

Contenido bloqueado

Las plataformas filtran los contenidos bajo criterios que no se condicen con la realidad de los territorios

Informe 2020

LIBERTAD DE EXPRESIÓN EN CHILE

Capítulo 4.

Censuras en Internet y redes sociales en Chile: Ataques, discursos de odio y algoritmos









CAPÍTULO 4. CENSURAS EN INTERNET Y REDES SOCIALES EN CHILE: ATAQUES, DISCURSOS DE ODIO Y ALGORITMOS

Por Patricia Peña, académica del Instituto de la Comunicación e Imagen de la Universidad de Chile y directora de Fundación Datos Protegidos.

- 1. LIBERTAD DE EXPRESIÓN EN AMÉRICA
 LATINA Y LAS POLÍTICAS DE MODERACIÓN
 DE CONTENIDOS
- II. LOS CAMBIOS EN LA MODERACIÓN DE CONTENIDOS DURANTE LA PANDEMIA Y EL IMPACTO DE LA CRISIS POLÍTICA DE ESTADOS UNIDOS
- III. CASOS EN CHILE

16 REFERENCIAS



CAPÍTULO 4. CENSURAS EN INTERNET Y REDES SOCIALES EN CHILE: ATAQUES, DISCURSOS DE ODIO Y ALGORITMOS

El Informe "Libertad de Expresión en el contexto de las protestas y movilizaciones sociales en Chile, durante el estallido social", preparado por el Observatorio del Derecho a la Comunicación, Fundación Datos Protegidos y el Instituto de la Comunicación e Imagen de la Universidad de Chile (2020) (1), fue pionero en establecer evidencia de lo que hoy ya se denomina "censura algorítmica" en debates de organizaciones de la sociedad civil latinoamericana y que se relaciona con una nueva dimensión en relación a la comprensión de los alcances del derecho a la comunicación y ejercicio de la libertad de expresión en Internet, en cuanto se refiere a los mecanismos y sistemas de moderación automatizados que implican filtración e invisibilización de contenidos y suspensión o bloqueo de cuentas en plataformas de redes sociales (Facebook, Twitter, Instagram y Youtube).

En ese informe, se dieron a conocer más de 200 situaciones o casos que fueron señalados como de censura algorítmica, lo que implica la aplicación de sistemas de filtrado automático a través de la programación de algoritmos, basados en inteligencia artificial y que están siendo constantemente entrenados para realizar la tarea de filtrar o evitar la publicación automática de contenidos de manera previa, de acuerdo a lo que indican las políticas o normas de uso, denominadas también como "normas comunitarias", definidas por las empresas dueñas de las plataformas.

En el presente informe, revisaremos casos ocurridos en el período mayo 2020 a febrero 2021 que nos permiten identificar distintas formas de limitar el ejercicio de la libertad de expresión en Internet, las que tienen consecuencias sobre el debate público, de particular relevancia en el contexto de situaciones de crisis sociales y políticas como la que Chile enfrenta desde el 18 de octubre de 2019 (o 18-O).

Estas situaciones incluyen censura de los algoritmos de las plataformas en cuentas de medios de comunicación y colectivos de información independientes, ataques a sitios web de medios digitales críticos al gobierno y situaciones de violencia y discursos de odio contra cuentas de redes sociales de mujeres feministas.

I. LIBERTAD DE EXPRESIÓN EN AMÉRICA LATINA Y LAS POLÍTICAS DE MODERACIÓN DE CONTENIDOS

En los primeros años de funcionamiento y crecimiento de las principales redes sociales (Facebook, Youtube y Twitter) para países latinoamericanos, la principal característica era potenciar estos espacios como de libre conversación y de creación de comunidades.

Las empresas dueñas adoptaron enfoques de mínima intervención en los contenidos y promovieron el uso de estos espacios para un libre intercambio de ideas, pensamientos y conexión de personas, tal como lo declaraban las misiones originales de Facebook o Twitter (en sus inicios en 2005) (2), pero desde el comienzo también establecieron un tipo de modelo de auto-regulación o "políticas de moderación de contenidos" (3) que establecían normas o reglas de uso de las plataformas en relación a los intercambios y reacciones que se generaban entre usuarias y usuarios respecto de conversaciones o circulación de informaciones que fueran dañinas o peligrosas para sus comunidades; y que también fueran perjudiciales a sus modelos de negocio, basados en publicidad y posicionamiento de marcas o contenidos (Cobbe, 2020) (4).

Es importante señalar que las plataformas de redes sociales, así como otras empresas de Internet son consideradas como "servicios o intermediarios de Internet".

En la primera década de los 2000 hubo un intenso debate en relación a la responsabilidad y la regulación de estas empresas en la medida que facilitan el acceso, la circulación y distribución de información y opiniones, así como también noticias y cualquier otro tipo de contenidos. En este contexto, la Organización para la Cooperación y el Desarrollo Económicos (OCDE) define que "los intermediarios de Internet facilitan transacciones entre terceros en Internet, dando acceso a, alojando, transmitiendo e indexando contenidos, productos y servicios originados por terceros en Internet o proveyendo a terceros de servicios basados en Internet" (5).

