibilidad de que sus criterios para definir dicho término se vean influenciados por la manera en que las compañías resuelven controversias similares. Frente a la novedad de estos casos, las autoridades públicas y judiciales no cuentan con estándares desarrollados. De este modo, hay muchas probabilidades de que miren los criterios privados para buscar una guía a sus decisiones. En definitiva, una regulación pública y democrática debería enmarcar el desarrollo de las políticas de moderación de contenido, no al revés.

El monitoreo de posteos en redes sociales (ciberpatrullaje)

La investigación y acusación de personas por comentarios en plataformas no pueden tener lugar sin que las fuerzas de seguridad hagan tareas de inteligencia en redes sociales (SOCMINT, por su denominación en inglés) para encontrar expresiones consideradas criminales.²³ Esta actividad es muy riesgosa ya que puede restringir severamente el espacio cívico online y de esta manera, transformarlo en un lugar donde las opiniones críticas son desalentadas.

Si bien Argentina viene haciendo uso de esta técnica desde hace varios años, el tema adquirió mayor repercusión en abril de 2020 cuando la ministra de Seguridad de la Nación afirmó que estaban recurriendo al ciberpatrullaje para detectar el humor social.²⁴ En particular, el gobierno se preocupaba por posibles convocatorias a saqueos y disturbios. Frente a las críticas, el gobierno elaboró un protocolo de actuación²⁵ que -según afirmó- se encuentra en línea con los estándares internacionales de derechos humanos.²⁶ Sin embargo, la autoridad de protección de datos recomendó la suspensión del protocolo ya que no protege de manera adecuada la privacidad de las personas.²⁷

La peligrosidad del uso de SOCMINT para la creación de un ambiente libre en internet resulta evidente por diversas razones. Entre ellas, podemos mencionar:

- **Legítima la prevención policial de expresiones, no de conductas.** Esta actividad ha sido denominada coloquialmente “ciberpatrullaje” ya que se la pretende asimilar a las tareas de prevención que la Policía realiza en los espacios públicos (calles, plazas, parques, etc). Sin embargo, esta identificación omite una diferencia sustancial. La prevención policial tiene como objeto

principal la disuasión de conductas ilícitas que puedan tener una consecuencia directa e inmediata en la vida, la integridad física o la propiedad de una persona. Tenemos policías en la calle para que actúen en caso de que una persona agrede a otra, le apunte con una pistola para asaltarla, destruya un negocio, intente entrar a una casa para robar, etc. Por el contrario, **el ciberpatrullaje se enfoca principalmente en el discurso, es decir, en comentarios, mensajes, posts y demás formas de comunicación en internet. Es decir, el objeto de la vigilancia son expresiones que en general, no ponen en riesgo de manera inmediata los bienes e intereses de una persona.** Desde luego que esto no significa que internet no pueda ser utilizada para enviar mensajes o intercambiar información para llevar a cabo una conducta ilícita, sea online o en la vida real. Pero la sensibilidad del uso de SOCMINT lleva a que finalmente se termine utilizando para control social o restringir discurso legítimo.

• **Dificulta que las fuerzas de seguridad respondan por su accionar.** El control ciudadano de las fuerzas de seguridad es esencial para que las personas mantengan la confianza en el sistema. En el caso del patrullaje en espacios físicos, hay varios aspectos que ayudan a esta tarea. En primer lugar, los policías se encuentran distinguidos por su uniforme, su placa o porque conducen un móvil que también está identificado. De este modo, los individuos pueden supervisar la actuación de los oficiales. En segundo lugar, si el patrullaje tiene lugar en espacios públicos, es probable que mucha gente sea testigo de la actividad policial. Finalmente, las personas pueden registrar casos de abuso policial mediante la toma de fotos o la grabación de videos. Incluso el periodismo ha sido importante para revelar instancias en las que la Policía violó los derechos humanos.²⁸ Estas circunstancias no están presentes en el ciberpatrullaje. No hay forma de saber si en este momento la Policía está revisando nuestros comentarios. No podemos acceder a las oficinas en donde se realizan actividades de

SOCMINT. No tenemos información acerca de las páginas que se visitan y si la tenemos, solo nos queda confiar en la palabra de las autoridades. La esperanza de que una ciudadana o una periodista registre un posible caso de abuso es ínfima. **El ciberpatrullaje posee un secretismo intrínseco a su modo de funcionamiento. Por eso, debemos tener mucho cuidado con su utilización.**

