



Tipo de documento: Tesina de Grado de Ciencias de la Comunicación

Título del documento: Un análisis del “bloqueo” a Donald Trump en Twitter en el contexto de las elecciones presidenciales de 2020 en Estados Unidos

Autores (en el caso de tesis y directores):

Juan Baltasar Amor

Mariela Baladron, tutora

Datos de edición (fecha, editorial, lugar,

fecha de defensa para el caso de tesis): 2022

Documento disponible para su consulta y descarga en el Repositorio Digital Institucional de la Facultad de Ciencias Sociales de la Universidad de Buenos Aires.
Para más información consulte: <http://repositorio.sociales.uba.ar/>

Esta obra está bajo una licencia Creative Commons Argentina.
Atribución-No comercial-Sin obras derivadas 4.0 (CC BY 4.0 AR)



La imagen se puede sacar de aca: https://creativecommons.org/choose/?lang=es_AR



La regulación de lo “decible” en las plataformas de redes sociales

Un análisis del “bloqueo” a Donald Trump en Twitter en el contexto de las elecciones presidenciales de 2020 en Estados Unidos

Tesina de grado – Licenciatura en Ciencias de la Comunicación Social

Facultad de Ciencias Sociales – Universidad de Buenos Aires



Autor: Juan Baltasar Amor

DNI: 39387763

Tutora: Mariela Baladron

Febrero 2022

Índice

1.	Introducción	2
2.	Objetivos y abordaje metodológico.....	7
2.1.	Objetivos de investigación	7
2.2.	Abordaje metodológico	9
2.2.1.	Herramientas de recolección de datos	10
2.3.	Estado del arte	12
2.3.1.	La relación entre Trump, Twitter y los medios masivos de comunicación	13
2.3.2.	Aproximaciones al concepto de discurso de odio en redes sociales y a su regulación.....	15
3.	Marco Teórico	18
3.1.	Discursos de odio y libertad de expresión: definiciones y la perspectiva de los Órganos Internacionales de Derechos Humanos.....	19
3.1.1.	Libertad de expresión y funcionarios públicos durante procesos electorales.....	25
3.2.	Legislación y jurisprudencia relevante en Estados Unidos	27
3.3.	Plataformas como posicionamiento discursivo: ¿empresas tecnológicas o empresas de medios?	31
3.4.	Responsabilidad de los intermediarios: definiciones y principios regulatorios	34
3.5.	¿Qué alcance tienen los términos de servicio?	38
3.6.	El acceso a la información en un escenario de concentración.....	42
4.	El ataque al Capitolio y el “bloqueo” a Donald Trump en Twitter: un estudio de caso.....	46
4.1.	Presentación y descripción del caso	46
4.2.	La relación Twitter - Donald Trump y el sistema de “etiquetado”	53
4.3.	Contenidos etiquetados por Twitter y propuesta de regulación	58
4.4.	Tweets eliminados a Donald Trump: las reglas de Twitter y la relación con el interés público64	
4.5.	La suspensión permanente de @realDonaldTrump: motivos, argumentos y reflexiones 70	
4.6.	Términos de servicio, reglas y políticas: propuestas y relación con otros modos de regulación	79
5.	Conclusiones	84
6.	Bibliografía	88

1. Introducción

La tensión entre la libertad de expresión y la existencia de sus límites (legítimos o no) es un fenómeno inherente a la comunicación. En relación con esto, una serie de casos que tomaron relevancia en los últimos años fueron ciertos discursos llamados “de odio” o de “incitación a la violencia” que reavivaron el debate acerca de hasta qué punto pueden realizarse restricciones sin afectar el derecho a la libre expresión.

La Organización de las Naciones Unidas define a estos discursos como:

Cualquier forma de comunicación de palabra, por escrito o a través del comportamiento, que sea un ataque o utilice lenguaje peyorativo o discriminatorio en relación con una persona o un grupo sobre la base de quiénes son o, en otras palabras, en razón de su religión, origen étnico, nacionalidad, raza, color, ascendencia, género u otro factor de identidad (ONU, 2019:3).