Las plataformas de redes sociales se caracterizan por su promesa de ser espacios en que los contenidos son generados y creados por usuarios en sus cuentas, sean personas individuales o perfiles de empresas, instituciones o gobiernos.

En este contexto, la controversia en torno a la regulación de estos espacios tiene que ver con la característica central de su modelo de funcionamiento: dejar que los contenidos, informaciones y mensajes se puedan publicar y visibilizar sin controles editoriales previos (o ex ante) como ocurría en el modelo de los medios de comunicación tradicional.

Estas empresas han contado tradicionalmente con equipos de moderadores, conformados por personas que revisan los contenidos que son marcados por sistemas automatizados de algoritmos como "sospechosos" o que "violan" algunas de las normas o reglas de la comunidad (6). Una cuestión fundamental en estos mecanismos ha sido que los moderadores para América Latina no viven en la región, ni tienen necesariamente relación con el contexto político, sociocultural o con lo propio del lenguaje e idiosincrasia comunicacional (las jergas, por ejemplo), lo que hace más complejo el entendimiento en relación a estas acciones de parte de las plataformas.

Estos sistemas de moderación humana, comienzan progresivamente a automatizarse a través de algoritmos para poder revisar y filtrar los grandes y masivos volúmenes de contenidos que se generan y que circulan en las redes sociales, lo que implica un sistema automatizado, entrenado con tecnologías de inteligencia artificial y que apunta hacia un cambio de formas ex ante de moderación que identifica y suprime el contenido clasificado como "prohibido" o "que viola algunas de reglas" de la comunidad o de la plataforma.

Esto es lo que finalmente comienza a ser debatido y denunciado como "censura algorítmica" (de las plataformas de redes sociales) tanto por estudios académicos como por organizaciones de la sociedad civil en todo el mundo (Cobbe, 2020) (7).

En ambas situaciones, debemos sumar que en redes sociales también circulan y actúan cuentas bot (8) que han estado detrás de ataques masivos a cuentas específicas como figuras políticas, públicas, activistas de derechos humanos, entre otras.

Las políticas y normas de moderación de contenidos de las plataformas de redes sociales tienen que ver con las siguientes disposiciones en la actualidad:

TABLA SÍNTESIS. NORMAS COMUNITARIAS DE MODERACIÓN DE CONTENIDOS Y POLÍTICAS DE USO DE PLATAFORMAS DE REDES SOCIALES

TWITTER (9)	YOUTUBE (10)	FACEBOOK (11)	INSTAGRAM (12)
> SEGURIDAD	> DESNUDOS Y CONTENIDO	>VIOLENCIA Y COMPORTAMIENTO	>VIOLENCIA Y
Violencia; terrorismo;	SEXUAL	CRIMINAL	COMPORTAMIENTO CRIMINAL
extremismo violento;	> CONTENIDO DAÑINO Y	Personas y organizaciones	Terrorismo, crimen organizado,
explotación sexual infantil;	PELIGROSO	peligrosas; Organización de	grupos que promueven el odio.
abuso; acoso; incitación al	> CONTENIDO VIOLENTO O	actividades nocivas y publicidad de	> SEGURIDAD
odio, suicidio -	GRÁFICO	la delincuencia, fraudes y estafas.	Suicidio o animación a las
autolesiones; Contenido	> CONTENIDO DE INCITACIÓN	> SEGURIDAD	autolesiones;
multimedia de carácter	AL ODIO	Autolesiones y suicidios;	No se permiten amenazas
delicado; incluida la	> ACOSO Y CIBERACOSO	explotación y abuso sexual de	graves; Imágenes con derecho
violencia gráfica y el	>SPAM, METADATOS	adultos y menores; bullying y	de autor; Compartir imágenes
contenido para adultos.	ENGAÑOSOS Y ESTAFAS	acoso; vulneraciones de privacidad	íntimas de otros usuarios/as.
>PRIVACIDAD	> AMENAZAS	y derechos de privacidad de	>CONTENIDO CUESTIONABLE
Información privada;	> DERECHOS DE AUTOR	imágenes.	Desnudos; Contenido que
desnudez no consensuada	> PRIVACIDAD	>CONTENIDO CUESTIONABLE	incluya amenazas creíbles o
>AUTENTICIDAD	> SUPLANTACIÓN DE	Incitación al discurso de odio;	lenguaje que incite al odio,
Spam y manipulación de la	IDENTIDAD	contenido gráfico y violento;	contenido dirigido a
plataforma; Integridad	> SEGURIDAD DE LOS	desnudos y actividad sexual de	particulares con el fin de
cívica; Suplantación de	MENORES	adultos; contenido cruel e	humillarlos o avergonzarlos;
identidad; Contenidos	> POLÍTICAS ADICIONALES	insensible.	Contenido sexual de menores.
multimedia falsos y		> AUTENTICIDAD E INTEGRIDAD	> AUTENTICIDAD E
alterados.		Circulación de spams; noticias	INTEGRIDAD
		falsas; contenido multimedia	Spam; Opiniones, calificaciones
		manipulado; Eliminación de	de usuarios falsas y engañosas;
		cuentas que "perjudican a la	Suplantación de identidad.
		comunidad"	

Fuente: elaboración propia en base a información recogida de las respectivas plataformas.