• **Supone que los datos provenientes de fuentes de acceso público no merecen un alto nivel de protección.** El mayor argumento para el uso de SOCMINT radica en que supuestamente se trataría de información o datos de acceso público. Impedir que las fuerzas de seguridad consulten las redes sociales para ver qué dice la gente sería análogo a ponerle una venda en los ojos al policía que recorre las calles de un barrio. Pero la comparación no es precisa. Las redes sociales son espacios peculiares ya que la información allí alojada está en principio disponible públicamente. Sin embargo, el funcionamiento de estos lugares dependen de entidades privadas que imponen sus propias reglas. Por lo tanto el carácter híbrido de las redes habilita a considerar el asunto de una manera particular. Es que en la era digital la protección de datos personales ha adquirido una importancia preponderante ya que permite que los individuos -de acuerdo al principio de autodeterminación informativa- controlen el uso y los destinatarios de toda información personal. **La gente expresa sus opiniones políticas, sus preferencias sexuales o su estado de salud en las redes sociales. La sensibilidad de estos datos requiere que su tratamiento no sea objeto de vigilancia o monitoreo encubierto por parte de las autoridades.** En su opinión sobre el protocolo para la regulación del ciberpatrullaje la autoridad de protección de datos recordó que los datos sensibles gozan de una protección especial y que por lo tanto, cualquier tratamiento está sometido a un escrutinio más estricto.²⁹ En ese sentido, toda actividad de ciberpatrullaje se parece más a la figura de un policía parado en la vereda de una casa todo el día para ver a través de la ventana qué

es lo que sucede adentro. Y en ese caso, el equivalente a “correr las cortinas” será el cierre de cuentas o la abstención a opinar en temas controversiales.

La influencia privada en la toma de decisiones sobre moderación de contenido

Tanto desde la sociedad civil³⁰ como la academia³¹ se viene insistiendo en la necesidad de adoptar los estándares internacionales de derechos humanos para la toma de decisiones sobre moderación de contenido. Si bien resulta obvio que ninguna solución resolverá mágicamente todos los problemas de la gobernanza del discurso online, la perspectiva de derechos humanos presenta ciertas ventajas. En términos de legitimidad, nos brinda un lenguaje universal en el cual podemos apoyarnos para iniciar una discusión igualitaria. Las controversias sobre remociones de posteos, suspensiones de cuentas y demás asuntos deberían ser resueltas en base a estándares que la comunidad internacional como tal pueda reconocer como suyos. Por otro lado, sirve como base para el diseño de regulación pública que evite la caída en tendencias autoritarias por parte de los gobiernos. Sin embargo, esta aspiración debería ser concebida como un primer paso indispensable pero insuficiente en sí mismo. **Cualquier implementación de un marco normativo -incluido el de los derechos humanos- se verá frustrada si no hay una desconcentración del poder que las compañías mantienen respecto a la toma de decisiones acerca de la moderación de contenido.** Un cambio de las reglas sin alterar profundamente el status quo en relación a los actores y el modo en que se aplicarán estas reglas puede dar como resultado los siguientes fenómenos:

- **Incrementar el poder de las plataformas.** Los estándares de derechos humanos necesitan ser interpretados antes de poder darnos una respuesta. Por lo tanto, es de crucial importancia saber quién va a ser el encargado de aplicar la norma. Si las plataformas continúan decidiendo de manera unilateral y exclusiva, reemplazar

directrices privadas por el Pacto Internacional de Derechos Civiles y Políticos no tendrá un efecto muy significativo. La discrecionalidad seguiría pero esta vez en nombre de los derechos humanos. De este modo, el poder de las compañías aumentaría ya que ahora contarían con una capa mayor de legitimidad.

• **Aumento de las probabilidades de que mecanismos externos de moderación de contenido influyeran en las decisiones de las autoridades públicas.** Las compañías están diariamente lidiando con miles de casos que suceden en sus plataformas. Este trabajo intensivo ha permitido que hayan desarrollado criterios expeditos para resolver las controversias sobre contenido. Al mismo tiempo, algunas de ellas han incorporado mecanismos de revisión externos de sus decisiones. El caso del Oversight Board (OB) es un ejemplo paradigmático. Las decisiones del OB han aplicado las normas comunitarias de Facebook pero han recurrido a los estándares internacionales de derechos humanos como referencia para interpretar las reglas internas. Frente a esta situación, las autoridades judiciales -en especial, aquellas de nuestro país o región- no cuentan con una experiencia equivalente. Por lo tanto, existe la posibilidad de que las autoridades judiciales se vean atraídas por esas interpretaciones en caso que deban intervenir. En consecuencia, los criterios utilizados por las autoridades podrían ser nada más que el reflejo del juicio previo de los mecanismos privados ya establecidos.