Cuando estos fenómenos ganan espacio en la discusión pública resurgen los debates que dan lugar a distintas posiciones. Por un lado, aquellas que se sostienen sobre la idea de que la libertad de expresión es un derecho que debe prevalecer frente a otros, más allá de lo posiblemente dañino de un discurso, por lo que todo intento de censura o restricción debería ser desalentado. Por otra parte, las posturas que consideran que existen otros derechos que también deben ponerse sobre la mesa y que, por lo tanto, habilitan algunas restricciones puntuales sobre el derecho a la libertad de expresión.

En el contexto de la expansión global de las redes sociales, estas plataformas se transforman también en actores preponderantes a la hora de definir qué tipo de restricciones pueden existir en la circulación de los discursos en línea.

En este sentido, la presente tesina propone trabajar la regulación de lo “decible” en las plataformas de redes sociales y los riesgos para la libertad de expresión de la autorregulación que éstas llevan adelante a partir del estudio de un caso paradigmático: el bloqueo de la cuenta de Donald Trump por parte de Twitter en 2021. Este análisis se realiza desde las perspectivas teóricas del derecho a la información y la economía política de la comunicación, ya que brindan conceptos para abordar la regulación y las

prácticas de las redes sociales con una mirada económica de estas plataformas como empresas de comunicación y también del derecho humano a la libertad de expresión.

Una breve descripción del caso y su contexto son importantes para comprender su elección para esta investigación y los interrogantes que se abordan en este trabajo. En enero de 2021, ocurrieron serios incidentes en las inmediaciones y en el interior del Capitolio de los Estados Unidos (edificio que reúne a las dos cámaras del Congreso de ese país). Estos recordados e inéditos hechos involucraron el ingreso al recinto de manifestantes a favor de Donald Trump, quien había cuestionado el resultado de las recientes elecciones en las que había perdido su reelección como presidente. Con su ingreso al edificio, los manifestantes lograron la interrupción de la sesión que se estaba llevando a cabo para confirmar la victoria del contrincante de Trump, Joe Biden. Los hechos continuaron con una importante represión de las fuerzas de seguridad que tuvo como resultado a cuatro personas fallecidas.

El entonces presidente Donald Trump realizó una serie de *tweets* en su cuenta personal (@realDonaldTrump) tanto antes, como durante y luego de los hechos. La plataforma de Twitter decidió que algunos de esos contenidos debían ser eliminados y, el 8 de enero de 2021, dictaminó la suspensión permanente de dicha cuenta. Según lo que expresó la propia plataforma a través de la cuenta oficial @TwitterSafety (traducido como Seguridad de Twitter), este bloqueo se debió a que existía un riesgo de que hubiera “más incitación a la violencia”².

Si bien en el desarrollo del trabajo se describen los hechos con mayor profundidad, cabe detallar en este punto un poco más acerca del contexto en el que se tomaron estas decisiones. El martes 3 de noviembre de 2020 se realizaron las elecciones presidenciales de los Estados Unidos, que enfrentaron como candidatos por los dos partidos más representativos del país a Donald Trump, quien buscaba su reelección, con Joe Biden.

Estas elecciones, a su vez, se vieron envueltas en un contexto muy particular debido a la pandemia del COVID-19 que provocó un crecimiento en la cantidad de votos por correo

² Traducción propia del *tweet* completo: “Luego de una revisión detallada de los *tweets* recientes de la cuenta de @realDonaldTrump y del contexto hemos suspendido permanentemente la cuenta debido a un riesgo de una mayor incitación a la violencia”. Disponible en: <https://twitter.com/TwitterSafety/status/1347684877634838528?s=19> (última consulta: 2/1/2022).

y una demora importante en el conteo de los votos³. Luego del conteo que se extendió por algunos días, Joe Biden fue proclamado como presidente de los Estados Unidos⁴. Sin embargo, Trump no aceptó ese resultado y reclamó en diversas oportunidades que había existido un fraude⁵.

