



Moderación privada de contenidos en Internet y su impacto en el periodismo

Ana Laura Pérez Bertrán
Carolina Martínez Elebi



OBSERVACOM

FRIEDRICH
EBERT
STIFTUNG



OBSERVACOM

Moderación privada de contenidos en internet y su impacto en el periodismo

Es una publicación de OBSERVACOM Observatorio Latinoamericano de Regulación, Medios y Convergencia

Gral. César Díaz 1239/101
Montevideo, Uruguay
www.observacom.org

Con el apoyo de Fundación Friedrich Ebert en Perú
Camino Real 456 Torre Real Of. 901
San Isidro - Lima, Perú
<https://peru.fes.de>



SOBRE LAS AUTORAS



Ana Laura Pérez Bertrán es Master in Business Administration del Instituto de Estudios Empresariales Montevideo de la Universidad de Montevideo (Uruguay) y es Licenciada en Ciencias de la Comunicación por la Universidad ORT (Uruguay), donde es docente desde 2016. Es consultora en temas vinculados a desinformación, libertad de expresión y moderación en plataformas sociales. Es gerenta de Innovación y Digital en el diario El País de Montevideo y conductora del programa Relatos de la señal pública TV Ciudad. Tiene una Diplomatura en Estudios Latinoamericanos de la Universidad de Montevideo.



Carolina Martínez Elebi es Licenciada en Ciencias de la Comunicación por la Universidad de Buenos Aires (UBA) de Argentina, donde se desempeña como docente desde 2011. Es consultora en temas vinculados al impacto de las tecnologías de la información y la comunicación en los derechos humanos y es directora del medio digital DHyTecno. En 2018 cursó el Programa de Derecho de Internet y Tecnologías de las Comunicaciones del Centro de Estudios en Tecnología y Sociedad (CETyS-UdeSA). Forma parte del Grupo de Estudios Regulación de los medios y las TIC del IEALC e integra el Observatorio de Impactos Sociales de la Inteligencia Artificial de la UNTREF.

01.

INTRODUCCIÓN

El presente documento aborda la problemática de la moderación de contenidos en las plataformas haciendo hincapié en su impacto fundamental en el periodismo y teniendo como eje los derechos a la libertad de prensa y la libertad de expresión.

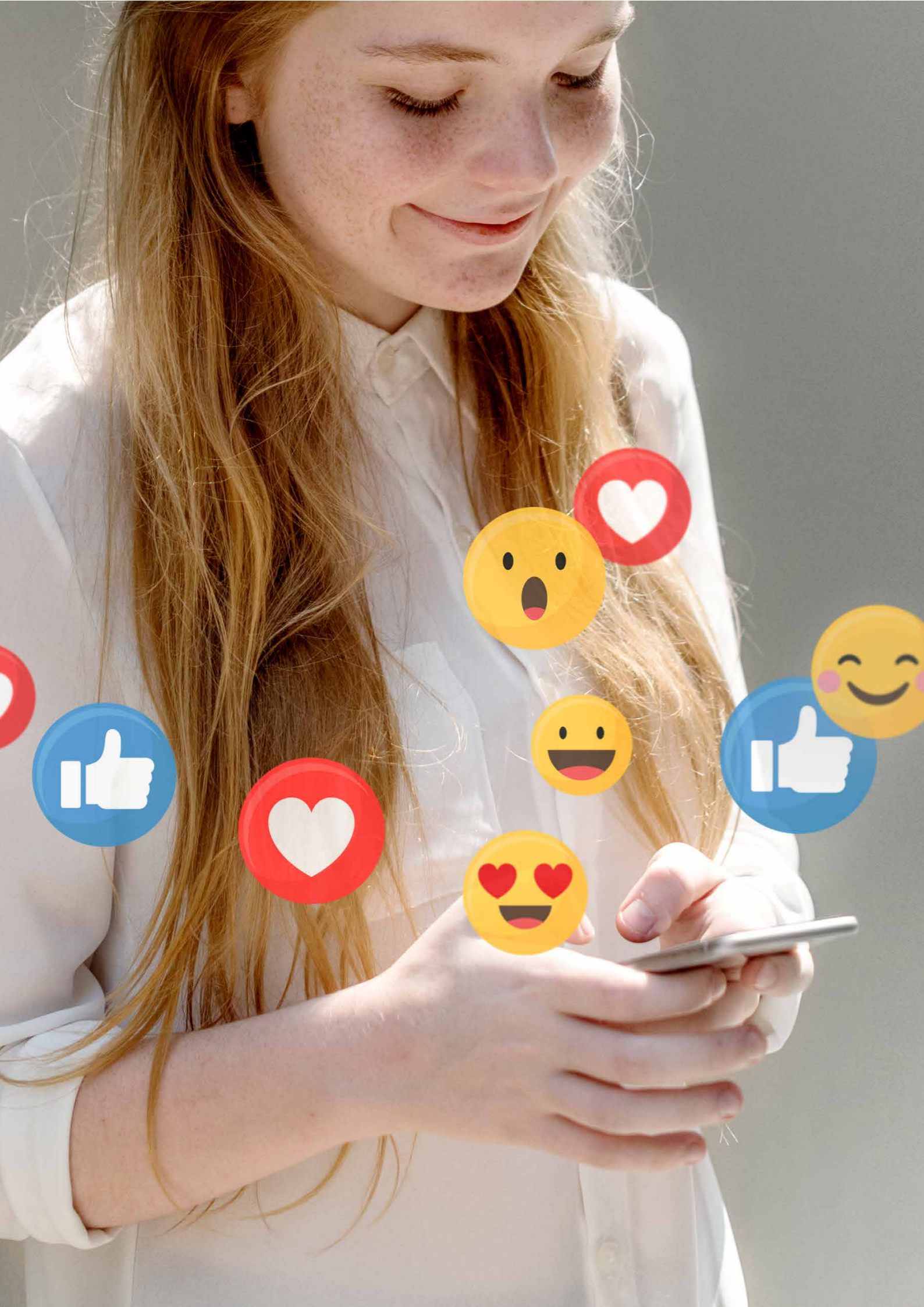
Para empezar, se describe qué se entiende por moderación, y se explica cómo funciona en las plataformas este sistema. En este sentido, se tienen en cuenta tanto los procesos automatizados como la moderación humana que deciden qué puede publicarse y qué no de acuerdo a sus términos de uso y normas comunitarias.

Además, se describen los tipos de moderación, dejando en claro que no se limita exclusivamente a la remoción de los contenidos y la eliminación de las cuentas de los usuarios -que son los dos ejemplos más extremos a través de los que puede ejercerse la censura por parte de estas plataformas-, sino detallando también los modos más sutiles que pueden invisibilizar discursos y actores en las redes sociales.

Esta moderación adquiere incluso otra rele-

vancia cuando se trata de moderar cuentas de medios de comunicación y de periodistas y que, para poder informar y comunicar, deben regirse bajo normas privadas que muchas veces no son del todo claras ni transparentes. A veces, incluso, son invisibles. Esto hace que no sepan qué pueden publicar sin correr el riesgo de que sus contenidos sean eliminados o sus cuentas bloqueadas.

Para que todo esto no se limite a descripciones abstractas, las autoras presentan una serie de casos que han ocurrido en la región en los últimos años y en diversas plataformas, como ejemplos del impacto de la moderación privada en el ejercicio diario del periodismo. Desde coberturas en tiempo real hasta profundas investigaciones colectivas pueden verse afectadas, eliminadas, marcadas, desindexadas o invisibilizadas por esta regulación privada, muchas veces sin que medien explicaciones claras sobre los motivos de tales decisiones. La libre circulación de información de interés público y el derecho a informarse de la ciudadanía están en manos de moderadores que en la mayoría de los casos no son humanos ni actúan en función de proteger el interés general.



02.

Moderación de contenidos en redes sociales

Las redes sociales son plataformas digitales que se basan en contenidos generados por terceros que se intercambian entre los miembros de esas comunidades. Por lo tanto, son intermediarias necesarias para que esa información se publique, se comparta y circule. Esto funciona igual para fotografías y reflexiones personales, como para el entretenimiento o la información que comparten a través de las redes sociales periodistas y medios de comunicación.

Pero estos espacios no son neutrales ni están libres de regulación. Cada plataforma tiene sus propias normas y condiciones de servicio a través de las que establecen lo que consideran apropiado o no para ser publicado y donde enumeran las sanciones que aplica la red social a los usuarios que no cumplan con estas reglas. A través de sus normas comunitarias, que son reglas privadas, indican, entre otras cosas, lo que se puede publicar y lo que no. Asimismo, establecen qué tipos de contenidos tendrán mayor o menor alcance. Es decir: actúan como moderadoras.

La moderación es algo que no se ve, aunque ocurre cada vez que un usuario sube un contenido a una plataforma. Desde la experiencia del usuario, se percibe que el elemento que sube -ya sea un texto, una foto, un documento o un video-, se publica automáticamente, sin filtros. Sin embargo, entre que el usuario lo sube y el elemento se publica ocurre la moderación que analiza si los contenidos publicados cumplen con las políticas de uso de la plataforma.

Los términos de servicio son un acuerdo legal entre quien provee la plataforma y la persona que quiere usarlo y que, para eso, debe aceptar regirse por las reglas que impone la otra parte. Por su parte, las normas comunitarias

son las reglas y guías que las plataformas establecen acerca de lo que es aceptable que se publique en ella. Este concepto deriva del término legal *community standards* que surgió a partir de la decisión de la Corte Suprema de 1957 del caso *Roth vs. EEUU*, donde la Corte sostuvo que un material podía considerarse obsceno “si para la persona promedio, aplicando los estándares comunitarios contemporáneos, el material es lascivo”. Tiene que ver con normas sociales y culturales que definen los valores de una comunidad sin necesidad de ser llevados a una ley o contrato explícito.

Estas decisiones impactan en la forma en que unos usuarios acceden a lo que otros publican. Sin embargo, muchas veces puede resultar difícil encontrar las normas o, en la práctica, es casi imposible que un usuario pueda leerlas en su totalidad, debido a su gran extensión. Además, en algunos casos puede pasar que no todos los textos estén escritos en el idioma del usuario, por lo que se vuelven inaccesibles. Todo esto hace que no sea tan claro para los usuarios cuáles son los criterios que los guían y, mucho menos, los procesos sobre cómo se toman cada una de las decisiones en cuanto a los contenidos que publican.