• **Impedirá que los usuarios tengan un rol central en las decisiones.** Las plataformas han manifestado que son conscientes del problema ocasionado por el hecho de que las decisiones sobre gobernanza del discurso en línea estén exclusivamente en sus manos. Es por eso que varias de ellas han comenzado a diseñar mecanismos para que actores externos puedan ejercer influencia

en aquellas resoluciones de manera vinculante. El ejemplo más relevante hasta el momento es el de Facebook. La compañía instituyó el Oversight Board que tendrá la palabra final en ciertos casos controversiales sobre remociones de contenido.³² La innovación constituye una buena noticia en tanto Facebook reconoce el problema de la falta de legitimidad para tomar decisiones tan sensibles. Pero esta solución también presenta sus dificultades. Los miembros de la Junta cuentan con la trayectoria e idoneidad suficiente para tratar estos temas. Al mismo tiempo, no hay evidencias de que su independencia se encuentre afectada. Sin embargo, la tarea sigue estando en manos de un número reducido de personas. **Las respuesta a problemas tan controversiales y globales debe apuntar a que la ciudadanía sea incorporada de manera significativa en la toma de decisiones.** Internet ya cuenta con modelos alternativos de moderación de contenido en donde las personas poseen una influencia considerable. Por ejemplo, el sistema de solución de controversias de Wikipedia reposa de manera fundamental en la discusión y decisión de las personas usuarias y voluntarias de la iniciativa. Por supuesto que no sería posible replicar este modelo de manera exacta a plataformas con características muy distintas. Pero de lo que se trata es de señalar que no hay barreras insalvables para ampliar la intervención ciudadana en la resolución de conflictos sobre moderación de contenido.

Las expresiones de figuras políticas y la violencia

A partir de la decisión de las grandes plataformas de expulsar al ex presidente de Estados Unidos Donald Trump de sus redes sociales, la discusión sobre moderación de contenido se ha enfocado en la relación entre el discurso de políticos y la violencia. Por un lado, nuestras intuiciones nos dicen que la ciudadanía debería acceder de la manera más sencilla posible a las expresiones de funcionarios públicos de tamaño relevancia. De este modo, la prohibición de un presidente de comunicarse por medio de las redes más usadas por la población equivale a una restricción significativa. Por otro lado, la influencia de este tipo de figuras puede llevar a que sus declaraciones incendiarias motiven a actos de violencia. A diferencia de alguien común, las figuras políticas son personas con ascendencia sobre buena parte de la población y por lo tanto, capaces de provocar conductas destructivas. Las respuestas a esta dicotomía son variadas. Sin embargo, lo importante es señalar ciertas preocupaciones subyacentes al proceso de decisión ya que son éstas las que van a influir en el diseño de un espacio cívico online saludable. Veamos algunas de ellas.

- **Idoneidad de las plataformas para analizar el contexto.** De acuerdo a las plataformas, la expulsión de Trump sucedió ya que un análisis del contexto estadounidense demostró que sus declaraciones tenían un claro potencial para provocar actos de violencia.³³ Este accionar demuestra lo importante que es para la moderación de contenido el hecho de conocer el terreno para poder tomar una decisión. Pero esto es algo que las principales compañías pueden hacer con efectividad -en cierta medida- en Estados Unidos ya que la mayor parte de ellas son originarias de allí. Por el contrario, en el resto del mundo probablemente carezcan por sí mismas de la habilidad de comprender los efectos de ciertos

discursos en un ambiente social o clima político determinado. En consecuencia, es posible que decisiones de este estilo no sean tomadas con la debida información cuando se trate de otros países, especialmente aquellos con una situación muy diferente a la de los Estados Unidos.