Esta tabla es un resumen que incluye los últimos cambios o modificaciones en las categorías de las normas comunitarias o políticas de uso realizadas por las 4 plataformas consideradas en el informe, a enero 2021. Para más detalles en profundidad de cada política, norma o categoría hay que ir al detalle del vínculo señalado.

- (*) Parte de las políticas de privacidad de Instagram y Facebook poseen disposiciones idénticas en varias categorías por pertenecer al mismo conglomerado de empresas.
- (*) La redacción de la política de privacidad de Instagram no contiene enunciados que clasifiquen las normas, por eso utilizamos la estructura de Facebook para proponer una clasificación.

Cabe destacar que los principios orientadores respecto a la garantía de la libertad de expresión en Internet han sido abordados por las Relatorías de Libertad de Expresión (RELE) de la Comisión Interamericana de Derechos Humanos (CIDH) como de Naciones Unidas (ONU), que han fijado parámetros en torno a las temáticas: acceso e infraestructura; pluralismo; no discriminación; respeto a la privacidad.

Según señala la Declaración conjunta de Relatores sobre Libertad de Expresión en Internet de 2011:

"La libertad de expresión se aplica a Internet del mismo modo que a todos los medios de comunicación. Las restricciones a la libertad de expresión en Internet solo resultan aceptables cuando cumplen con los estándares internacionales que disponen, entre otras cosas, que deberán estar previstas por la ley y perseguir una finalidad legítima reconocida por el derecho internacional y ser necesarias para alcanzar dicha finalidad" (13).

La RELE señaló el 2013 que "los intermediarios deben evitar que sus actividades provoquen o contribuyan a provocar consecuencias negativas sobre el derecho a la libertad de expresión" en sus acciones voluntarias de moderación de contenidos, las que "solamente deben considerarse legítimas cuando dichas restricciones no impidan u obstaculicen de manera arbitraria las posibilidades de expresión de una persona en Internet" (14).

Recientemente, la declaración conjunta de las cuatro Relatorías de la Libertad de Expresión que existen regionalmente, en una declaración sobre Internet han señalado que el "control privado" (que realizan estas empresas), es uno de los tres principales desafíos para la próxima década y una "amenaza" a la libertad de expresión ya que "una característica transformativa del entorno de las comunicaciones digitales es el poder de las empresas privadas y, particularmente, de las redes sociales, las plataformas de búsqueda y otros intermediarios, sobre las comunicaciones, con un enorme poder concentrado en unas pocas empresas" (15).

Particularmente en el caso de la censura algorítmica se critica la falta de transparencia en relación a los softwares con que funcionan estos mecanismos y la aplicación de filtros previos o posteriores a contenidos audiovisuales, gráficos o escritos; cuentas o perfiles de distintas categorías y cantidad de seguidores; hashtags, entre otros.

Las situaciones de bloqueos, moderación de contenidos o censura previa que realizan las plataformas impacta directamente en los registros y evidencias digitales sobre violaciones a los derechos humanos o crisis sociales y políticas que son publicados en las plataformas de redes sociales (16).

En octubre de 2020, el Observatorio Latinoamericano de Regulación de Medios y Convergencia (OBSERVACOM) junto a Intervozes e IDEC (Brasil), CAInfo (Uruguay), TEDIC (Paraguay), PROLEDI (Costa Rica), Usuarios Digitales (Ecuador), Desarrollo Digital (Argentina), IPANDETEC (Panamá) y Fundación Datos Protegidos (Chile), todas organizaciones de América Latina, desarrollaron la propuesta "Estándares para una regulación democrática de las grandes plataformas que garantice la libertad de expresión en línea y una Internet libre y abierta" (2020) (16).

Esta propuesta busca avanzar en el debate y desarrollo de recomendaciones que permitan a los países y sociedades enfocadas en mejores prácticas de co-regulación y regulación pública en aspectos como: transparencia; alcance y carácter de la regulación; términos y condiciones del servicio; entre otras categorías (17).

II. LOS CAMBIOS EN LA MODERACIÓN DE CONTENIDOS DURANTE LA PANDEMIA Y EL IMPACTO DE LA CRISIS POLÍTICA DE ESTADOS UNIDOS

Durante 2020, dos nuevos contextos han visibilizado el funcionamiento de los sistemas automatizados de moderación de contenidos en las plataformas: la pandemia de COVID-19 y los esfuerzos para frenar flujos de información falsa, engañosa o perjudicial; y por otro, las elecciones presidenciales en Estados Unidos, en particular los hechos ocurridos en el contexto del asalto al Capitolio en Washington, D.C ocurrido en enero de 2021 y la vocería del ex Presidente D. Trump en sus propias redes sociales, particularmente en Twitter.

Luego de la declaración de la pandemia en marzo de 2020 y los cambios en las dinámicas cotidianas que esto implicó, las plataformas Google, Facebook, Instagram y Twitter establecieron una serie de nuevas reglas en relación a la publicación de contenidos sobre la crisis sanitaria en su amplia variedad de formatos y narrativas en medida que fuera propagación de información errónea, dañina o peligrosa para las personas.