El 6 de enero de 2021, ambas cámaras del Congreso de los Estados Unidos debían reunirse para validar los resultados de las elecciones y proclamar formalmente a Biden en su nuevo rol. Ese día, los manifestantes irrumpieron en el Capitolio y ocuparon el edificio para interrumpir la sesión.

Esa misma noche, la plataforma solicitó la eliminación de tres *tweets* del hasta entonces presidente Donald Trump en relación con lo ocurrido. Según enuncia Twitter, esta decisión se tomó debido a “repetidas y severas violaciones” a su política de Integridad Civil⁶.

En adición a esto, y como se mencionó anteriormente, dos días después de estos hechos se dio la suspensión definitiva de la cuenta de Trump que evitó que el presidente pudiera manifestarse a través de esa vía a partir de ese momento. Si bien el presente trabajo se enfocará en el caso de Twitter, cabe destacar que otras plataformas también suspendieron las cuentas oficiales de Trump o tomaron decisiones para restringir el contenido relacionado con lo ocurrido: Facebook, YouTube, Snapchat, Twitch, TikTok, Reddit y Pinterest, entre otras⁷.

El caso se toma como objeto de análisis ya que dispara numerosas preguntas y caminos de investigación posibles, junto con una serie de interesantes discusiones que se abarcan en el presente trabajo.

³ Redacción BBC Mundo (6/11/2020). “Elecciones de EE.UU. 2020: ¿puede el resultado decidirse en los tribunales?”. BBC. Disponible en: <https://www.bbc.com/mundo/noticias-internacional-54824790> (última consulta: 2/1/2022).

⁴ Redacción BBC Mundo (7/11/2020). “Joe Biden gana las elecciones presidenciales en Estados Unidos”. BBC. Disponible en: <https://www.bbc.com/mundo/noticias-internacional-54845093> (última consulta: 2/1/2022).

⁵ Redacción BBC Mundo (5/11/2020). “Elecciones en Estados Unidos: a qué se refiere Trump cuando habla de “fraude” a favor de Biden y por qué sus acusaciones son infundadas”. BBC. Disponible en: <https://www.bbc.com/mundo/noticias-internacional-54820053> (última consulta: 2/1/2022).

⁶ Traducción propia del *tweet* completo: “Como resultado de la situación violenta sin precedentes y en curso en Washington, D.C., hemos pedido la remoción de tres Tweets de @realDonaldTrump que se publicaron hoy por violaciones graves y repetidas de nuestra política de integridad cívica”. Disponible en: <https://twitter.com/TwitterSafety/status/1346970430062485505> (última consulta: 2/1/2022).

⁷ Redacción Télam (9/1/2021). “Una por una, las redes sociales y plataformas que prohibieron o restringieron a Trump”. Agencia Télam. Disponible en: <https://www.telam.com.ar/notas/202101/541080-redes-sociales-plataformas-trump-elecciones-estados-unidos.html> (última consulta: 2/1/2022)

En primera medida, se reflexiona acerca de la magnitud que puede alcanzar en términos del acceso a la información en una sociedad democrática la decisión editorial de una empresa. En este caso, la restricción generó que millones de ciudadanos estadounidenses perdieran una fuente importante de acceso a su presidente en ejercicio.

Por otra parte, se considera de qué manera las decisiones de las plataformas (fundadas en sus términos de servicio) implican en cierta medida una contradicción con otras formas de regulación. Por ejemplo, Estados Unidos se encuentra regido por un marco legal que tiende a la defensa irrestricta de la libertad de expresión casi ante cualquier circunstancia, algo que puede no verse reflejado en el escenario que se trae a colación.

Así, el objetivo principal que encauza esta investigación es comprender cuáles son los riesgos para el ejercicio de la libertad de expresión que implica la autorregulación de las plataformas de redes sociales.

De este objetivo central, se desprenden los objetivos específicos: entender los modos en los que se regula el contenido en las plataformas de redes sociales, la responsabilidad que le cabe a los intermediarios y las posibilidades de intervención de los Estados, entre otros.