• **Soluciones locales con impactos globales.** Si bien la expulsión de Trump estuvo motivada bajo el argumento de prevenir un recrudecimiento de la violencia en Estados Unidos, las consecuencias alcanzaron a usuarios de todo el mundo. Es decir, la situación interna de Estados Unidos fue el motivo por el cual personas de Argentina, Dinamarca, India, Australia y demás países del mundo se vieron privadas de continuar viendo los mensajes de Donald Trump. Resulta evidente que no había ninguna chance de que aquellos países se vieran expuestos a hechos de violencia debido a los tuits de Trump. Por lo tanto, en un cierto sentido, la medida adoptada puede ser considerada desproporcionada. Es verdad que las plataformas implementan sus decisiones de manera uniforme en todos los países en que operan. Además de las económicas, hay razones técnicas y operativas para justificar ese modus operandi. Por otro lado, una fragmentación de las reglas -un Trump prohibido en Twitter en EE. UU. pero permitido en el resto del mundo- tampoco es una solución de fácil implementación y puede generar problemas mayores de los que pretende resolver. Sin embargo, el fenómeno merece ser considerado. **Las remociones globales por orden judicial suelen ser cuestionadas porque pueden conducir a una “carrera hacia el abismo” en la aplicación de normas restrictivas de la libertad de expresión. La conclusión no debería ser distinta cuando esa decisión es tomada por una empresa privada.**

• **La coincidencia de criterios entre las empresas.** Cada vez más las grandes plataformas adoptan decisiones similares en temas de moderación de contenido. La época en la cual las compañías

adoptaban diversas estrategias para enfrentar la desinformación o aplicaban diferentes criterios para determinar qué contenido puede permanecer en sus redes está siendo reemplazada por una coincidencia -espontánea o deliberada- de las acciones a tomar, fenómeno que ha sido descrito por algunas autoras con el nombre de "carteles de internet".³⁴ Si esta tendencia a la uniformización se acentúa, la correspondiente pérdida de diversidad traería varios problemas. En primer lugar, se expanden los efectos dañinos de una decisión equivocada. La homogeneidad impide que, por ejemplo, una remoción injusta pueda ser mitigada por la decisión de otra plataforma de mantener el contenido. En segundo lugar, se reducen las posibilidades de que diferentes políticas de moderación de contenido nos permitan experimentar, comparar y evaluar los beneficios y desventajas de cada una de ellas. De este modo, se pierde la riqueza necesaria para poder evolucionar en la comprensión de un tema como la gobernanza de internet que recién está en sus primeros estadios de desarrollo. Por último, la cartelización deja cada vez más en evidencia el inmenso grado de concentración de poder en manos de pocos actores. Frente a este panorama, una acción coordinada de las grandes plataformas puede dejar en el ostracismo a una persona al privarla de toda interacción significativa en las redes sociales más importantes.³⁵

Recomendaciones preliminares

Los desafíos son complejos y no hay disponible ninguna solución mágica. Pero sí existen acciones que de ser implementadas, pueden contribuir al diseño de un espacio cívico online favorable para el ejercicio de derechos. A continuación, brindamos algunas recomendaciones preliminares para abordar los problemas mencionados anteriormente.

Criminalización del discurso online

- **No recurrir a conceptos previstos en el Código Penal como “intimidación pública” o similares para perseguir comentarios hechos en redes sociales.** Las fiscalías y juzgados penales deben implementar los estándares de libertad de expresión al momento de evaluar el inicio de investigaciones criminales o la aplicación de tipos penales a personas por su actividad en redes sociales. Estos principios nos indican que únicamente puede ser perseguido aquel discurso que esté dirigido a producir una acción ilegal inminente y que cuente con posibilidades ciertas de hacerlo. Bajo este estándar, la inmensa mayoría de los comentarios hechos en internet no deben ser criminalizados aunque se trate de expresiones desagradables o molestas. Por otro lado, resultaría conveniente la sanción de una ley que establezca de manera clara que las expresiones referidas a asuntos de interés público no están incluidas en los delitos de intimidación pública. Al igual que lo sucedido con los delitos de calumnias e injurias, la aclaración sería útil para evitar la persecución penal a personas que recurren a internet para difundir información.

- **Consagrar y expandir los alcances del derecho a la protesta en internet.** En un país con profundas desigualdades como la Argentina, diferentes grupos sociales cuentan con pocos medios

para hacer escuchar sus demandas a los poderes políticos. En este contexto, internet puede colaborar en mitigar -aunque sea mínimamente- esa ausencia de recursos. Esta situación debería ser motivo para que el Poder Judicial y las autoridades resguarden especialmente aquellas expresiones de protesta. De este modo, ejemplos como el tuit de Kevin Guerra podrían haber sido utilizados para atender el problema subyacente de la asistencia social a gente afectada por la cuarentena y no para iniciar investigaciones penales por bromas que no tenían posibilidades de generar ningún impacto serio.