Dos ejemplos de la aplicación de estas medidas han sido:



Pantallazo tomado de la cuenta de twitter de @JairBolsonaro (29/03/2020)

- 1. La eliminación de tres videos, en tres tweets de la cuenta del Presidente de Brasil Jair Bolsonaro informando de actividades masivas en las que estaba participando en Brasilia, cuando la Organización Mundial de la Salud ya había establecido las recomendaciones de distanciamiento social para la prevención de contagios. La empresa Twitter estableció esta eliminación señalando que había incumplido las reglas en relación a publicaciones e informaciones compartidas en relación al contexto de la pandemia del COVID-19 (19).
- 2. La empresa YouTube anunció en un comunicado que a la fecha ha removido de su plataforma más de 30 mil videos que emitían afirmaciones falsas u ofrecían información engañosa acerca de las vacunas de COVID-19 durante los últimos seis meses, según los criterios sanitarios de la Organización Mundial de la Salud (OMS) (20).

La segunda situación se genera en el contexto del proceso eleccionario de Estados Unidos, el 6 de enero de 2021, con el asalto al Capitolio en Washington D.C., cuando el ex Presidente Trump desconoció el triunfo del Presidente electo Joe Biden.

En la medida que los hechos se sucedieron, Twitter fue la primera plataforma que decidió aplicar una eliminación de tweets de la cuenta del ex Presidente (@realDonalTrump), por incumplir sus reglas de publicación y luego establecer un bloqueo o suspensión permanente y en paralelo las plataformas de Facebook, Instagram, Youtube y Tik Tok, también implementaron medidas que durante menos de 48 horas no le permitieron acceder a publicar contenidos en ninguna de estas plataformas (21).

Este último hecho ha marcado el debate sobre libertad de expresión y regulaciones de Internet y contextualiza hacia las situaciones que se observan en Chile durante 2020 e inicios de 2021.

III. CASOS EN CHILE

Al igual que el contexto del estallido social de 2019 – 2020, la Fundación Datos Protegidos publicó un formulario de denuncia y registro de situaciones de censura en Internet y en redes sociales, particularmente en relación a situaciones o hechos vinculados a violaciones de derechos humanos en el país.

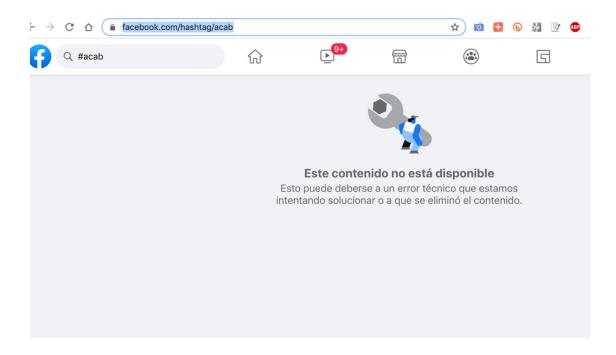
Durante 2020, este formulario (22) permaneció disponible. Sin embargo, el nivel de casos recibidos por esta vía fue muy bajo, debido a que las restricciones de movilidad con motivo del contexto sanitario frenaron el flujo y participación en protestas y movilizaciones sociales en espacios públicos.

Es para los propósitos de este informe que se expone una selección de casos que fueron compartidos directamente a través contactos vía redes sociales con la Fundación Datos Protegidos mediante Twitter (23) e Instagram (24).

En este contexto, se observan casos que grafican limitaciones al ejercicio de la libertad de expresión en Internet con consecuencias en el debate público, desde la aplicación de la censura algorítmica; bloqueos de cuentas; hostigamiento y ataques a medios independientes y críticos con el gobierno; ataques y discursos de odio hacia mujeres públicas, comunicadoras y comunidades.

> CENSURA ALGORÍTMICA Y BLOQUEO DE CUENTAS

- INVISIBILIZACIÓN DEL HASHTAG #ACAB EN FACEBOOK



Pantallazo de buscador en Facebook del hashtag #ACAB (septiembre 2020)

Entre agosto y septiembre de 2020, en países como Colombia, Chile, Argentina, Brasil al buscar contenidos o publicaciones etiquetados con las siglas o hashtag #ACAB (sigla inglesa para la frase *All Cops are Bastards (En castellano: "Todos los policías son bastardos")*, particularmente utilizada para denunciar la violencia o represión policial ejercida en estos países contra manifestantes, no aparecía visible ni tampoco era posible etiquetar porque infringía normas comunitarias.

Luego de varios reclamos y consultas de organizaciones latinoamericanas de la sociedad civil (25), la empresa señaló que el bloqueo era global y se debía a "contenidos de baja calidad".

El bloqueo para los países latinoamericanos se resolvió hacia fines de septiembre de 2020, sin mayores explicaciones o entendimiento acerca de los mecanismos que aplicó la plataforma.