En el segundo capítulo de esta tesina de investigación se desarrollan el objetivo general y los objetivos específicos de este trabajo, en conjunto con el abordaje metodológico y las herramientas de recolección de información que fueron utilizadas. En pocas líneas, se puede indicar que el análisis se realiza bajo la metodología de un estudio de caso para el cual se realiza una descripción densa de las especificidades del evento a analizar, con el objetivo de generar conocimiento acerca de las formas de regulación actuales de los discursos en redes sociales. Por otro lado, el capítulo también cuenta con una recopilación del estado del arte con el fin de recuperar materiales relevantes ya publicados sobre el objeto de análisis.

A continuación, el tercer capítulo está dedicado al marco teórico que aporta los conceptos y perspectivas teóricas seleccionados para el abordaje del caso. Este capítulo está dividido en seis apartados en los que se trabajan los siguientes temas: definiciones de los discursos de odio y su tensión con la libertad de expresión con el aporte de la perspectiva de los Órganos Internacionales de Derechos Humanos; la legislación y jurisprudencia relevante de los Estados Unidos, considerando el lugar donde sucedieron

los hechos; el debate acerca de la consideración de las plataformas de redes sociales como empresas de medios o empresas de tecnología; la cuestión de la responsabilidad que les cabe a las plataformas en tanto intermediarios; el alcance que pueden tener los términos de servicio diseñados por las plataformas; y la problemática del acceso a la información en un escenario de concentración.

El cuarto capítulo se dedica al análisis del caso elegido, basado en un corpus elaborado a partir de *tweets* y declaraciones públicas del ex presidente Trump y de los responsables de la red social Twitter. Para ello, se traza un recorrido que comienza en una descripción detallada del bloqueo de la cuenta de Twitter de Donald Trump. Luego, se recuperan los antecedentes de la relación entre esta red social y el ex presidente estadounidense con el fin de poder observar las primeras tensiones entre ambos actores y la aparición de distintas propuestas de modificaciones regulatorias.

Por último, se realiza un análisis respecto de la eliminación de contenidos y la suspensión definitiva de la cuenta del entonces presidente, sobre la base de las declaraciones públicas de su CEO en el momento de los hechos, los *tweets* publicados por la propia red social y los cambios en las reglas comunitarias que se sucedieron. Allí, se apunta a una elaboración tanto de los argumentos de las decisiones tomadas como de las consecuencias que éstas pueden tener en términos de la regulación de lo “decible” en las plataformas de redes sociales, a partir del caso de estudio trabajado.

A modo de cierre, en las conclusiones se sistematizan los principales hallazgos de esta tesina y se proponen nuevos interrogantes para continuar en futuras investigaciones que continúen la producción de conocimiento sobre esta problemática contemporánea, a la cual los estudios de comunicación tienen importantes aportes que realizar.

2. Objetivos y abordaje metodológico

2.1. Objetivos de investigación

A continuación, se presenta el objetivo general y los objetivos específicos que guían esta investigación, con una breve descripción de la forma en que se llevan adelante en los subsiguientes capítulos.

Objetivo general:

- Reflexionar acerca de los riesgos para el ejercicio de la libertad de expresión que implica la autorregulación de las plataformas de redes sociales a partir del análisis del bloqueo a la cuenta de Twitter de Donald Trump en 2021.

Objetivos específicos:

- Comprender la manera en que se regula la información a la que los usuarios pueden acceder en redes sociales.
- Indagar acerca del alcance que pueden tener las decisiones corporativas de las plataformas, fundadas en los términos de servicio.
- Explorar las responsabilidades que les caben a las plataformas -en tanto intermediarios- por los discursos que emiten sus usuarios.
- Identificar las posibilidades de intervención de los Estados a la hora de la regulación de lo “decible” en redes sociales.
- Analizar los argumentos de Twitter respecto de la eliminación de contenido y la posterior suspensión permanente de la cuenta del entonces presidente en ejercicio de los Estados Unidos.