Ciberpatrullaje

- **Las tareas de prevención policial en internet deberían ser sometidas a un escrutinio estricto.** No podemos normalizar el patrullaje en internet de la misma manera en que lo hacemos con el patrullaje en la calle por una sencilla razón: nuestra ley prohíbe obtener información o almacenar datos sobre personas por sus actividades lícitas.³⁶ Precisamente, esto es lo que hace una actividad de prevención masiva en línea que no se enmarca en una investigación judicial determinada. De este modo, es el Estado el que debe demostrar que no existen alternativas menos lesivas para llevar a cabo su misión de brindar seguridad. Y a su vez, debe cuidar que su accionar sea limitado y específico. En términos de protección del espacio cívico, esto supone que no debería usarse el ciberpatrullaje para buscar activamente presuntos delitos de opinión o expresión.

- **Crear mecanismos de rendición de cuentas independientes y con participación ciudadana.** Las tarea policial en internet debe contar con un sistema externo de frenos y contrapesos que asegure que las fuerzas de seguridad realicen su labor con respeto a los derechos de las personas. Esto supone múltiples

condiciones pero como punto de partida exige que los organismos policiales difundan públicamente su actividad y sus políticas de actuación en redes sociales para que puedan ser evaluadas de manera imparcial. Por otro lado, requiere la creación de comités de supervisión formado por la ciudadanía que asegure una revisión democrática de los procedimientos policiales. Finalmente, demanda la realización de evaluaciones y auditorías de impacto en derechos humanos que sean llevadas a cabo por instituciones independientes de reconocida trayectoria e idoneidad.

La toma de decisiones sobre moderación de contenido

- **Profundizar y desarrollar los estándares existentes en materia de libertad de expresión.** Los Estados, los poderes judiciales y los sistemas de derechos humanos -cada uno dentro del rol que les toca- deben continuar trabajando para especificar y detallar los principios vigentes sobre libertad de expresiones. De lo contrario, la vaguedad de los términos puede ser aprovechada por las grandes plataformas para justificar decisiones tomadas en vista de sus intereses mientras proclaman que están aplicando el marco internacional de derechos humanos. Es cierto que temas como el discurso de odio o la desinformación generan legítimos desacuerdos en relación a cómo afrontarlos. Sin embargo, esto no impide que los poderes públicos impulsen la discusión y realicen propuestas que puedan ser discutidas de manera abierta y democrática.

- **Explorar vías alternativas para incrementar la legitimidad democrática de los órganos y procedimientos por los cuales se diseñan y aplican las políticas de moderación de contenido.** Más allá de lo justo o injusto de la solución, las decisiones de moderación de contenido de las grandes plataformas digitales

adolecen de un profundo déficit democrático. Las organizaciones de la sociedad civil, la academia y las personas en general deben ser incorporados de manera significativa en los procesos internos que cada plataforma realiza para moderar contenido. Significativo, en este caso, implica no participar únicamente como órganos de consulta o de asesoramiento sino tener verdadero poder e influencia en las decisiones más controversiales. Al mismo tiempo, la democratización de los modelos de gobernanza del discurso en línea requiere que las plataformas diseñen nuevos mecanismos que permitan la incorporación directa de la ciudadanía en la toma de decisiones. Si bien es demasiado pronto para afirmar con certeza que esta es la solución más eficaz, parece ser una vía posible de explorar para -junto con otras medidas- abordar el ya señalado déficit democrático.

Violencia y expresiones de funcionarios públicos

• **En contextos de alta conflictividad, los funcionarios públicos tienen un deber especial de evitar que sus declaraciones afecten derechos fundamentales.** La libertad de expresión pertenece a toda la población y por lo tanto, también a los funcionarios públicos. Sin embargo, los líderes políticos tienen el deber de proteger los derechos fundamentales de la ciudadanía no solo con sus conductas sino también con sus dichos. Tal como lo ha sostenido el sistema interamericano, en situaciones de alta polarización política y/o social los funcionarios deben evitar discursos que pongan en una situación de vulnerabilidad a ciertos grupos sociales.³⁷ Por ejemplo, la Corte Interamericana de Derechos Humanos ha indicado que los funcionarios públicos deben abstenerse de hacer declaraciones que, en el marco de un contexto de polarización social, aumenten el riesgo de periodistas y medios de sufrir agresiones por parte de terceros.³⁸ Bajo este principio,

pueden justificarse algunas decisiones drásticas tomadas por las plataformas. Pero es necesario que dichas resoluciones estén respaldadas por un conocimiento del contexto local. Para esto, es indispensable un vínculo permanente con actores nacionales, quienes -en sintonía con lo dicho anteriormente- deben tener la posibilidad de influir de manera cierta en las determinaciones hechas por las compañías.