Las reglas de moderación de las plataformas deberían presentarse de forma clara, precisa, comprensible, de fácil acceso y en el idioma del usuario, para que pueda entenderlas cualquier persona. En este sentido, se debería informar claramente cuáles son las condiciones que pueden llevar a la rescisión de ese acuerdo (que se elimine la cuenta del usuario, por ejemplo) o a la remoción del contenido publicado, entre otras medidas posibles que impacten en los derechos del usuario de la plataforma.

a o ¿Cómo funciona la moderación?

Teniendo en cuenta que la moderación tiene un impacto central en la circulación de la información, es importante comprender cómo funciona en las redes sociales y, para eso, primero es necesario conocer de qué forma circula el contenido en estas plataformas. Esto podría clasificarse de la siguiente manera: alcance orgánico, alcance pago y microsegmentación.

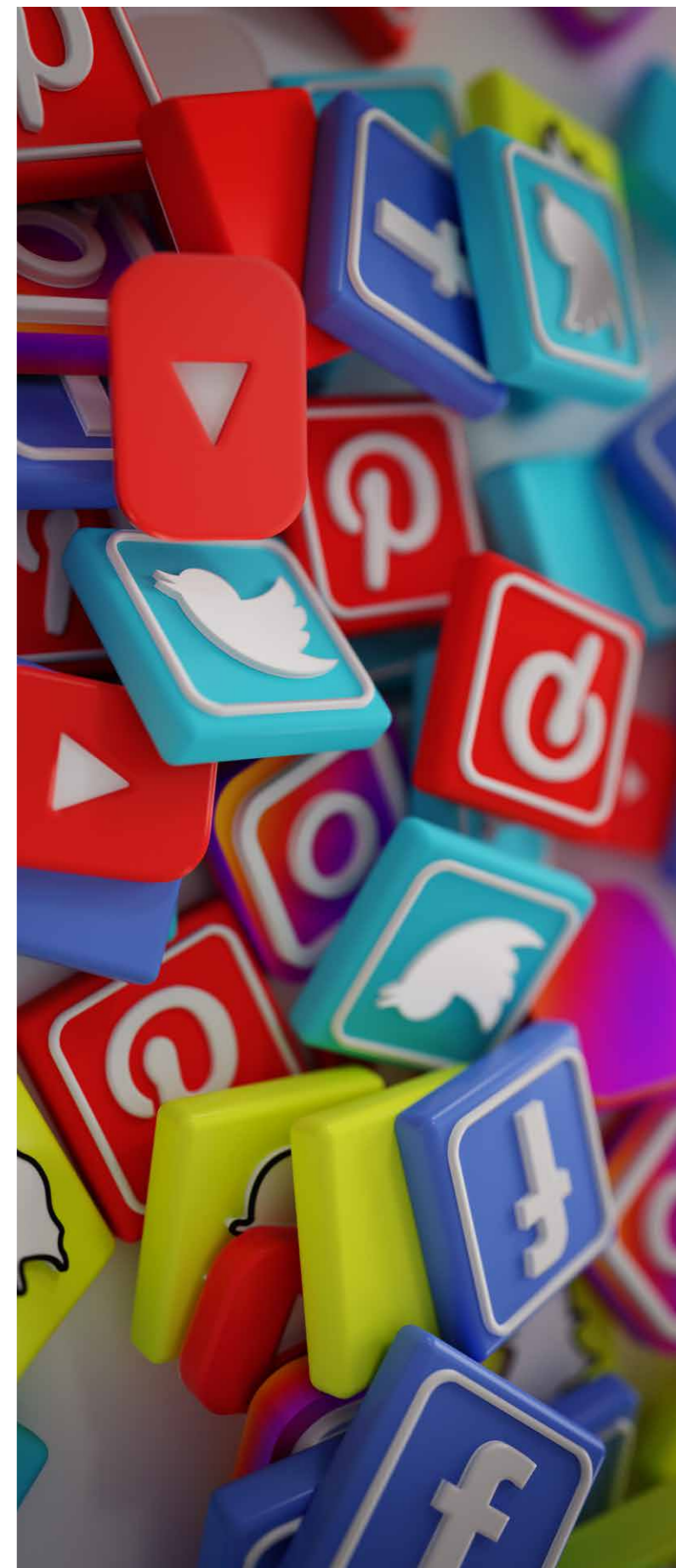
El alcance orgánico se define por el número de contactos que alcanza una publicación de forma natural. La diferencia con el alcance pago es que, como indica su nombre, involucra una inversión de dinero en la plataforma para aumentar la cantidad de contactos alcanzados.

El alcance pago puede definirse a partir de un objetivo o de un monto de inversión en relación al tiempo. Cuando una persona paga por el alcance de una publicación puede, y es uno de los diferenciales más importantes de la publicidad en plataformas sociales, definir las características de los segmentos de usuarios que estarán expuestos a esa publicación y definirlos con enorme detalle y precisión.

La microsegmentación es, entonces, la capacidad de alcanzar a determinados segmentos de los usuarios en función del objetivo que se tenga.

En el alcance orgánico, entonces, el número de personas alcanzadas dependerá de los seguidores de una cuenta (en general, entre 3% y 20% del total de seguidores del usuario, no más) y las acciones de moderación sobre ese contenido que ocurran.

¿Pero quiénes y cómo moderan los contenidos en las redes sociales? Las plataformas llevan adelante acciones concretas para verificar y asegurar que los contenidos que se comparten a través de sus servicios cumplan con sus reglas. Para eso tienen equipos humanos que revisan miles de contenidos diarios para definir si respetan las normas comunitarias o infringen alguno de los puntos de los términos de servicio. De hecho, el trabajo de moderación de la enorme diversidad de contenidos que se suben a las redes sociales



¹ "#TusDerechosEnInternet ¿Qué es la 'moderación de contenidos' en redes sociales?", por OBSERVACOM. Disponible en: <https://www.observacom.org/tusderechoseninternet/#MODERACION>

² Decisión de la Corte Suprema de 1957 del caso *Roth vs. EEUU*. Disponible en: <https://tile.loc.gov/storage-services/service/l/usrep/usrep354/usrep354476/usrep354476.pdf>

(discurso de odio, ataques violentos y pornografía gráfica, entre otros) ha sido noticia en los últimos años debido a los impactos nocivos en la salud mental de las y los moderadores.

Algunos de los contenidos que las plataformas consideran que deben ser sancionados, son: pornografía, abuso sexual infantil documentado, sexo explícito, desnudos, discurso de odio, cualquier acto de violencia, apología del delito, acoso sistemático, desinformación, y el incumplimiento de las normativas de derechos de autor y copyright.

Sin embargo, los moderadores humanos no trabajan solos. Debido al creciente volumen de información que circula en las redes y a las exigencias por parte de diversos actores sociales para que ejerzan una moderación más estricta de determinados contenidos -como los vinculados a discurso de odio o desinformación-, las plataformas han ido incorporando el uso de técnicas de machine learning, inteligencia artificial y algoritmos para revisar el contenido que se sube a diario.

Esta situación presenta un nuevo problema, ya que no siempre es preciso qué contenido está infringiendo las normas comunitarias de la plataforma. Para determinarlo, es fundamental analizar el contexto de cada publicación, algo que todavía resulta un enorme desafío para los sistemas automatizados de moderación de contenidos. En este sentido, la ausencia de supervisión humana sobre esa moderación aumenta los riesgos de falsos positivos y falsos negativos.

Independientemente de cómo actúan los moderadores humanos o los sistemas automatizados de moderación de contenidos, está también el cumplimiento o no de las reglas por parte de la misma compañía.

A pesar de que los términos del servicio y las condiciones de uso de Facebook se presume que son iguales para todos sus usuarios, en 2020, a través de una filtración de documentos internos de la compañía que fueron publicados por el medio de comunicación The Wall Street Journal, se conoció una lista de 5.800.000 personas llamada XCheck o Crosscheck que se trata de una serie de usuarios VIP de la red social que reciben un trato diferencial con respecto al contenido que publican. Es decir, si uno de estos usuarios VIP publica un contenido que viola las reglas de la red social, este no es eliminado o no es tan fácil ni tan frecuente que suceda.

Ese mismo año, en octubre, Facebook formó un Consejo Asesor de Contenido (Oversight Board) para resolver apelaciones de casos de decisiones de la plataforma que podrían haber afectado la libertad de expresión online y establecer recomendaciones sobre qué contenido eliminar, cuál conservar y por qué.

Pero la eliminación de los contenidos, una de las medidas más extremas y visibles, no es la única decisión que se toma desde la moderación privada. Existen diversos tipos de sanciones entre los que se encuentran mecanismos más sutiles y hasta desconocidos por la mayoría de los usuarios, pero que pueden tener el mismo resultado de afectación de la libertad de expresión en línea.

b. Tipos de moderación

Dentro de lo que se conoce como moderación existe un amplio abanico de opciones. Los contenidos que los usuarios crean, publican o comparten en las redes sociales pueden ser marcados, ocultados e, incluso, eliminados. A su vez, las cuentas y perfiles que crean, y que son su medio de comunicación en ese espacio, pueden ser suspendidas temporalmente o bloqueadas de manera definitiva. En el presente documento se hará referencia a los siguientes tipos de acciones que pueden tomar las plataformas: jerarquización o priorización, recomendación, reducción de alcance, invisibilización o shadowban, etiquetado o flagging, remoción de contenidos, suspensión y eliminación de cuentas



³ "The trauma floor. The secret lives of Facebook moderators in America" por Casey Newton, para The Verge. Publicado el 25 de febrero de 2019. Disponible en: <https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>

⁴ <https://www.oversightboard.com/>

JERARQUIZACIÓN O PRIORIZACIÓN

Una de las principales acciones de moderación privada que realizan todas las redes sociales es la jerarquización o priorización de lo que vemos en el inicio de cada una. Establecen el orden en que veremos lo que otros usuarios publican. El criterio de ese orden puede variar entre las diversas plataformas e incluso algunas, en los últimos años, han empezado a ofrecer la posibilidad a los usuarios de elegir entre dos o tres criterios para definir ese orden.

En sus orígenes, redes sociales como Facebook o Twitter ordenaban las publicaciones en forma cronológica y solo mostraban los contenidos de la red de contactos del usuario. Con el tiempo, las empresas optaron por establecer la jerarquización o priorización de los contenidos a partir de algoritmos de recomendación. En lugar de ordenarlo en forma cronológica y mostrar exclusivamente lo que publicaban los contactos de la red del usuario, empezaron a mostrar contenidos “destacados”, de acuerdo a lo que los algoritmos de recomendación seleccionan para cada usuario. Muestran lo que le gusta al usuario, dicen las empresas. Los algoritmos aprenden a partir del comportamiento del usuario (a quién le comentan, le dan “me gusta”, re-tuitean) y de sus decisiones (si silencian contactos, si bloquean palabras determinadas o a otros usuarios).