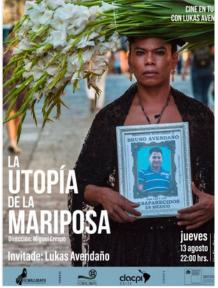
> OCULTAMIENTO DE CONTENIDOS, MEDIOS Y COLECTIVOS INDEPENDIENTES EN REDES SOCIALES

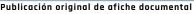
El informe de Libertad de Expresión en Chile en el contexto del estallido, que revisó el periodo entre el 18 de octubre y el 22 de noviembre de 2019, permitió identificar los casos de dos medios informativos independientes, como objeto de diversos tipos de censuras o bloqueo de acceso a cuentas y publicación de contenidos: Señal3 La VictoriaTV y PiensaPrensa. Se trata de proyectos de comunicación que han tenido un rol central en la información, tanto en tiempo real como amplificadores de informaciones y noticias sobre protestas sociales, movilizaciones sociales y situaciones de violencia y represión policial. Esta tendencia se mantuvo y se ha mantenido en distintos momentos y fechas durante el 2020, pero también sumaremos los casos de otros medios y colectivos informativos independientes.

- FICWALLMAPU

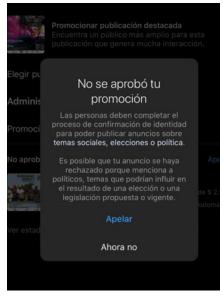
La cuenta de Instagram del Festival Internacional de Cine indígena de Wallmapu (26) y también productora audiovisual multicultural, tuvo una situación de censura previa al tratar de comprar publicidad para la publicación de un poster que anunciaba el lanzamiento de un documental online en su canal de YouTube. La promoción no fue aprobada por el sistema de Instagram, y se aplicó una prohibición de publicar ex ante.

Durante el 2020, hubo al menos 3 momentos en que sus cuentas registraron problemas de bloqueo de cuentas o censura de contenidos:





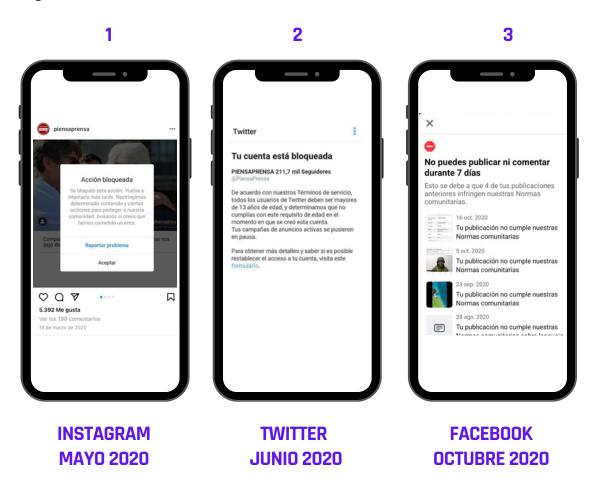




- PIENSA PRENSA

Al tratarse de un colectivo de prensa independiente que publicaba mayormente hasta 2019 sus noticias y contenidos en sus cuentas de redes sociales en Instagram (26), Twitter (27) y Facebook (28) estaba más expuesto a tener situaciones de censura de contenidos o publicaciones, porque su trabajo se focaliza en la cobertura de situaciones de protestas en Santiago, particularmente en la zona de Plaza Italia - rebautizada popularmente como Plaza de La Dignidad-, y las situaciones de violaciones de derechos humanos y represión policial a movilizaciones o manifestaciones sociales en el país.

Durante el 2020, hubo al menos 3 momentos en que sus cuentas registraron problemas de bloqueo de cuentas o censura de contenidos:



- 1. Invisibilización de contenido cuando aún se desarrollaban situaciones de protesta social en Santiago (previo a las medidas de cuarentena).
- 2. Bloqueo de cuenta.
- 3. Censura publicación relacionada a la presentación de un documental sobre el estallido social de octubre de 2019.

A la fecha, la cuenta original de Instagram @piensaprensa sigue con problemas de bloqueo. Con motivo de esto es que existen otras 3 cuentas de respaldo. Cabe destacar que este colectivo ya cuenta con un sitio web con identidad propia como medio de comunicación en www.piensaprensa.org .

> ATAQUES DIGITALES A MEDIOS Y EXPRESIONES CULTURALES CRÍTICOS CON EL GOBIERNO

- COLECTIVO PLIEGUE

Autodefinido como un colectivo estético - filosófico independiente, desarrollan una serie de actividades y acciones artísticas de divulgación y de creación de contenidos sobre estética, filosofía y saber contemporáneo. Su principal cuenta de red social es Instagram (30).

Se hicieron conocidos por presentar el 2019 un documental que cuestionaba críticamente a la academia en Chile y el 2020 preparaban el estreno de un documental sobre el estallido social o 18O (18 Octubre de 2019), en su canal de YouTube.

Días previos a esta presentación, recibieron un serie de amenazas en su principal cuenta de Instagram que hablaban de cerrar su canal en esta plataforma de videos online. Por ello decidieron suspender su estreno masivo, previo a la fecha de realización del plebiscito por una nueva constitución.



- CHILETODAY.CL

Desde el estallido social de 2019 Chile Today (31) se ha destacado por la difusión de contenidos en inglés sobre los hechos que estaban ocurriendo en Chile, lo que ha permitido un mayor conocimiento a nivel internacional de la gravedad de las violaciones de derechos humanos.