• **Las medidas para combatir las expresiones de violencia por funcionarios públicos deben ser lo suficientemente adecuadas como para no afectar el principio de proporcionalidad.** La suspensión indefinida de una cuenta implica quizás la sanción más severa en términos de la garantía de la libertad de expresión. Por lo tanto, es indispensable que las plataformas exploren otras alternativas menos lesivas que puedan lograr el mismo fin. Un ejemplo claro fue lo sucedido en el caso de Trump. Si el argumento para expulsar de las redes sociales al ex presidente fueron las extraordinarias circunstancias políticas en Estados Unidos, la duración de la suspensión podría haberse extendido hasta la finalización de la transición de poder. Y si las plataformas hubieran considerado que el peligro aún persistía, podrían haber renovado la suspensión por otro periodo de tiempo y así sucesivamente. Una prohibición indefinida deja a las compañías con absoluta discrecionalidad para evaluar posibles nuevas circunstancias. En cambio, una decisión con plazo de vencimiento puede obligar a las plataformas a analizar cambios en el contexto social con el fin de justificar públicamente si van a seguir o no con la suspensión.

* * *

Notas

- 1 Corte Interamericana de Derechos Humanos, Caso Palamara Iribarne Vs. Chile. Fondo, Reparaciones y Costas. Sentencia de 22 de noviembre de 2005. Serie C No. 135, párr. 79: https://www.corteidh.or.cr/docs/casos/articulos/Seriec_135_esp.pdf
- 2 El art. 211 del Código Penal argentino sostiene que “será reprimido con prisión de dos a seis años, el que, para infundir un temor público o suscitar tumultos o desórdenes, hiciere señales, diere voces de alarma, amenazare con la comisión de un delito de peligro común, o empleare otros medios materiales normalmente idóneos para producir tales efectos”: <http://servicios.infoleg.gob.ar/infolegInternet/anexos/15000-19999/16546/texact.htm#22>
- 3 <https://twitter.com/KevinGuerra99/status/1247709948554903554>
- 4 *La Nación*, “Ciberpatrullaje: hizo un chiste sobre saqueos en Twitter y le abrieron una causa”, 15 de abril de 2020: <https://www.lanacion.com.ar/sociedad/ciberpatrullaje-hizo-chiste-saqueos-twitter-le-abrieron-nid2354402/>
Clarín, “Ciberpatrullaje: ‘ Sigue en pie lo del saqueo, no?’ , el tuit que le valió una causa judicial a un joven de Balcarce”, 14 de abril de 2020: https://www.clarin.com/policiales/-sigue-pie-saqueo-tuit-valio-causa-judicial-joven-balcarce_0_Q2GKqAle1.html
- 5 *La Izquierda Diario*, “Efecto ciberpatrullaje. Kevin Guerra: La DDI me dijo ‘al presidente no le gustó tu tuit’ y ahora estoy procesado”, 15 de abril de 2020: <https://www.laizquierdadiario.com/Kevin-Guerra-La-DDI-me-dijo-al-Presidente-no-le-gusto-tu-tuit-y-ahora-estoy-procesado>
- 6 CELS, “La justicia federal sobreseyó a Kevin Guerra por sus expresiones en Twitter”, 14 de enero de 2021: <https://www.cels.org.ar/web/2021/01/la-justicia-federal-sobreseyo-a-kevin-guerra-por-sus-expresiones-en-twitter/>
- 7 Ibid.

8 Diario *El Tiempo*, "Se está instruyendo otro sumario penal por 'intimidación pública'," 29 de marzo de 2020: <https://www.diarioeltiempo.com.ar/nota-se-esta-instruyendo-otro-sumario---penal-por-intimidacion-publica-163794>

El Popular, "Dos policías y un vecino de Azul imputados por difundir información falsa", 25 de marzo de 2020: <http://www.elpopular.com.ar/142512>

9 *Infobae*, "Allanaron a una mujer por difundir noticias falsas sobre el coronavirus y encontraron al secretario de Energía durmiendo en su casa", 4 de abril de 2020: <https://www.infobae.com/politica/2020/04/04/allanaron-a-una-mujer-por-difundir-noticias-falsas-sobre-el-coronavirus-y-encontraron-al-secretario-de-energia-durmiendo-en-su-casa/>