Otra forma de establecer una jerarquización entre los contenidos que se publican en internet es la propia indexación de esos contenidos, principalmente -aunque no únicamente- por los motores de búsqueda. Debido a la sobreabundancia informativa que existe en internet, el rol de los buscadores en el acceso a esa información es fundamental. Son los que permiten, en primera instancia, encontrar aquello que un usuario busca en inter-

net. Es decir, que si un buscador decide activamente no indexar contenidos provenientes de un determinado sitio web o decide excluir de sus resultados de búsqueda cualquier contenido que incluya una palabra (o conjunto de palabras) o una imagen, esta decisión tiene un efecto concreto en el acceso a esa información por parte de los usuarios.

RECOMENDACIÓN

Las redes sociales son espacios a través de los que se publica y comparte una gran cantidad de información y contenidos que responden a diversas categorías, géneros, estilos y calidades, desde algo casero y amateur, hasta los más profesionales. A través de los datos que recopilan del comportamiento pasado de sus usuarios (en qué enlaces hace clic, qué publicaciones le gustan o comparte) y de sus bases de perfiles de usuario, las plataformas pueden identificar, con mayor o menor margen de error, qué contenidos son interesantes y relevantes para cada usuario.

De esta manera, sus algoritmos recomiendan contenidos que se presume que serán interesantes para el usuario, por lo que se mantendrá más tiempo conectado y atento. La contraparte de este proceso es que a ese usuario no se le recomendarán otros contenidos u otra información que la plataforma considere que no es interesante para el mismo, sin importar si es un contenido relevante por su valor social, si aporta una mayor diversidad de puntos de vista o si es de interés público.

REDUCCIÓN DE ALCANCE

En algunos casos, por publicar algún contenido que infringe las normas comunitarias, la plataforma puede tomar la decisión de reducir al mínimo la cantidad de usuarios que pueden verlo. Es decir que, a pesar de que el contenido no es estrictamente eliminado de la red social, el impacto de esta decisión es parecido, ya que, aunque siga publicado, muy pocos usuarios lo van a ver.

A pesar de ser un tipo de sanción que algunas plataformas utilizan, el usuario no recibe ninguna notificación que le avise que sus publicaciones tendrán menor alcance. Debido a esta falta de aviso, el usuario no tiene conocimiento informado acerca del problema, ya que él ve que su contenido sigue publicado y continúa recibiendo interacciones (me gusta o comentarios). Sin embargo, estas interacciones son menos de las que podría recibir si su contenido se mostrara y circulara libremente por la red social.

INVISIBILIZACIÓN O SHADOWBAN

Se conoce como shadowban o ghostban a la invisibilización que se produce a partir del bloqueo de un usuario en una plataforma sin recibir ninguna notificación en la que se explicita que se tomó esta medida y sin que éste perciba que está siendo bloqueado. El shadowban no elimina el contenido o la cuenta del usuario, pero le resta relevancia disminuyendo su visibilidad casi a cero. Esto significa que quien sube contenido sigue viéndolo, pero ignora que probablemente sea el único que puede hacerlo.

Una de las técnicas que más se utilizan para lograrlo es no mostrar las publicaciones de un usuario en la lista de posts de un hashtag relacionado, sino solo mostrarlo a quienes son sus seguidores. El efecto de esta medida es impedirle formar parte de la conversación global y sumar nuevos seguidores, lo que disminuye su capacidad de influenciar sobre ese u otros temas relacionados.

⁵ Centro de ayuda de Twitter, Acerca de las etiquetas de cuentas automatizadas <https://help.twitter.com/es/using-twitter/automated-account-labels>

ETIQUETADO O FLAGGING

El etiquetado o flagging puede referirse a la marca que agrega la propia plataforma a contenidos de terceros o sus cuentas. Es una herramienta que usan las redes sociales para identificar contenidos que ellos consideran de dudosa procedencia o veracidad y hacerlos visibles para sus usuarios. En este sentido, el etiquetado es una forma que las plataformas digitales tienen de “editar” los contenidos publicados en su espacio, lo que da cuenta de que no son un mero soporte técnico para la circulación de contenidos de terceros.

En los casos en los que buscan detectar noticias falsas o desinformación, Meta, por ejemplo, trabaja con organizaciones de periodistas que se especializan en la verificación de datos (fact-checkers) que analizan contenidos de interés público. Si se considera que una publicación es falsa, la plataforma la etiqueta y, en algunos casos, puede vincularla con la información que la contrasta. Es decir que no se elimina este contenido, sino que aparece un mensaje que avisa a los usuarios que una parte o la totalidad de esa publicación ha sido calificada como falsa.

También en casi todas las plataformas se utiliza el flagging para marcar contenido violento o sensible que por sus características es considerado por los equipos de moderación como de interés público y, por ende, no eliminable de la plataforma.

En el caso de Twitter, hace tiempo que también se implementa el etiquetado de cuentas para identificar o dar contexto al contenido que publican algunas de ellas. Los casos en los que esto se hace con regularidad son las etiquetas a las cuentas au-

tomatizadas (denominadas “bots”), precisamente para diferenciarlas de aquellas que son administradas directamente por personas. En este sentido, en febrero de 2022, la empresa anunció que iba a habilitar una etiqueta para destacar a los “bots buenos” que cumplen con el rol de informar y ofrecer contenido de calidad a los usuarios.

Desde 2020, la plataforma también aplica etiquetas a cuentas que considera de medios afiliados al gobierno y al Estado de diversos países. En su centro de ayuda, Twitter explica que estas etiquetas “proporcionan más contexto sobre las cuentas controladas por determinados representantes oficiales de gobiernos, entidades de medios afiliadas al Estado y personas relacionadas con esas entidades”.

REMOCIÓN DE CONTENIDOS

Una de las medidas más graves en lo que se refiere a la moderación de contenidos en redes sociales es la remoción de contenidos. En los últimos años, este mecanismo se ha utilizado cada vez más con la justificación de proteger los intereses de aquellas personas titulares de derechos de propiedad intelectual (por ejemplo, derechos de autor) e incluso de derechos vinculados a la privacidad y la protección de datos personales.

Esta práctica de las redes sociales se origina a través de dos vías principales. En primer lugar, a través de solicitudes de remoción de contenidos a las plataformas digitales, que pueden estar realizadas por parte del Estado o de actores privados. En segundo lugar, puede ser por la ejecución de sanciones que impone la plataforma por incumplimiento de sus términos de servicio o normas comunitarias de las plataformas.

La remoción o eliminación de contenidos publicados por un usuario se trata de una medida que las plataformas toman de manera unilateral y, en general, sin notificación antes de la aplicación de la medida ni la posibilidad de ejercer defensa previa por parte del autor del contenido. La remoción puede ser definitiva o temporal. En el último caso, la decisión puede estar sujeta a modificaciones en el contenido publicado originalmente.

SUSPENSIÓN DE USO

En algunos casos, las plataformas mantienen visibles las cuentas que han recibido una sanción, pero impiden a los administradores algunas acciones como por ejemplo publicar contenido, durante un tiempo determinado o hasta que se elimine una publicación específica.

Por ejemplo, en el caso de Twitter, si se considera que una cuenta infringió sus reglas de funcionamiento, la plataforma puede solicitar la modificación del perfil y suspender el uso de la cuenta hasta tanto no se haga ese cambio. Es decir, puede configurar la cuenta para que sea de solo lectura y la persona esté impedida de generar ninguna clase de contenido. Esta modificación puede durar de 12 horas a 7 días y es el resultado de lo que la plataforma llama “comportamientos abusivos”. La medida no incluye la suspensión del uso de los mensajes directos.

SUSPENSIÓN Y ELIMINACIÓN DE CUENTAS

En los casos en que la empresa considere que un usuario ha infringido las normas comunitarias de la plataforma más allá de un contenido puntual, puede aplicar una suspensión temporal de acceso a su cuenta y a todos sus contenidos. Esto puede revertirse, de acuerdo a la consideración de la empresa, luego de que el usuario aplique una serie de ajustes que se le soliciten para cumplir con las reglas.

La suspensión permanente de una cuenta es una medida más seria que hará que se elimine la vista de la cuenta e impedirá la creación de nuevas. En este caso se informa al usuario acerca de la suspensión, así como cuáles fueron las normas incumplidas y en qué contenidos.

Un paso más extremo es la eliminación de cuentas. Esto ocurre en los casos en que las plataformas entienden que las infracciones a los términos del servicio fueron de una importante gravedad o se trata de una cuenta reincidente. En estos casos, además de la eliminación de la cuenta, se eliminan en forma unilateral todos los contenidos generados por esa cuenta mientras estuvo activa, sin considerar el valor social o cultural que pudieran tener algunos o todos esos contenidos para el resto de los usuarios y de la comunidad.



03.

La importancia de las redes sociales para los medios de comunicación y los periodistas



La importancia de las redes sociales para los medios de comunicación y los periodistas

En octubre de 2015, 66,9 millones de usuarios visitaron el sitio del diario norteamericano The Washington Post, mientras que en agosto de ese mismo año Facebook anunció que 1.500 millones de personas habían iniciado sesión en su plataforma. Esto dejó claro, hace varios años, que uno de los medios más prestigiosos y visitados del mundo, recibe apenas 4% del tráfico que circula dentro de una sola de estas redes sociales.

A pesar de que en sus orígenes no fue así, Internet es hoy un territorio intermediado. Un pequeño oligopolio de transnacionales que en su mayoría generan ingresos a partir de la publicidad y monopolizan lo que la mayoría de los usuarios llaman “navegar por internet” en el mundo occidental. Aunque a primera vista el surgimiento de Internet y las plataformas sociales le dio a los medios y periodistas una oportunidad enorme de alcanzar a más personas y en lugares lejanos, en los hechos esta potencialidad de la red de redes depende cada vez más de la intermediación de estas plataformas y sus modelos de negocios, de sus decisiones de moderación de contenidos y de cómo los medios y periodistas se adaptan a éstas.