El objetivo general de este trabajo es reflexionar acerca de los riesgos para el ejercicio de la libertad de expresión que implica la autorregulación de las plataformas de redes sociales. Con este fin, se realiza el análisis de un caso paradigmático: el bloqueo a la cuenta de Twitter de Donald Trump en enero de 2021, en el contexto de las elecciones presidenciales de los Estados Unidos ocurridas a finales de 2020.

De forma complementaria, se trabaja con los objetivos específicos que se desprenden de este objetivo general. El primero de ellos apunta a comprender de qué modo se regula la información a la que los usuarios de las redes sociales pueden acceder. Para ello, se realizó una revisión bibliográfica y se incorporó en el marco teórico una síntesis de los debates actuales sobre estas regulaciones, algunas normativas existentes y sus propuestas de modificación.

El segundo objetivo específico consiste en analizar qué alcance tienen las decisiones de quienes gestionan las plataformas. En este caso en particular, se hace referencia al hecho de observar si existen maneras alternativas de definir qué contenidos pueden ser publicados en línea, más allá de los criterios que las propias plataformas establecen. Para ello, se sistematizan los documentos y recomendaciones de los órganos internacionales de derechos humanos y otras propuestas elaboradas por organizaciones de la sociedad civil especializadas en derechos digitales. A su vez, se focaliza en debates actuales respecto de líderes mundiales y procesos electorales, ya que el usuario involucrado en las medidas restrictivas es una figura de relevancia pública mundial como Donald Trump, presidente en ejercicio de los Estados Unidos en el momento de los hechos.

El tercer objetivo específico propone estudiar la responsabilidad que les cabe a los intermediarios (en este caso encarnados en las plataformas de redes sociales y particularmente en Twitter) por los contenidos que sus usuarios emiten. Con este fin, se realiza una sistematización sobre el rol de las plataformas como editores, los riesgos de la concentración de estos gigantes tecnológicos y la legislación estadounidense respecto de los intermediarios, en consonancia con el caso elegido.

Adicionalmente, se busca entender qué intervención pueden tener los Estados a la hora de esta regulación de lo “decible” en redes sociales. En este sentido, se trabaja sobre las leyes y la jurisprudencia existente en la materia para considerar qué contradicciones existen entre ellas y los términos de servicio que elaboran las propias plataformas y en las cuales basan sus decisiones.

Finalmente, se analizan los argumentos esgrimidos por la propia plataforma de Twitter respecto de la eliminación de contenido y la posterior suspensión permanente de la cuenta del entonces presidente en ejercicio de los Estados Unidos. Con este objetivo, se construyó un corpus basado en *tweets* y declaraciones públicas de Twitter al respecto, el

cual es analizado en su contexto histórico y situado desde la perspectiva del derecho a la información y la economía política de la comunicación.

2.2. Abordaje metodológico

El abordaje metodológico con el que se procede durante el análisis se puede categorizar bajo el concepto de estudio de caso.

Esta elección se basa en el acercamiento a la comprensión de un fenómeno general a través de un caso que se considera que cuenta con “características paradigmáticas o ejemplificadoras” (Marradi, Archenti y Piovani, 2007:241). En el presente trabajo, el caso “paradigmático” a estudiar es el bloqueo de la cuenta de Twitter del ex presidente estadounidense Donald Trump cuando aún estaba en ejercicio de sus funciones. Sus características únicas -respecto de los límites y las atribuciones de la moderación privada de contenidos sobre temas de interés público y funcionarios públicos- permiten analizar a lo largo del desarrollo de esta tesina la forma de regulación del discurso público en las redes sociales.

En esta misma dirección se pueden encontrar también las tipologías de Robert Stake (1999) y de Robert Yin (1994) que ayudan a describir con más precisión las características del estudio de caso de este trabajo.

Siguiendo los conceptos de Stake (1999) se apunta hacia un estudio de caso de tipo intrínseco, ya que se trabaja en una “descripción densa” de las especificidades del caso que despertó el interés en este fenómeno. Sin embargo, también existen características del tipo instrumental ya que este estudio brinda algunas líneas que permiten comprender “un fenómeno que lo trasciende” como lo es la regulación de los discursos en las redes sociales.