Misiones Online, "Coronavirus: por una causa judicial de fake news allanan una vivienda en Oberá y encuentran al secretario de Energía de la Nación Sergio Lanziani", 4 de abril de 2020: <https://misionesonline.net/2020/04/04/coronavirus-por-una-causa-judicial-de-fake-news-allanan-una-vivienda-en-obera-y-encontran-a-sergio-lanziani-secretario-de-energia-de-la-nacion-violando-la-cuarentena1/>

10 *Página/12*, El intendente de Pichanal multó a un comunicador 23 de junio de 2020: <https://www.pagina12.com.ar/274042-el-intendente-de-pichanal-multo-a-un-comunicador>

Diario Clarín, "Salta: un intendente multó con \$40 mil a un periodista por reportar casos de coronavirus en su ciudad", 24 de junio de 2020: https://www.clarin.com/politica/salta-intendente-multo-40-mil-periodista-reportar-casos-coronavirus-ciudad_0_bev8G3nMu.html

11 Comunicado de prensa R78/20, "CIDH y su RELE expresan preocupación por las restricciones a la libertad de expresión y el acceso a la información en la respuesta de Estados a la pandemia del COVID-19", 18 de abril de 2020: <https://www.oas.org/es/cidh/expresion/showarticle.asp?artID=1173&IID=2>

12 CELS, "El derecho a la protesta en peligro", 15 de diciembre de 2017: <https://www.cels.org.ar/web/2017/12/el-derecho-a-la-protesta-en-peligro/>

- 13 *Diario Judicial*, “La protesta social ganó una batalla”, 9 de diciembre de 2005, <https://www.diariojudicial.com/nota/51879>
- 14 *Página/12*, “Una figura para criminalizar la protesta social”, 6 de julio de 2017: <https://www.pagina12.com.ar/48389-una-figura-para-criminalizar-la-protستا>
- 15 En Estados Unidos activistas socialistas y comunistas fueron condenados penalmente por distribuir panfletos en oposición al servicio militar obligatorio para combatir en la Primera Guerra Mundial. Las condenas fueron ratificadas por la Corte Suprema de aquel país bajo el estándar del “peligro claro y presente”, Cfr. “Schenck v. United States, 249 U.S. 47 (1919)” [<https://www.law.cornell.edu/supremecourt/text/249/47>] y “Abrams v. United States, 250 U.S. 616 (1919)” [<https://www.law.cornell.edu/supremecourt/text/250/616>].
- 16 El estándar de la “acción legal inminente” es consagrado por la Corte Suprema de Justicia de Estados Unidos en el fallo “Brandenburg v. Ohio”: <https://www.law.cornell.edu/supremecourt/text/395/444>. En Argentina ha sido adoptado para analizar casos de libertad de expresión en internet como “Vita, Leonardo G. Y González Eggers, Matías s/procesamiento” de la sala 1 de la Cámara Federal: <http://www.hfernandezdelpech.com.ar/JurisprudenciaArgLiberExpreFalloCamaraFed.htm>
- 17 *La Vaca*, “Tuitear contra Macri no es delito”, 22 de mayo de 2019: <https://lavaca.org/notas/tuitear-contramacri-no-es-delito/>
Perfil, “Declaran ´ inocente ´ al joven que estuvo preso por un tuit contra Mauricio Macri”, 23 de marzo de 2018: <https://www.perfil.com/noticias/sociedad/declaran-inocente-al-joven-que-escribio-un-tuit-contramacri.phtml>
- 18 Para ver un resumen de las explicaciones de Facebook ver la decisión del Oversight Board sobre el caso: <https://www.oversightboard.com/sr/decision/2021/001/pdf-english>