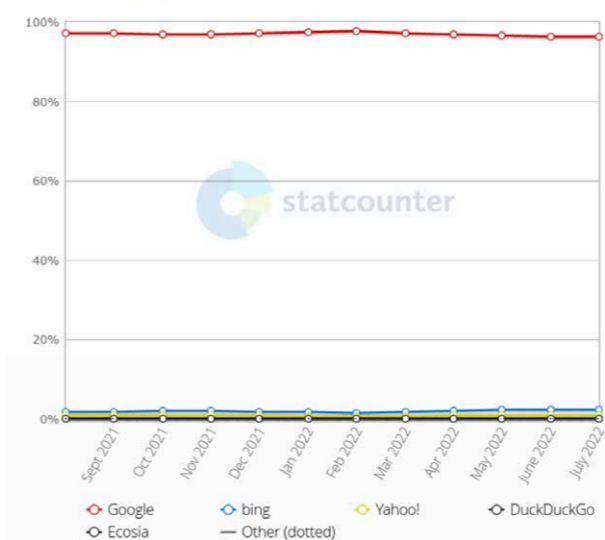
Las consecuencias de no jugar con las reglas de plataformas como Facebook, Twitter, YouTube y Google tienen un enorme impacto en el futuro de los medios del mundo y pueden ir desde ver cómo su audiencia se reduce significativamente hasta directamente perder un canal o una cuenta, así como todos sus archivos históricos.

⁶ <http://digiday.com/publishers/comscore-washington-post-tops-new-york-times-online-first-time-ever/>
⁷ <http://www.bbc.com/news/world-us-canada-34082393>

Noticias, medios y buscadores

Google es el principal motor de búsqueda de informaciones y cualquier otro contenido en Internet, a tal punto que sus competidores sumados no logran alcanzar el 10% de las búsquedas que cada día realizan usuarios de todo el mundo. Un 96,3% de los millones de búsquedas realizadas en Sudamérica en agosto de 2022 ocurrieron en el buscador de Google y, en ellas, 28% de las veces los usuarios hicieron clic en el primer resultado, sin mirar ninguno de los ubicados debajo. Más del 91% de los sitios web del mundo no reciben ningún tráfico orgánico (que no implique pagar al buscador para aparecer en los resultados destacados).

Search Engine Market Share
South America
Aug 2021 - July 2022



El buscador rankea millones de sitios de todo el mundo para ubicarlos en un orden determinado al entregar resultados de cada búsqueda. Ese orden está determinado por algoritmos que han sido modificados miles de veces desde su nacimiento y si bien Google hace públicos algunos criterios generales, nadie sabe a ciencia cierta cómo funcionan las reglas sobre las que decide mostrar algunas cosas y no otras.

⁹ Las compañías más grandes de Internet https://en.wikipedia.org/wiki/List_of_largest_Internet_companies
¹⁰ <https://blog.google/products/search/how-google-delivers-reliable-information-search/>

Google asegura que su algoritmo está orientado a ofrecer información de calidad a quienes realizan búsquedas. Define “calidad” sobre la base de tres conceptos: un ranking diseñado para identificar la información que es más probable que las personas encuentren útil y confiable, un sistema de identificación de “autoridades” cuyos sitios aparecen destacados para preguntas específicas (como por ejemplo gobiernos o Ministerios de Salud) y finalmente Google asegura que tiene “políticas acerca de que puede aparecer para asegurarse que se está mostrando alta calidad y contenido útil”. Sobre cómo se hace eso, bajo qué parámetros, qué controles, tasas de error y acierto, poco sabemos.

Los medios de comunicación y periodistas dependen en gran medida de Google para que millones de usuarios de todo el mundo accedan a sus contenidos y es por esto que las redacciones han creado, desde hace unos años, figuras de análisis SEO (Search Engine Optimization). El SEO es el conjunto de saberes, técnicas y herramientas utilizadas para posicionar contenidos lo más alto posible en los resultados de búsqueda de Google. El SEO implica desde el uso de técnicas de redacción (repetición o destaque de palabras y conceptos, tipos de titulación, uso de términos específicos, etc.), hasta la mejora tecnológica del sitio que aloja los contenidos para acelerar la velocidad de carga o colaborar en el trabajo de Google para la indexación acelerada de contenidos.

Este trabajo se realiza en gran medida a ciegas, debido a la ya mencionada falta de transparencia de Google respecto a los cambios que realiza en la distribución de contenidos en los resultados, cambios que a veces tienen un impacto muy significativo en el tráfico de los sitios de noticias de un día para otro.

A modo de ejemplo, el 18 de agosto de 2022, Google informó en su blog de una actualiza-



ción de su algoritmo de búsquedas que ocurriría la semana siguiente con la intención de asegurarse de penalizar contenido “que parece creado primero para posicionar mejor en buscadores más que para ayudar a la gente a informarse”. Esta actualización “aseguraré que el contenido no original y de baja calidad no se ubique arriba en los resultados”. Google no explica a qué se refiere con contenido “no original y de baja calidad”, con qué parámetros lo define y tampoco dice en esa publicación que esto aplica al contenido orgánico dado que los primeros resultados que una persona ve al buscar seguirán siendo de aquellos que pagan para aparecer primero, sin importar la calidad de los sitios que tengan¹¹.

Pero Google no intermedia el consumo de noticias o contenidos periodísticos solamente a través de su buscador. También a través de su sistema de recomendación personalizada para teléfonos, conocido como Google Discovery, así como a través de los sistemas de recomendación algorítmica con los que recomienda contenidos en YouTube. En este último, 70% de los contenidos consumidos provienen de esos sistemas de recomendación. Allí también se hace necesario que los periodistas y medios adapten sus temas y las formas en las que los plantean utilizando técnicas SEO como forma de alcanzar a las audiencias.

Noticias, medios y redes sociales

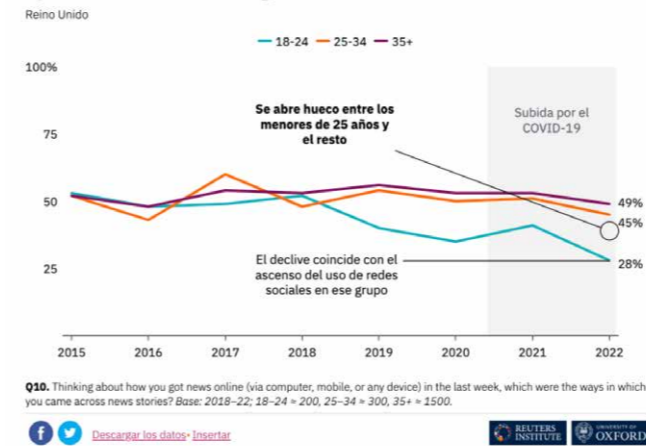
Las redes sociales también se han transformado en una de las principales fuentes de tráfico para los medios. Muchos usuarios leen noticias a partir de verlas publicadas en los muros o líneas de tiempo de sus contactos, de influencers o de medios de comunicación dentro de esas plataformas. Si se enfoca este comportamiento con una mirada generacional se verá incluso que el fenómeno se vuelve más y más marcado a medida que se vuelve más joven el grupo analizado.

Según el 2022 Digital News Report del Instituto Reuters, de todos las personas consultadas para el informe (46 países, 6 continentes y más de 93.000 personas entrevistadas), menos de una cuarta parte (23%) prefiere comenzar la jornada informativa con un sitio web o una app de un medio o periodista. Esto representa una caída de nueve puntos desde 2018. En todos los países se encontró que desde 2019 las redes sociales (39%) han superado a los sitios web y las aplicaciones (31%); luego siguen la pantalla de inicio del móvil (12%) y agregadores como Apple News (9%).

Si se analiza a quienes tienen entre 18 y 24

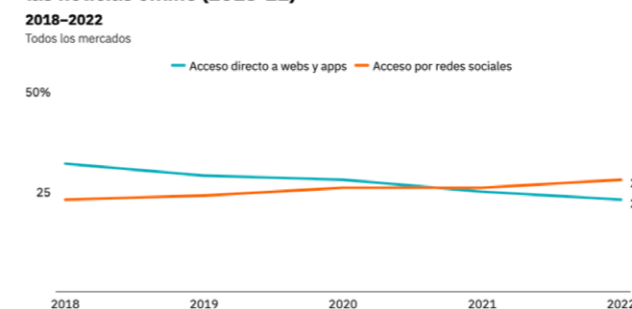
años, la conexión es aún más débil ya que, en su mayoría, aseguran que prefieren acceder a las noticias “mediante caminos alternativos como las redes sociales, las búsquedas y los agregadores móviles”.

Porcentaje que comienza su jornada informativa con un sitio o una aplicación de noticias, según edad



En todos los dispositivos y los usuarios, los datos del informe de Reuters indican que pierde importancia el acceso directo a las aplicaciones y los sitios web de los medios de comunicación, mientras que las redes sociales crecen, “en parte debido a su ubicuidad y conveniencia”. 2022 se transformó en un punto de inflexión en cuanto al acceso a noticias en Internet, ya que la preferencia por las redes sociales (28%) ha superado al acceso directo (23%). De todas formas, se trata de promedios de los mercados, y entre ellos se presentan grandes diferencias. Las audiencias en los países nórdicos y en el Reino Unido aún poseen fuertes conexiones directas, en tanto que en los mercados de México y Brasil la caída es incluso más pronunciada, llegando a apenas 11% de tráfico directo. En Perú, 73% de los encuestados dicen consumir noticias a través de redes sociales, 3 puntos porcentuales más que en el estudio de 2021.

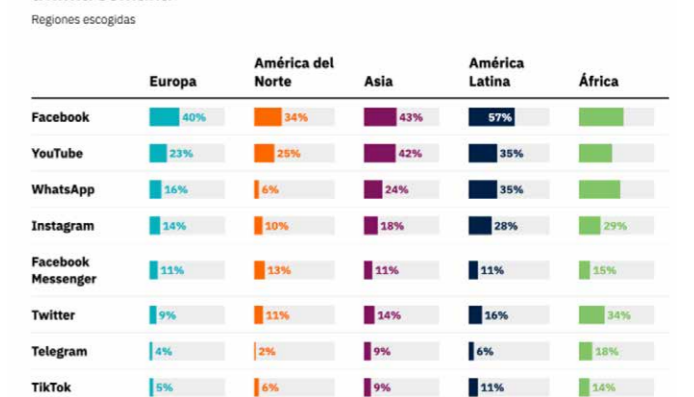
Porcentaje que dice que cada opción es su principal vía de acceso a las noticias online (2018-22)



Si se analizan redes específicas, el estudio de Reuters muestra que Facebook ha caído notoriamente en su peso en el consumo de noticias, se ha derrumbado 12 puntos porcentuales desde 2016, pero se mantiene como la red más importante. Twitter tiene en gran medida su base de usuarios estancada durante la última década, aunque conserva enorme influencia entre periodistas y políticos. Instagram, propiedad de Meta, ahora se usa más para las noticias, mientras que TikTok ha superado a Snapchat, aunque con una base baja.