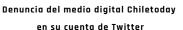
Partidarios del rechazo a la reforma constitucional y grupos de extrema derecha han sido especialmente hostiles con este tipo de medios, por sus críticas hacia Chile, lo que se ha traducido en campañas de acoso y hostigamiento a través de medios digitales. El 3 de octubre su sitio web fue objeto de un ataque de denegación de servicio (DDoS) que impidió que los contenidos del medio pudieran ser accesibles por parte de sus usuarios en las semanas anteriores al plebiscito por una nueva constitución. Esta situación fue denunciada por el mismo medio a través de su cuenta de Twitter (32), diez días después.

Además, el 18 de octubre de 2020 el corresponsal holandés del medio Chile Today News, Boris van der Spek, recibió, por correo electrónico, amenazas explícitas de muerte por parte de una persona no identificada, según informó la Asociación de Corresponsales de la prensa internacional en Chile en una declaración pública en su sitio web (33).

- LA RAZÓN

El 25 de octubre de 2020, día en el que se celebró el plebiscito sobre la reforma constitucional en Chile, el medio digital La Razón.cl sufrió un ataque informático a su sitio web de Noticias que impedía acceder a sus contenidos. Horas después del ataque el personal técnico del medio logró restablecer el servidor y el medio volvió a estar disponible a través de su sitio web (34).







Denuncia del medio digital La Razón sobre ataques a su servidor web

> DISCURSOS DE ODIO Y ATAQUES ONLINE A MUJERES EN EL CONTEXTO DE PANDEMIA

Como señalamos, este punto ha sido incorporado a modo de cierre de este informe para alertar de una situación que se ha incrementado en redes sociales: las situaciones de ataques, acoso y discurso de odio contra mujeres lideresas, periodistas, comunicadoras, activistas feministas y en particular de las comunidades LGTBQ+.

El contexto de la pandemia del COVID19 al decretarse extensos períodos de confinamiento y cuarentenas, aumentó las alertas en relación a las situaciones de violencia de género de tipo doméstico, pero también en y a través de plataformas digitales, en la medida que buena parte de las interacciones sociales están medidas por plataformas de redes sociales y de comunicación online.

- ATAQUES POR LA PLATAFORMA DE VIDEOCONFERENCIAS ZOOM

En el primer semestre de 2020, se hicieron frecuentes las denuncias de ataques denominados *zoombombing* (35) en espacios de reunión, conversatorios o de realización de clases online. Acá presentamos 3 casos que generaron atención pública, tanto en relación a los protocolos y medidas de seguridad que tienen estas plataformas, como del impacto en aspectos de libertad de expresión, seguridad y privacidad de estos espacios de comunicación.



16 de abril de 2020

Instituto de Filosofía de la Universidad Diego Portales, denuncia ataque de conversatorio vía zoom "Feminismo y revolución. Crónica de una inquietud" (36).



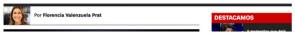
20 de abril de 2020

Instituto de Física de la Pontificia Universidad Católica de Valparaíso (PUCV) denunció ataque en defensa de tesis (37).

Colina: Irrumnieron en clase no

Colina: Irrumpieron en clase por Google Meet, compartieron pornografía e insultaron a profesora

Dos usuarios solicitaron acceso a un curso que se impartía por videollamada, usando el nombre de dos estudiantes del colegio. Expertos en ciberseguridad entregan recomendaciones para que casos como estos no vuelvan a ocurrir.



10 de agosto de 2020

Colegio San Antonio de Colina anunció medidas legales contra quienes resulten responsables (38).

- DENUNCIA DE LA ACTRIZ DANIELA VEGA POR VIDEOJUEGO QUE INSTABA A ASESINARLA

El **28 de octubre de 2020**, la actriz Daniela Vega denunció en su cuenta de twitter una querella en contra del responsable de la creación de un video juego que instaba a su asesinato, junto a otros personajes públicos. De igual manera, esta situación fue denunciada públicamente en la red social Twitter por la diputada de la Asamblea de Madrid Carla Antonelli.



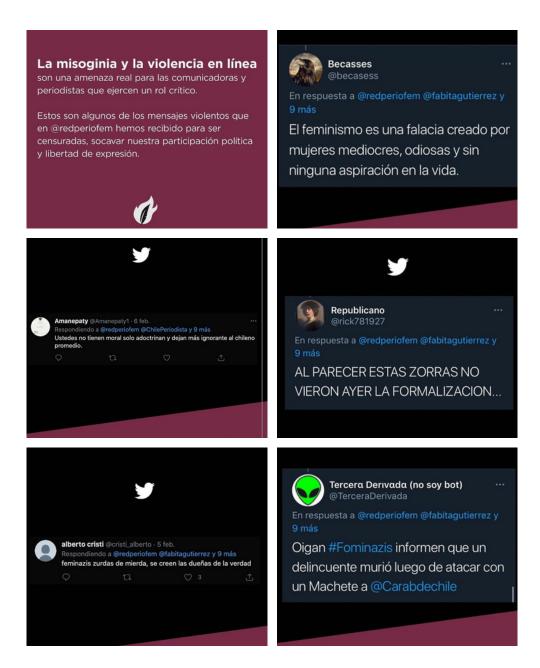
El videojuego estaba asociado a la cuenta de Instagram @cristianismo_revolucionario. Si bien la Policía de investigaciones de Chile logró detener a la persona que estaba detrás de este desarrollo de videojuego, la cuenta de esta red social que estuvo suspendida por unos 4 meses, al día de hoy está nuevamente activa (39).