- 19 El sitio web <https://letrachica.digital/> monitorea y analiza las normas de moderación de contenido de Facebook, Twitter y YouTube.
- 20 Normas comunitarias de Facebook. Parte III Contenido inaceptable. Lenguaje que incita al odio: <https://www.facebook.com/communitystandards/objectionable>
- 21 Reglas de Twitter. Política relativa a la glorificación de la violencia: <https://help.twitter.com/es/rules-and-policies/hateful-conduct-policy>
- 22 Ibid.
- 23 Cfr. ADC, *Seguidores que no vemos. Una primera aproximación al uso estatal del Open-source intelligence (OSINT) y Social media intelligence (SOCMINT)*, octubre de 2018: <https://adc.org.ar/informes/seguidores-que-no-vemos/>
- 24 *Infobae*, "Polémica revelación: la ministra de Seguridad admitió que las fuerzas a su cargo realizan ciberpatrullaje en redes sociales para ´ detectar el humor social ´", 8 de abril de 2020: <https://www.infobae.com/politica/2020/04/09/polemica-revelacion-la-ministra-de-seguridad-admitio-que-las-fuerzas-a-su-cargo-realizan-ciberpatrullaje-en-redes-sociales-para-detectar-el-humor-social/>
- Perfil*, "Frederic habló de ´ ciberpatrullaje para medir humor social ´ y desató otra polémica", 8 de abril de 2020: <https://www.perfil.com/noticias/actualidad/sabina-frederic-ciberpatrullaje-redes-sociales-medir-humor-social-no-me-van-a-callar.phtml>
- 25 Resolución 144/2020: Protocolo de Ciberpatrullaje en fuentes abiertas, Ministerio de Seguridad de la Nación, 31 de mayo de 2020: <https://www.boletinoficial.gob.ar/detalleAviso/primera/230060/20200602>
- 26 Télam, "Seguridad creó un protocolo para detectar ciberdelitos relacionados con la pandemia de coronavirus", 2 de junio de 2020: <https://www.telam.com.ar/notas/202006/471150-crean-un-protocolo-para-detectar-delitos-relacionados-con-la-pandemia-del-covid-19-en-internet.html>

Página/12, "Coronavirus: Un protocolo para el ciberpatrullaje", 2 de junio de 2020

27 Agencia de Acceso a la Información Pública Respuesta a Nota NO-2020-41096462-APN-UGA-MSG - Mesa Consultiva para la evaluación y seguimiento del Protocolo General para la Prevención Policial del Delito con uso de Fuentes Digitales Abiertas, 23 de julio de 2020: <https://www.argentina.gob.ar/sites/default/files/no-2020-47326285-apn-aaip.pdf>

28 *La Izquierda Diario*, "Maxi Kosteki y Dario Santillán. Avellaneda hostil: la historia de la foto de Sergio Kowalewski", 26 de junio de 2020: <https://www.laizquierdadiario.com/Avellaneda-hostil-la-historia-de-la-foto-de-Sergio-Kowalewski>

Página/12, "Imágenes que prueban quien disparó", 26 de junio de 2005, <https://www.pagina12.com.ar/diario/elpais/1-52917-2005-06-26.html>

29 Agencia de Acceso a la Información Pública Respuesta a Nota NO-2020-41096462-APN-UGA-MSG - Mesa Consultiva para la evaluación y seguimiento del Protocolo General para la Prevención Policial del Delito con uso de Fuentes Digitales Abiertas, 23 de julio de 2020: <https://www.argentina.gob.ar/sites/default/files/no-2020-47326285-apn-aaip.pdf>

30 Observacom et al., "Estándares para una regulación democrática de las grandes plataformas que garantice la libertad de expresión en línea y una internet libre y abierta", julio de 2020: <https://www.observacom.org/wp-content/uploads/2020/09/Estandares-para-una-regulacion-democratica-de-las-grandes-plataformas.pdf>

31 David Kaye, *Speech Police: The Global Struggle to Govern the Internet*, junio de 2019.

32 Consejo Asesor de Contenido, "Garantizamos el respeto por la libertad de expresión basándonos en un criterio independiente": <https://oversightboard.com/>

- 33 Cfr. Post de Mark Zuckerberg del 7 de enero de 2021:
https://www.facebook.com/story.php?story_fbid=10112681480907401&id=4
- 34 Cfr. Evelyn Douek, "The rise of content cartels", 11 de febrero de 2020:
<https://knightcolumbia.org/content/the-rise-of-content-cartels>
- 35 Cfr. CELE, "El ostracismo de Trump", 30 de enero de 2021:
<https://observatoriolegislativocele.com/el-ostracismo-de-trump/>
- 36 Art. 4 inc. 2 Ley 25520: <http://servicios.infoleg.gob.ar/infolegInternet/anexos/70000-74999/70496/texact.htm>
- 37 Cfr. Relatoría Especial para la Libertad de Expresión, Comisión Interamericana de Derechos Humanos, "Marco jurídico interamericano sobre el derecho a la libertad de expresión", 2010, párrafos 201 a 212:
http://www.oas.org/es/cidh/expresion/docs/cd/sistema_interamericano_de_derechos_humanos/index_MJIAS.html
- 38 Ibid. párr 178



por los Derechos Civiles

adc.org.ar