En Argentina, 56% de los usuarios usan Facebook para leer noticias, mientras que 35% usan WhatsApp y 31% lo hacen a través de Instagram. En Perú, es el 67% de los usuarios de Facebook, 39% de YouTube y 34% de WhatsApp. Esto implica que muchos de quienes leen noticias en los medios de ese país llegan a ellos a través de las redes sociales. De unas poquísimas redes sociales.

Porcentaje que utilizó cada red social para consumir noticias en la última semana



Q12b. Which, if any, of the following have you used for news in the last week? Base: Total sample in each region. Europe = 48,836, North America = 4048, Asia = 20,349, Latin America = 12,104, Africa = 6057. Note: Africa average is Kenya, South Africa, and Nigeria only (English speakers in South Africa and Nigeria).

Descargar los datos - Insertar

¹¹ <https://blog.google/products/search/more-content-by-people-for-people-in-search/>
¹² <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022>

países que no son los suyos y no siempre alineados con las leyes de los países en los que esos medios están ubicados. E implica hacer periodismo todos los días sabiendo que es posible perder a muchos de sus usuarios si las plataformas entienden que se ha infringido alguna de esas normas o, incluso, deciden de un día para otro cambiar las reglas.

Un caso paradigmático que permite ilustrar el impacto de la moderación de contenidos por parte de las plataformas en el periodismo y los periodistas fue el apagón de noticias llevado adelante por Facebook como mecanismo de presión frente a una ley con la que Meta no estaba de acuerdo. En febrero de 2021, los australianos dejaron de poder ver páginas de Facebook de medios locales e incluso de poder publicar links de noticias publicadas en ellos. La medida fue tomada por Meta en rechazo a un proyecto de ley impulsado por el gobierno australiano para que la empresa de Mark Zuckerberg, así como también Google, debieran pagar a los medios por los enlaces que se publicaban. La decisión desactivó la principal fuente de noticias para casi uno de cada cinco australianos y silenció, según la empresa accidentalmente, una serie de páginas gubernamentales que contenían consejos de salud pública sobre el coronavirus, advertencias de la oficina meteorológica e incluso la de un hospital infantil.

Facebook dio marcha atrás a la medida días después, pero tras obtener el compromiso del gobierno australiano de que dejaría a Facebook y Google negociar con los medios en vez de obligarlos a pagar. Esta medida puso sobre la mesa no solo el impacto que tiene la mediación de las plataformas en la llegada de noticias a los ciudadanos sino también el poder de estas empresas que, como se vio claramente en este episodio, están dispuestas a privar a miles y miles de usuarios de información de calidad para mantener su lucro incambiado.

Si ésta es la asimetría de poder frente a los nuevos distribuidores de contenido a la que se enfrentan los medios más importantes de todo el mundo, ¿en qué situación pueden encontrarse los medios comunitarios, más pequeños y con menos peso dentro de sus propias comunidades? Muchos de ellos utilizan plataformas como YouTube, Facebook o Twitter como la única forma de hacer contacto con sus audiencias, produciendo específicamente contenido pensado para esas plataformas e incluso usándola como únicos repositorios de sus archivos.

Muchas veces los periodistas y medios no cuentan con la información necesaria para saber qué hacer el día que encuentran sus cuentas en redes cerradas o cuándo algunos de sus contenidos son eliminados. En muchos de los casos, las plataformas no comunican con claridad acerca de qué normas específicas fueron infringidas y esto hace muy difícil iniciar un proceso de apelación con alguna chance de éxito. Si no sabemos qué se dice que hicimos mal, ¿cómo podemos defendernos?

Si efectivamente se inicia un proceso de apelación frente a decisiones consideradas injustas, en muchos casos esto implica tener que aceptar legislación y procesos judiciales en países distintos a donde viven quienes son sancionados y reclaman por ellos, sino donde se encuentran las oficinas centrales de las plataformas (generalmente en EEUU).

Frente a estos escenarios, muchas veces los periodistas y pequeños medios optan por aceptar las sanciones y adaptarse a esas reglas. Creando nuevas cuentas, dejando de postear sobre contenidos específicos, haciéndolo de formas que no eligen, sino que les son impuestas.



¹³ <https://elpais.com/tecnologia/2021-02-18/facebook-recrudece-su-pulso-global-a-los-editores-con-un-apagon-de-noticias-para-18-millones-de-australianos.html>

¹⁴ <https://www.bloomberg.com/news/articles/2021-02-23/facebook-to-restore-news-pages-in-australia-in-coming-days>



04.

Ejemplos acerca del impacto de la moderación en el periodismo

En tanto en el ejercicio concreto de los procesos, diagnósticos y definiciones de la moderación llevada adelante por las plataformas es donde se producen los impactos más significativos en el periodismo y la libertad de expresión, es que las autoras de este trabajo incluimos a continuación cuatro casos que involucran periodistas y medios latinoame-

ricanos. El análisis de estos casos funciona como un modo de poner sobre la mesa los problemas que surgen del efectivo rol de moderar el debate público que han tomado para sí las plataformas -sin transparencia, debido proceso ni supervisión alguna- y del impacto que esto tiene para el ejercicio efectivo del derecho a informar e informarnos.

CASO 1

Taggeo de perfiles de periodistas independientes como medios afiliados al gobierno de Rusia en Twitter. El caso Marco Teruggi de Argentina



Desde 2020, Twitter aplica etiquetas a cuentas que considera de medios afiliados al gobierno y al Estado de diversos países. Por ejemplo, medios como RT o Sputnik comenzaron a ser identificados con la etiqueta “Medios afiliados al gobierno, Rusia”. En su centro de ayuda, Twitter explica que estas etiquetas “proporcionan más contexto sobre las cuentas controladas por determinados representantes oficiales de gobiernos, entidades de medios afiliadas al Estado y personas relacionadas con esas entidades”.

Según la compañía, las etiquetas se aplican a las cuentas que pertenecen a “medios donde el Estado ejerce control sobre el contenido editorial mediante recursos financieros, presiones políticas directas o indirectas o el control sobre la producción y distribución” y las etiquetas son para entidades de medios afiliadas al Estado, a sus editores en jefe y su personal de mayor nivel jerárquico. A fines de febrero, luego de que Rusia hubiera dado inicio al conflicto bélico en Ucrania, se conocieron una serie de cuentas que habían sido etiquetadas como “Medios afiliados al gobierno,

Rusia”. Sin embargo, no cumplían con estos requisitos, sino que se trataba de las cuentas personales de periodistas que colaboraban para algunos de ellos de manera freelance e independiente.

En el caso de los impactos sobre la circulación de los contenidos, la decisión es que Twitter no recomienda ni amplifica entre los usuarios las cuentas o Tweets que estén marcados con estas etiquetas. Esto fue lo que le sucedió a Marco Teruggi, periodista independiente argentino, que en ese momento trabajaba como corresponsal para Sputnik, y cuyo caso se describe en el presente documento.

El 28 de febrero de 2022, el periodista argentino Marco Teruggi entró a su cuenta personal de Twitter y se encontró con una sorpresa. En su perfil, la red social había colocado la etiqueta “medio afiliado al gobierno, Rusia”. Teruggi, periodista independiente que en ese momento trabajaba como corresponsal para la agencia rusa Sputnik, publicó su descargo en la misma cuenta pidiendo que fuera retirada esa etiqueta.



Esta etiqueta, como se mencionó antes, no cumple una función meramente descriptiva, sino que tiene efectos concretos en la circulación de los contenidos publicados por el usuario titular de la cuenta etiquetada.

La cuenta personal de Marco Teruggi, al igual que la de otras y otros periodistas del mundo que recibieron el mismo etiquetado por ser igualada con un medio afiliado al gobierno ruso, iba a tener una serie de restricciones en cuanto al alcance de sus publicaciones. La consecuencia concreta es que Twitter “no recomendará ni amplificará entre los usuarios cuentas o sus Tweets con estas etiquetas”, tal como dicen las condiciones de uso de la plataforma.

Después de que pasaran las primeras horas de haber notado la etiqueta que lo estigmatizaba como periodista afiliado al gobierno de Rusia, y sin haber recibido ningún aviso, información o aclaración por parte de la plataforma, Marco Teruggi se descargó en la red social: “Aún sin respuesta de @TwitterLatAm por haberme puesto ‘medio afiliado al gobierno ruso’ y las consecuencias que trae. Mi cuenta es personal, trabajo como corresponsal para varios medios, no escribí para Sputnik sobre lo que pasa en Ucrania. Somos varios en esta situación”.

Para un periodista, una plataforma como Twitter es más que una red social donde estar en contacto con otros usuarios y a dónde puede ir a buscar fuentes primarias para sus artículos o columnas. Es el espacio a través del que un medio de comunicación o un periodista difunde la información que tiene, el trabajo que realiza es donde, muchas veces, publica en tiempo real lo que por su carácter de urgente elige compartir en apenas unos pocos caracteres y es el espacio que, por su propia lógica, le permite amplificar su voz, sus escritos, sus publicaciones.

Durante 48 horas el periodista tuvo la etiqueta en su cuenta y, a pesar de intentar ponerse en contacto con la empresa para pedir explicaciones y solicitar que la retiraran, no había recibido una respuesta. Teruggi, además, puso en evidencia uno de los mayores inconvenientes que enfrentan los usuarios de la plataforma ante situaciones similares: “No existen formularios de reclamo ante esto”.

¹⁵ https://twitter.com/Marco_Teruggi/status/1498324426454036483

¹⁶ https://twitter.com/Marco_Teruggi/status/1498470843894785989

Es decir que, ante la medida unilateral de la plataforma -quizás a través de la decisión de un moderador humano o como resultado de un proceso automatizado-, el usuario no tiene forma de ponerse en contacto con la empresa ni para pedir explicaciones o más información, ni para reclamar ante una medida injusta cuyas consecuencias son la vulneración del ejercicio del derecho a la libertad de expresión y de prensa de ese periodista.