> HOSTIGAMIENTO CONTRA LA RED CHILENA DE PERIODISTAS FEMINISTAS

La cuenta de Twitter (40) e Instagram (41) del colectivo de la Red de Periodistas Feministas en Chile @redperiofem, recibe constantemente ataques, especialmente cuando denuncia situaciones o noticias asociadas a feminicidios y casos de violencia hacia mujeres, adolescentes y niñas, o promueve prácticas y lineamientos en relación a la cobertura de los medios de comunicación y prensa en estos casos.

Algunos ejemplos del mes de febrero de 2021, cedidos para el informe por la propia red:



Selección de imágenes de tweets recibidos en la cuenta de @redperiofem en febrero 2021, a raíz de casos públicos de políticos involucrados en violencia contra mujeres.

REFERENCIAS

- (1) El informe está disponible para su descarga en este URL https://datosprotegidos.org/descarga-informe-libertad-de-expresion-en-el-contexto-de-las-protestas-y-movilizaciones-sociales-en-chile-durante-el-estallido-social/>
- (2) Sobre la evolución de la misión de Facebook, revisar el artículo de TheObserver.com en https://observer.com/2009/07/the-evolution-of-facebooks-mission-statement/
- (3) Como señala Guerrero (2020) la política de moderación de contenidos es "una herramienta empleada con el fin de excluir o impedir el acceso a contenido específico dentro de una plataforma o servicio web. Esta prerrogativa suele ser empleada de forma discrecional por las empresas a través de sus políticas de uso o términos del servicio, pero a veces también vienen impuestas por leyes y, en algunos casos, se ejercen por causa de la presión de los gobiernos" en blog HiperDerecho, Perú < https://hiperderecho.org/2020/02/politicas-de-moderacion-de-contenidos/>
- (4) Cobbe, J. (2020), Algorithmic Censorship by Social Platforms: power and resistance, en Philosophy & Technology Journal, disponible para descarga en la URL https://link.springer.com/article/10.1007/s13347-020-00429-0> y en https://doi.org/10.1007/s13347-020-00429-0>
- (5) Perset, K (2019), The Economic and Social Role of Internet Intermediaries, OCDE, 2010, p. 9. Disponible en: https://bit.ly/2frQGWC. Esta definición de intermediarios de Internet también incluye a las empresas de proveedores de servicios de Internet (ISPs, por sus siglas en inglés), empresas de motores de búsqueda, proveedores de servicios en la nube, plataformas de comercio electrónico, empresas proveedoras de hosting, registros de nombres de dominio, entre otros.
- (6) Para más referencias sobre los equipos de moderadores de contenidos trabajando en empresas como Facebook o Tik Tok revisar los siguientes artículos
- El Diario.es- The Guardian (2017) sobre Facebook < https://www.eldiario.es/internacional/theguardian/trabajan-moderadores-facebook-psicologico-necesario 1 3377141.html > y el reportaje de The Intercept sobre la censura invisible y discriminatoria en Tik Tok (2020) < https://theintercept.com/2020/03/16/tiktok-app-moderators-users-discrimination >
- (7) Cobbe, J. (2020), Algorithmic Censorship by Social Platforms: power and resistance, en Philosophy & Technology Journal, disponible para descarga en la URL https://link.springer.com/article/10.1007/s13347-020-00429-0> y en

<<u>https://doi.org/10.1007/s13347-020-00429-0</u>>

- (8) Se denomina cuenta bot a una cuenta o perfil creado para redes sociales, por softwares automatizados que son programados y creados para simular ser personas en una red social y son creadas con un fin determinado, que puede ser desde aumentar artificialmente la cantidad de seguidores de una cuenta o perfil a tener comportamientos de ataques con mensajes automatizados o para reportar y bajar a otras cuentas en una red social.
- (9) Reglas y políticas de Twitter https://help.twitter.com/es/rules-and-policies/twitter-rules
- (10) Reglas y políticas de YouTube https://www.youtube.com/intl/es/about/policies/
- (11) Reglas y políticas de Facebook https://www.facebook.com/communitystandards/
- (12) Reglas y políticas de Instagram https://help.instagram.com/477434105621119?
- (13) Declaración conjunta de Relatores de Libertad de Expresión sobre Libertad de Expresión e Internet, 1 de junio de 2011, suscrita por el Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, la Representante para la Libertad de los Medios de Comunicación de la Organización para la Seguridad y la Cooperación en Europa (OSCE), la Relatora Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP). Disponible en línea:
- < http://www.oas.org/es/cidh/expresion/showarticle.asp?artID=849>
- (14) Comisión Interamericana de Derechos Humanos (CIDH), Libertad de Expresión e Internet, Relatoría Especial para la Libertad de Expresión, documento OEA /Ser.L/V/II.149 Doc. 50.31, diciembre 2013, párr. 111. Disponible en línea:
- < http://www.oas.org/es/cidh/expresion/docs/informes/2014 04 08 Internet WEB.pdf>
- (15) Declaración Conjunta: Desafíos para la libertad de expresión en la próxima década, 2019, del Relator Especial de las Naciones Unidas (ONU) para la Libertad de Opinión y de Expresión, el Representante de la Organización para la Seguridad y la Cooperación en Europa (OSCE) para la Libertad de los Medios de Comunicación, el Relator Especial de la Organización de Estados Americanos (OEA) para la Libertad de Expresión y la Relatora Especial sobre Libertad de Expresión y Acceso a la Información de la Comisión Africana de Derechos Humanos y de los Pueblos (CADHP). Disponible en línea:

< http://www.oas.org/es/cidh/expresion/showarticle.asp?artID=1146&IID=2>

- (16) https://www.observacom.org/regulacion-2/
- (17) El documento propuesta de OBSERVACOM puede ser leído y descargado desde este link https://www.observacom.org/wp-content-uploads-2020-09-estandares-regulacion-grandes-plataformas-internet-pdf/
- (18) https://www.bbc.com/mundo/noticias-internacional-55671725
- (19) Para más detalles, revisar la noticia publicada por la cadena informativa Deutshe Welle https://www.dw.com/es/twitter-censura-a-bolsonaro-por-videos-contra-el-aislamiento-ante-el-covid-19/a-52951405
- (20) Para más información, revisa la noticia en la web del medio informativo Axios.com en https://www.axios.com/youtube-removed-30000-covid19-vaccine-videos-misinformation-a8968086-95a4-4d5e-86da-0e22ddbclb6a.html
- (21) Para un resumen del debate y polémica sobre este caso, revisar las notas publicadas por Xataxa.com https://www.xataka.com/legislacion-y-derechos/trump-twitter-gran-debate-censura-quien-tiene-poder-para-marcar-que-puede-decirse-redes-sociales
 TheConversation.com https://theconversation.com/trumps-twitter-ban-obscures-the-real-problem-state-backed-manipulation-is-rampant-on-social-media-153136">https://theconversation.com/trumps-twitter-ban-obscures-the-real-problem-state-backed-manipulation-is-rampant-on-social-media-153136
- (22) Este formulario está disponible en la URL en el vínculo https://es.surveymonkey.com/r/internetlibrechile
- (23) https://twitter.com/datosprotegidos
- (24) https://www.instagram.com/datosprotegidos/
- (25) Fundación Karisma de Colombia, el Observatorio de Latinoamericano de Regulación de Medios y Convergencia denunciaron esta situación a través de medios y sus redes sociales. Para referencia ver la web de Observacom < https://www.observacom.org/censurabot/facebook-bloqueo-temporalmente-el-hashtag-acab-utilizado-en-las-protestas-sociales-contra-la-represion-policial
- (26) https://www.instagram.com/ficwallmapu/
- (27) https://www.instagram.com/piensaprensa/
- (28) https://twitter.com/PiensaPrensa

- (29) https://www.facebook.com/piensaprensa.org
- (30) https://www.instagram.com/pliegue/
- (31) https://chiletoday.cl/
- (32) https://twitter.com/ChileTodayNews/status/1316023425039704064
- (33) La declaración pública puede leerse en este link del sitio web de la Asociación de Corresponsales de Chile https://www.corresponsaleschile.cl/index.php/2020/10/18/nueva-amenaza-de-muerte-contra-corresponsal-de-la-prensa-extranjera-en-chile/
- (34) La declaración pública sobre este ataque a la web del medio digital puede leerse en el siguiente link https://www.larazon.cl/2020/10/25/la-razon-denuncia-un-ataque-informatico-a-su-sitio-web-de-noticias/
- Zoombombing se denomina a la intrusión indeseada y perturbadora en una videoconferencia, reunión online, generalmente realizada por perfiles troles, personas o cuentas bots o hackers de manera intencional. Usualmente los ataques consisten en la inserción de material obsceno, lascivo, pornográfico o racista. El juego de la palabra se refiere a la plataforma zoom pero puede ser cualquier otra y boombing, en cuanto ocurre como un bombardeo inesperado. Para mayores referencias sobre el tema, revisar el artículo del New York Times en el link < https://www.nytimes.com/2020/04/03/technology/zoom-harassment-abuse-racism-fbiwarning.html>
- (36) https://www.t13.cl/noticia/nacional/udp-denuncia-pornografia-infantil-zoom-17-04-2020
- (37) https://www.tl3.cl/noticia/nacional/hackeo-pornografia-infantil-pucv-tesis-zoom-21-04-2020
- (38) <u>https://www.cnnchile.com/tecnologias/irrumpieron-clase-google-meet-pornoinsultaron-profesora 20200814/</u>
- (39) Para más antecedentes del caso revisar la noticia en 24 horas de TVN en este link https://www.youtube.com/watch?v=xnL3bP8yzS4
- (40) https://twitter.com/redperiofem
- (41) https://www.instagram.com/redperiofem/

CONTACTO





- **o**dcchile
- @OdcChile



- **f** Fundación Datos Protegidos
- @datosprotegidos
- @datosprotegidos