Estas etiquetas, que operan como “marcas”, no se limitan solo a las cuentas de los periodistas, sino que también a tuits que contienen enlaces a artículos de RT o Sputnik. Esta marca negativa es persecutoria y uno de sus efectos es la autocensura. Aquellos usuarios que temen que les cierren sus cuentas o que los etiqueten y puedan perder, entonces, la capacidad de hacer visibles sus trabajos y sus publicaciones, pueden verse forzados a no compartir enlaces a los sitios marcados o tuits de las cuentas etiquetadas.

Sobre este resultado de autocensura, la periodista Alejandra Loucau, especializada en noticias internacionales, relató lo siguiente: “Yo escribía para Sputnik News hasta el año pasado. Había escrito una nota sobre la Corte Penal Internacional (CPI) y los crímenes de guerra de Estados Unidos e Inglaterra. La iba a compartir, pero me contuve. Solo compartí el título de la nota para quien quiera buscarla. Me autocensuré porque no quiero que me bloqueen la cuenta o me la restrinjan de alguna manera”. A su vez, se refirió a cuentas de Twitter que se dedican a compartir información y que reconocieron comenzar a autocensurarse debido al mismo temor, sobre todo en cuanto a los videos o imágenes que suben.

La agencia china Xinhua también tiene su

cuenta marcada como medio del gobierno de China. La particularidad de estas etiquetas de “medios afiliados al gobierno” es su carácter arbitrario, ya que no todas las agencias de noticias vinculadas a gobiernos la tienen. Si bien es cierto que Twitter en su centro de ayuda menciona que “Las organizaciones de medios financiados por el Estado con independencia editorial, como por ejemplo la BBC en el Reino Unido o NPR en los EE.UU., no entran en la categoría de medios afiliados al Estado para los efectos de esta política”, existen otros medios financiados directamente por el gobierno de Estados Unidos y vinculados a estrategias de comunicación a favor de sus intereses como, por ejemplo, Radio Free Asia o Radio Free Europe/Radio Liberty que no están etiquetados como tales.

Igualmente, a pesar de lo cuestionable, en los casos de Xinhua, RT o Sputnik, la etiqueta se aplica sobre el medio de comunicación, es decir sobre la empresa. Lo llamativo y preocupante del caso Teruggi, al igual que el de otros colegas que también fueron etiquetados, como el periodista uruguayo Sergio Pintado, es que la plataforma aplica la etiqueta sobre las cuentas personales de periodistas independientes que trabajan para distintos medios. La periodista Alejandra Loucau destacó que “en el caso de los periodistas independientes o freelance es más grave, porque ya de por sí es una forma de trabajo precarizado”.

Desde las organizaciones Sindicato de Prensa de Buenos Aires (SiPreBA), la Federación de Periodistas de América Latina y el Caribe (FEPALC), y la Federación Internacional de Periodistas (FIP) se pronunciaron en contra de la medida tomada por Twitter y a favor de la libertad de prensa y de expresión.

CASO 2

Desindexación en Google de notas periodísticas de un medio digital. El caso Sudestada de Uruguay



En noviembre de 2021, un caso resonó en los medios de comunicación de Uruguay y rápidamente fue noticia en otros países de la región. A raíz de denuncias que recibió la empresa, Google removió de sus resultados de búsquedas dos investigaciones periodísticas del medio digital Sudestada que trataban sobre la participación de estudios jurídicos uruguayos en complejas estructuras internacionales de lavado de activos, conocidas como “Pandora Papers”. A partir de esta investigación colaborativa a nivel mundial realizada por diversos medios de comunicación se reveló cómo políticos, empresarios y otras personas influyentes utilizan guaridas fiscales para lavar dinero.

Un mes antes, el 11 de octubre de 2021, Google le había notificado a Sudestada la decisión de quitar de sus resultados de búsqueda (desindexar) una nota titulada “Lava Jato: estudio De Posadas fue un centro de blanqueo de los sobornos” que se había publicado el 1 de agosto de 2017. La decisión se debía, según el mensaje que la empresa había enviado por correo electrónico -en inglés- al medio de comunicación, a una presunta violación de la Digital Millennium Copyright Act (DMCA), que es una ley de derechos de autor de los Estados Unidos reconocida por ser el motivo de remoción de contenidos publicados que incluyen algún material protegidos por derechos de autor en otras plataformas, como YouTube.

En el correo enviado a Sudestada, la compañía aclaraba: “Estamos en el proceso de eliminar de los resultados de la búsqueda de Google el material que presuntamente infringe los derechos de autor de otros. Si no lo hacemos, podríamos estar sujetos a un reclamo por infracción de derechos de autor, independientemente de sus méritos”.



¹⁷ https://twitter.com/Marco_Teruggi/status/1499039371047149568

¹⁸ <https://www.tiempoar.com.ar/informacion-general/denuncian-que-twitter-censura-y-estigmatiza-medios-y-periodistas/>

¹⁹ <https://help.twitter.com/es/rules-and-policies/state-affiliated>

²⁰ https://es.wikipedia.org/wiki/Radio_Free_Asia

²¹ https://es.wikipedia.org/wiki/Radio_Free_Europe/Radio_Liberty

²² <https://twitter.com/sergiopintado/status/1498351517107200000>

²³ <https://twitter.com/sipreba/status/1498425820246953987>

²⁴ <https://ladiaria.com.uy/politica/articulo/2022/3/federacion-de-periodistas-latinoamericanos-rechaza-etiquetas-de-twitter-a-trabajadores-de-medios-rusos/>

²⁵ https://www.sudestada.com.uy/articledId_483deda0-9d05-4b59-af45-14a91e56529a/10893/Detalle-de-Noticia

En el correo enviado a Sudestada, la compañía aclaraba: “Estamos en el proceso de eliminar de los resultados de la búsqueda de Google el material que presuntamente infringe los derechos de autor de otros. Si no lo hacemos, podríamos estar sujetos a un reclamo por infracción de derechos de autor, independientemente de sus méritos”.



A pesar de que los contenidos no fueron eliminados del sitio original -es decir, que no se le pidió a Sudestada que bajara sus investigaciones de su propio sitio-, ser desindexado de los resultados del buscador más utilizado del mundo tiene un efecto concreto de censura, violando no solo el derecho a la libertad de prensa y de expresión en internet, sino también el derecho a la ciudadanía a estar informada.

La desindexación es una medida tomada en forma unilateral por Google antes de notificar al titular del enlace que será desindexado. Es decir, que el aviso por parte de la plataforma llega luego de que se efectúa la remoción del enlace de los resultados que arroja el motor de búsqueda a un usuario que esté buscando información sobre el caso. Es por eso que no hay manera de evitarla. Sólo se puede actuar luego de consumado el hecho a través de la única vía de reclamo que provee la plataforma para los casos en los que la desindexación es motivada por la supuesta infracción a la DMCA, la “contra-notificación”. Esta con-

tra-notificación funciona como una forma de apelación a la que los usuarios pueden recurrir directamente a través de la plataforma.

Sin embargo, tomar la decisión de avanzar con una contra-notificación implica un riesgo y Google lo expresa en tono de advertencia: “Tenga en cuenta que se le pueden exigir responsabilidades por perjuicios (que pueden incluir el pago de costas procesales y honorarios de abogados) si denuncia sin fundamento que el contenido se ha retirado o inhabilitado por error o porque se identificó incorrectamente. Por lo tanto, si no sabe con seguridad si un determinado material infringe los derechos de autor de otros, le sugerimos que busque asesoramiento legal antes de presentar la contra-notificación”. Al final del formulario, la persona debe aceptar otra imposición -que no tiene la opción de rechazar para poder seguir apelando-: “Acepto someterme a la jurisdicción del tribunal federal del distrito judicial en el que resido (o el distrito norte de California si mi residencia se encuentra fuera de Estados Unidos)”.

Aún en ese contexto, Sudestada decidió apelar ante la plataforma por la desindexación de su investigación de Google, ante lo que la empresa respondió: “Recibimos muchas solicitudes de este tipo cada día; su mensaje está en nuestra cola y lo recibiremos tan pronto como lo permita nuestra carga de trabajo. Debido al gran volumen de solicitudes que recibimos, solo podremos brindarle una respuesta si determinamos que su solicitud puede ser una queja legal válida y procesable, y podemos responder con preguntas o solicitudes de aclaración”.

A los pocos días, Google envió dos nuevas notificaciones a Sudestada acerca de su decisión de remover otra nota sobre el caso Lava Jato de sus resultados de búsqueda: “Se afianza el vínculo de estudio uruguayo en la trama de corrupción española”. En esta oportunidad el motivo no fue por la supuesta violación de

un derecho de autor aplicando la DMCA, sino que la medida se basaba en el denominado “derecho al olvido” contemplado en el Reglamento General de Protección de Datos de la Unión Europea.

Esta era la única información proporcionada por Google para justificar la desindexación del segundo contenido. No se informó quién era la persona denunciante, ni qué parte del artículo violaba los datos personales. Tampoco hubo posibilidad de contestar la denuncia antes de que se desindexara el contenido.

Google Search Console

Aviso de eliminación de la Búsqueda de Google en virtud de la ley europea sobre protección de datos

Para: Webmaster de <https://www.sudestada.com.uy/>:

Debido a una solicitud en virtud de la ley europea sobre protección de datos, Google ya no puede mostrar una o más páginas de tu sitio en los resultados de la Búsqueda de Google. Esto solo afecta las respuestas de algunas búsquedas de nombres o de otros identificadores personales que posiblemente aparezcan en las páginas. Solo se ven afectados los resultados de las versiones de resultados de búsqueda de Google correspondientes a los países que aplican la ley europea sobre protección de datos. No es necesario que realices ninguna acción.

Información importante:

- 1 No se bloquearon por completo esas páginas en nuestros resultados de la búsqueda. Solo se bloquearon en ciertas búsquedas de nombres en versiones de resultados de la búsqueda de Google correspondientes a países que aplican la ley europea sobre protección de datos. Por lo tanto, seguirán apareciendo en otras búsquedas.
- 2 No revelamos qué búsquedas se vieron afectadas.
- 3 En muchos casos, las búsquedas afectadas no se relacionan con el nombre de ninguna persona mencionada visiblemente en la página.

Ante el segundo hecho, Sudestada lo identificó como una estrategia de silenciamiento y no como acontecimiento aislado, por lo que decidió presentar una acción de amparo en la Justicia de Uruguay “ante la vulneración y ataque contra los derechos a la libertad de expresión y a ser oído, a una decisión fundada y al debido proceso”. La acción de amparo tenía como objetivo restituir “de manera inmediata los contenidos desindexados y que se ordene la adopción de medidas de no repetición de violación de derechos humanos y garantías procesales aquí afectados, respecto de futuras denuncias contra el medio y el periodista”.

Ante el recurso de amparo, los tribunales uruguayos resolvieron que no podían actuar, al no tener competencia para decidir sobre

la aplicación del derecho al olvido europeo por parte de Google. Por otro lado, el tribunal argumentó que no se afecta la libertad de expresión o de prensa ya que “el contenido no fue eliminado de internet a nivel mundial, sino desindexado por un motor de búsqueda en una región”.

Esto difiere de lo planteado por el reporte “Estándares para una internet libre, abierta e inclusiva” de la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos (CIDH), donde se afirma que la desindexación hace que “la información sea más difícil de encontrar y la invisibiliza. Ambos (bloqueo y desindexación) tienen un efecto limitante del derecho a la libertad de expresión porque restringen la posibilidad de buscar, recibir y difundir información e ideas sin importar las fronteras nacionales”. Por último, la justicia uruguaya consideró que el demandante debió haber buscado una reparación en Europa, donde se había bloqueado la información.

El departamento de Libertad de Expresión Global de la Universidad de Columbia, Estados Unidos, analizó el caso y, en un informe que publicado al respecto, lo consideró “un precedente regresivo puesto que la sentencia desaconseja futuros reclamos contra plataformas de internet a pesar de que las personas se vean perjudicadas por leyes extranjeras”.

²⁶ Google domina el negocio de las búsquedas en internet en América del Sur con el 96.38%.

²⁷ <https://www.observacom.org/universidad-de-columbia-califico-la-sentencia-del-caso-sudestada-de-uruguay-como-un-precedente-regresivo/>

²⁸ http://www.oas.org/en/iachr/expression/docs/publications/internet_2016_eng.pdf

²⁹ <https://globalfreedomofexpression.columbia.edu/cases/sudestada-and-fabian-werner-v-google/>

CASO 3

Suspensión temporal y exigencia de eliminar una denuncia contra violencia policial a un medio comunitario de Ecuador. El caso Wambra de Ecuador



El 9 de marzo de 2022, la cuenta en Twitter de Wambra, medio digital comunitario ecuatoriano, recibió la notificación de que tenía limitado su derecho a publicar nuevo contenido. Twitter le notificó que su cuenta fue bloqueada temporalmente por violar sus normas comunitarias respecto a la publicación de información sensible de terceros sin consentimiento, aunque no aclaró a qué información refería, si la supuesta infracción se relacionaba con una imagen publicada o con el texto de los tuits.

La red social sí informaba que la infracción se refería a un tuit específico en el que se denunciaba que una de las periodistas de Wambra había sido agredida por la Policía en la marcha por el Día de la Mujer realizada el 8 de marzo en Quito (Ecuador). Para devolver el uso inmediato de la cuenta, Twitter exigía al medio comunitario que eliminaran el tuit en el que se denunciaba la agresión de la Policía a la periodista mientras realizaba la tarea de informar.

El tuit de Wambra decía: “URGENTE Nuestra compañera periodista Viviana Erazo @vivianaerz fue afectada por el gas pimienta y fue golpeada con toletes por la Policía mientras cubría la marcha de mujeres y feministas #8Marzo2022 en #Quito varias mujeres fueron golpeadas por la Policía”. Incluía además

un video en el que se podía ver el momento en que Erazo recibía gas pimienta por parte de la Policía y luego escuchar (porque apenas puede verse) los momentos en los que sus compañeras la ayudaban.

Tras el cierre de la cuenta de su lugar de trabajo, Erazo dijo al portal ecuatoriano Fundamedios que la marcha no fue violenta y que los manifestantes “no tenían armas”. Opinó que las acciones de la Policía fueron “desproporcionadas” pues utilizaron garrotes y gas, y agredieron a mujeres con el riesgo de exponer a niños que acompañaban la marcha. La comunicadora dijo que no vio muchas colegas de prensa en el momento de la agresión, pero que sí vio a hombres con cámaras y contó que a ellos no les hicieron nada.

“Yo tenía el micrófono, y estaba desde ahí grabando y se acercó un policía, me dio un toletazo..., todo lo que estaba dentro de mi maleta estaba roto...; al rato en que me regreso, para irme después del ataque, me lanza gas pimienta directamente en los ojos”.

Como resultado de la sanción, Wambra emitió un comunicado en el que informaba que no veía otra alternativa que eliminar el tuit denunciado -lo que finalmente hizo-, para recuperar el uso de la cuenta:

“Esta agresión la denunciarnos a través de nuestra cuenta Twitter @wambraEc donde difundimos imágenes y videos de la represión. Sin embargo, la publicación donde informamos sobre la agresión a Viviana fue denunciada y nuestra cuenta bloqueada, a partir de las 23H00 del 08 de marzo. Este bloqueo implica una restricción al acceso a la cuenta de Wambra, la suspensión del tuit y la restricción a generar nuevos contenidos, aunque la cuenta siga activa. Esto nos preocupa de sobremanera ya que puede convertirse en una forma de accionar de personas que buscan censurar nuestros contenidos como medio de comunicación. Frente a esto, nos hemos visto obligados a eliminar el tuit para retomar el acceso a nuestra cuenta y continuar nuestra labor de comunicación.”

Somos un Medio Comunitario que trabaja un periodismo con enfoque de derechos humanos, feminista e interseccional que, desde nuestro nacimiento en 2010, realizamos la cobertura de la movilización social y cubrimos la agenda por los derechos de las mujeres, las diversidades sexo genéricas, los pueblos y nacionalidades, los trabajadorxs, estudiantes y las comunidades diversas. Esto lo hacemos con coberturas informativas constantes, cercanas, sensibles con las demandas sociales y rigurosas en el manejo de la información. Desde 2016, con la primera marcha Vivas Nos Queremos hemos cubierto las acciones feministas y de mujeres contra la violencia de género y el feminicidio y hemos podido visibilizar estas demandas. Este



³⁰ https://help-twitter-com.translate.google/en/rules-and-policies/personal-information?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sc

³¹ Garrotes

³² <https://wambra.ec/8m-mujeres-piden-derechos-reciben-violencia-policial-ecuador/>

³³ <https://www.fundamedios.org.ec/alertas/en-el-de-la-mujer-la-fuerza-publica-agrede-con-toletazo-y-gas-pimienta-a-una-periodista/>

³⁴ <https://wambra.ec/rechazamos-agresion-censura-mujeres-8-marzo/>

8 de marzo de 2022 es la primera vez que, en nuestras coberturas, la Policía Nacional responde con violencia una manifestación de mujeres. Algo que no ocurrió solo en Quito, ya que mujeres también fueron violentadas y amedrentadas por la Policía en Cuenca y Guayaquil.

Rechazamos la actuación desproporcionada de la Policía contra las mujeres activistas feministas y contra las mujeres periodistas que cubrían la movilización y nos solidarizamos con cada una de ellas.

Exigimos al Estado que garantice la labor de las comunicadoras y periodistas para informar en el marco del ejercicio del Derecho a la Libertad de Expresión y la seguridad de los y las trabajadoras de la comunicación y que incluya el enfoque de género en estas decisiones. De la misma manera, pedimos al gobierno nacional que rechace el hostigamiento y la censura de partidarios del gobierno contra medios de comunicación y periodistas en redes sociales, algo que puede convertirse en una forma de censura de la información y limitación de la Libertad de Expresión”.

En ningún momento queda clara cuál es la información personal no autorizada que fue divulgada y que, supuestamente, contravino las reglas de Twitter, algo que la plataforma no lo aclara en la información que envía al medio.

Las reglas a las que refiere el aviso de Twitter se relacionan con la publicación privada de terceros y tienen por objetivo impedir una práctica conocida como doxing (acosar a una persona a partir de difundir datos de su vida como su dirección, fotos familiares, lugares donde circula, etc.). Las normas sobre este aspecto fueron actualizadas por Twitter en abril de 2022 y establecen que “no (se) puede postear ni publicar información privada de otras personas sin su autorización y permiso expresos” así como “amenazar con exponer información privada o incentivar a otros a hacerlo”. “No puede compartir medios privados, como imágenes o videos de personas privadas, sin su consentimiento. Sin embargo, reconocemos que hay casos en los que los usuarios pueden compartir imágenes o videos de personas privadas, que no son figuras públicas, como parte de un evento de interés periodístico o para promover el discurso público sobre temas o eventos de interés público”, se añade.

Twitter define específicamente qué información no puede hacerse pública sin el permiso de terceros: domicilio o ubicación física (direcciones de calles, coordenadas de GPS u otra información de identificación relacionada con ubicaciones que se consideran privadas), documentos de identidad, incluidas las identificaciones emitidas por el gobierno u otros números de identidad nacional (aunque Twitter asegura que hace excepciones en zonas en las que esta información no se considera privada), teléfonos personales o direcciones de correo electrónico, números de cuenta o detalles de tarjetas de crédito, datos biométricos o registros médicos, registros fotográficos o de video de personas particulares sin su permiso, registros del mismo tipo de prisioneros de guerra, incluso los publicados por medios de gobiernos o estatales.

Como se ve, no es clara qué información de difusión prohibida fue hecha pública en el tuit por el que se recibe la sanción ni siquiera a qué personas pertenece esa información, dato importante si se tiene en cuenta que las únicas personas que aparecen en el video, además de las propias periodistas, son -y apenas visibles- agentes policiales. Es al menos discutible que la imagen de funcionarios públicos reprimiendo una marcha pueda ser considerada divulgación de información privada.

Asimismo, tampoco se entiende por qué la cobertura de un medio de comunicación (comunitario o no) sobre un hecho de notorio interés público -como lo es la marcha por el Día Internacional de la Mujer-, no es considerado como una excepción prevista en la reglamentación de Twitter que refiere específicamente al periodismo y el interés público.

Nota aparte es el hecho de que la única posibilidad de apelar la decisión que se ofrece en el anuncio de suspensión de publicación es comunicarse con un “equipo de soporte” que maneja todo tipo de reclamos técnicos y no se ofrece ningún camino de apelación específico. Para la recuperación de la cuenta se exige como única alternativa eliminar un contenido. Esto implica una presión sobre el medio de comunicación involucrado, máxime cuando se trata de una denuncia por violencia policial, así como una censura del derecho del medio a informar a sus seguidores de lo que está ocurriendo y del derecho de éstos a saberlo.

Es curioso que meses después Ecuador fuera escenario de un caso similar. El 22 de agosto de 2022, la Alianza por los Derechos Humanos de Ecuador emitió un comunicado denunciando que, luego que Twitter les cerrara su cuenta y que pasaran 20 días sin recibir respuesta, habían tenido que eliminar un tuit como única alternativa para recuperarla. El motivo alegado también fue la supuesta difusión de información privada, en este caso

una imagen del activista Henry Quezada durante el Paro Nacional realizado el 30 de julio de 2022. Aunque entendió que “la denuncia y la suspensión de Twitter obstaculizan la labor de la Alianza y es otra forma de persecución a la defensa de derechos”, sus representantes no vieron más alternativa que eliminar la publicación para poder retomar, luego de 20 días de espera, ese canal de difusión.

Entendemos este caso como de gran relevancia para el análisis del impacto de la moderación de contenidos en plataformas sobre los medios y el periodismo por varios aspectos antes mencionados, pero en particular porque la exigencia directa de eliminar un contenido como única alternativa para recuperar el uso de una cuenta es abusiva, en tanto su incumplimiento conlleva el cierre del principal o uno de sus principales canales de difusión con la sociedad. Todo esto, sin dar explicaciones claras sobre los motivos ni tener la posibilidad cierta de apelación para revertir la decisión, configura un acto de censura directa que, a juzgar por lo que ocurrió meses después con la Alianza para los Derechos Humanos, lejos está de ser un caso aislado.

³⁵ https://help-twitter-com.translate.goog/en/rules-and-policies/personal-information?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sc

³⁶ “Sin embargo, reconocemos que hay casos en los que los usuarios pueden compartir imágenes o videos de personas privadas, que no son figuras públicas, como parte de un evento de interés periodístico o para promover el discurso público sobre temas o eventos de interés público” https://help-twitter-com.translate.goog/en/rules-and-policies/personal-information?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sc

³⁷ <https://twitter.com/SurkunaEc/status/1554202734496301060>

³⁸ <https://www.fundamedios.org.ec/alertas/actualizacion-alianza-por-los-derechos-humanos-debio-borrar-un-tuit-para-tratar-de-recuperar-su-cuenta/>

CASO 4

Eliminación de miles de publicaciones de Instagram sobre el Paro Nacional de Colombia



En mayo de 2021, miles de usuarios -algunos con miles de seguidores y otros con muy pocos- denunciaron la desaparición de sus publicaciones o historias de Instagram y Facebook (ambas, propiedad de la empresa Meta) que informaban sobre las movilizaciones y protestas ocurridas en distintas ciudades de Colombia durante el Paro Nacional.

Las protestas habían comenzado el 28 de abril para manifestar la oposición contra un proyecto de reforma fiscal presentado por el gobierno del entonces presidente colombiano Iván Duque. Como consecuencia de las protestas, murieron más de 35 personas, Duque dio marcha atrás con el proyecto y el ministro de Hacienda del momento, Alberto Carrasquilla, presentó su renuncia.

La actriz colombiana Carolina Guerra, que vive en Estados Unidos, escribió en Twitter: “Mis historias en @instagram de las últimas 24 hrs tienen un promedio de 35k vistas excepto unas en donde me dirijo a la comunidad colombiana en City Hall de la ciudad de LA en una velatón. Esas tienen sólo 1k. Las han restringido y después borrado. ¿CENSURA? Si”. Horas después, compartió el video de su discurso en la manifestación de Los Ángeles. “Este vídeo me llegó a mi cuenta ayer, lo compartí en mis historias y esta mañana @ Instagram lo bajó. ¿Por qué, si no viola ninguno de los lineamientos establecidos?”, se preguntó.

La desaparición de publicaciones afectó también a medios y periodistas dentro y fuera de Colombia que informaban en sus países sobre lo que estaba ocurriendo, muchas veces utilizando el hashtag #SOSColombia. La Fundación para la Libertad de Prensa (FLIP) de

Colombia ⁴¹ denunció que había recibido denuncias individuales y colectivas de “reporteros, fotógrafos, artistas, colectivos feministas y medios independientes”.

La FLIP y el resto de las organizaciones integrantes del índice de Derechos Digitales emitieron un comunicado en el que aseguraron que “las acciones de “apagar” o “interrumpir” Internet son una forma de censura previa prohibida por la Constitución y por la Convención Americana sobre Derechos Humanos, y el Estado tiene el deber de investigar y clarificar las causas de los hechos denunciados en los últimos días.

Tras los reclamos, Instagram -que se había mantenido en silencio y tampoco había informado a los usuarios de forma individual sobre lo que ocurría-, emitió un breve comunicado desde su cuenta de Twitter en el que aseguraba que “algunas personas tienen problemas para cargar y ver historias” y que se debía a “un problema técnico global generalizado que no está relacionado con ningún tema en particular”, sin mencionar el episodio específico en Colombia. Horas después, Instagram actualizó la información asegurando que había “solucionado este problema” que “afectó a muchas historias que contienen publicaciones compartidas creadas ayer y esta mañana [miércoles y jueves]”, además de las opciones para publicar imágenes de archivo.

Pero muchos usuarios y medios de comunicación recordaron que algo muy similar había ocurrido con el hashtag #BlackLivesMatter durante las protestas por el asesinato de George Floyd a manos de policías norteamericanos. Lo que había ocurrido en ese caso, al menos lo que la plataforma hizo público



luego, fue una confusión en el etiquetado del contenido de ese hashtag como spam debido a la velocidad con la que nuevo contenido aparecía, lo que hizo que los sistemas de inteligencia artificial lo confundieran con contenido automatizado.

En setiembre de 2020, Facebook censuró los posts que incluían el hashtag #ACAB (cuya sigla significa “All Cops are Bastards”) en, al menos, Colombia, Brasil, Chile, México, Costa Rica y Perú. Meta fue consultado al respecto y sostuvo que “pasaba desde agosto y que era mundial, ya que encontraron que #ACAB estaba relacionado con contenido de baja calidad”, pero semanas después y tras varios reclamos por eliminación de contenido de distintas organizaciones, Facebook dio marcha atrás a la medida y dejó de eliminar los posts con ese hashtag.

Respecto al caso del Paro Nacional en Colombia, en una declaración de un portavoz de la empresa al diario colombiano El Tiempo, los motivos parecen tener que ver con la dificultad para manejar un gran volumen de contenido que las normas comunitarias de la plataforma podrían considerar sensible y por ende sujeto a un flaggeeo que lleva más tiempo de procesar que la velocidad en la que se estaban publicando contenidos. “Sabemos que las personas acuden a Meta para debatir

los temas que más les interesan y tenemos normas para mantener un ambiente seguro en nuestras plataformas, que no permite publicar imágenes violentas. En el marco de las manifestaciones en Colombia, estamos trabajando para habilitar publicaciones que sean de interés público, advirtiendo a los usuarios que puede tratarse de imágenes sensibles”, dijo el portavoz.

Este caso resulta relevante para comprender el impacto que tiene la moderación de contenido a escala global, en tiempo real y de forma masiva para la difusión de hechos de interés público que están ocurriendo y por ende para el derecho a informar y a estar informado. Las plataformas pueden, de quererlo, generar prácticamente un apagón informativo sobre un tema determinado y de esa forma acallar a medios y periodistas, así como mantener a las personas alejadas de información relevante para sus vidas.

Es importante además tener en cuenta que para el periodismo el momento presente es parte central del contexto informativo, por lo que el hecho de que las publicaciones desaparezcan de forma temporal y luego sean reinstaladas horas después, altera de forma significativa el impacto de la información que contienen y, eventualmente, las consecuencias para el público de recibir esa información.

³⁹ <https://twitter.com/carolinaguerram/status/1390483879883018243>

⁴⁰ <https://ladiaria.com.uy/politica/articulo/2021/5/denuncian-censura-de-instagram-a-publicaciones-sobre-represion-en-colombia-la-red-social-dice-que-fue-un-problema-tecnico-global/>

⁴¹ https://www.instagram.com/p/COiayXhse3K/?utm_source=ig_web_copy_link

⁴² https://twitter.com/FLIP_org/status/1390336091803332608?s=20&t=A9VkhW3Tq4Lh6334Wj9cQ

⁴³ <https://twitter.com/InstagramComms>

⁴⁴ <https://twitter.com/InstagramComms/status/1267541207015661568>

⁴⁵ <https://cerosetenta.uniandes.edu.co/la-censura-de-acab-o-el-poder-de-facebook-sobre-la-libertad-de-expresion/>

⁴⁶ <https://www.eltiempo.com/tecnosfera/apps/instagram-esta-borrando-las-stories-sobre-el-paro-nacional-586471>



05.

Palabras finales

05. Palabras finales

Con la llegada de Internet y las redes sociales se democratizó de manera muy significativa la producción y creación de contenido. Miles de personas pudieron comenzar a manifestar su opinión e informar a otros sobre distintos temas. El periodismo ganó en recursos para generar contenido y en capacidad para trascender las fronteras. Sin embargo, su distribución y visibilización quedó en manos de un oligopolio de redes sociales y buscadores que intermedian diariamente la relación entre quienes generan contenidos y sus audiencias.

En ese marco, la moderación -sus procesos, sistemas y protagonistas- se ha transformado en un punto clave en la construcción de espacios de debate público. Quiénes moderan, cómo lo hacen, sobre qué bases, cuántas veces se equivocan, en qué lugares, son preguntas que aún hoy continúan sin respuesta. De fondo se plantea la interrogante de si son los ejecutivos de empresas multinacionales, sin ninguna clase de controles o de transparencia acerca de sus procesos, quienes deben decidir sobre qué hablamos, cómo lo hacemos y qué reglas se aplican para callar a algunos y aumentar el alcance de otros.

El impacto de estas opacidades y discrecionalidades en el ejercicio concreto del periodismo, así como en la libertad de expresión en su doble dimensión de derecho a difundir y derecho a recibir información, es una preocupación creciente entre investigadores, expertos y también de periodistas y sus organizaciones representativas. Es imposible pensar el periodismo sin sus audiencias y estas decisiones privadas tienen la potencialidad de crear espacios en los que los contenidos existan, pero nadie pueda verlos.



OBSERVACOM

Gral. César Díaz 1239 AP. 101 - CP 11800 -
Montevideo, Uruguay
www.observacom.org