Si se toman las tipologías de Robert Yin (1994), se puede hablar de un estudio del tipo holístico entendiendo que se trabaja sobre un único caso. A su vez, con respecto al objetivo de la investigación, y tal como se ha sugerido, se trabaja principalmente con un estudio descriptivo ya que se busca dar la mayor profundidad posible a la descripción del caso.

Así, se realiza una descripción profunda del caso de estudio que tiene su inicio en los antecedentes de la relación de Donald Trump y Twitter (durante los años previos a su asunción como presidente) y la marcada preferencia que él tenía para con esta plataforma.

Posteriormente, se describen los hechos que llevaron a los primeros conflictos entre ambas partes, principalmente encarnados en las acusaciones hacia Trump por la publicación de “información engañosa” relacionada con las elecciones y el sistema de votación por correo. Finalmente, se abordan las escenas de violencia ocurridas en el Capitolio y el contenido compartido por Trump durante esos días, que se convertiría en el desencadenante para el bloqueo definitivo de sus cuentas.

Por otra parte, se encuentran en este trabajo también muchas características de un estudio exploratorio, comprendiendo que se realiza un acercamiento a un fenómeno que podría derivar en investigaciones posteriores. Esto se debe principalmente a que, al tratarse de un caso reciente, no se encuentran en circulación demasiados trabajos en el ámbito académico que analicen la relación entre este caso y la regulación de los discursos en las redes sociales.

En este “estudio denso” del caso de análisis se narran las especificidades, antecedentes y situaciones relacionadas con el fin de obtener “indicios” que permitan un acercamiento a los objetivos (Marradi, Archenti y Piovani, 2007:240). De este modo, revisten importancia tanto el trabajo de archivo con respecto a la particular relación entre Trump y Twitter como el análisis exhaustivo del contexto que rodeaba a los hechos que ocurrieron en el Capitolio: la preocupación por la circulación de “información engañosa” y la polémica por el sistema de votos por correo son algunos ejemplos de esto.

2.2.1. Herramientas de recolección de datos

En primer lugar, este trabajo recurre a la revisión bibliográfica que permitió construir el estado del arte y el desarrollo del marco teórico de este trabajo, que incluye los debates actuales en torno a la regulación de los discursos de odio y la incitación a la violencia en las redes sociales, la normativa y jurisprudencia estadounidense al respecto, las

recomendaciones de los órganos internacionales de derechos humanos sobre la temática y el rol de las plataformas como editores en un ecosistema altamente concentrado, entre los principales.

Luego, ante la dificultad de acceder a los actores involucrados en este caso de estudios a través de fuentes directas, se recurrió a un relevamiento y sistematización de las declaraciones públicas en medios de comunicación y redes sociales oficiales, tanto de Donald Trump como de las máximas autoridades y representantes de las plataformas de redes sociales involucradas.

El corpus que se utiliza para el análisis está compuesto en su mayoría por publicaciones hechas en la plataforma Twitter (popularmente conocidos como *tweets*). Principalmente, se toman aquellas publicaciones que han sido compartidas por la cuenta oficial de Donald Trump (@realDonaldTrump) en un intervalo que va desde abril de 2020 – cuando comenzaron las publicaciones de Trump en contra de las elecciones por correo– hasta enero de 2021, momento en el que la cuenta fue suspendida.

Cabe destacar que esta cuenta continuaba suspendida al momento de la escritura de esta tesina. Por este motivo, se tomaron como fuentes distintos medios de comunicación que recuperaron los *tweets* del ex presidente, así como también se recurrió a un archivo⁸ que recopiló todas las publicaciones de la cuenta en cuestión.

Por otro lado, se trabaja también sobre las publicaciones oficiales de Twitter (y de sus representantes) que la plataforma utilizó para argumentar sus decisiones de restricción de contenido, así como también las defensas que utilizó ante ciertos comentarios de Trump.

La elección de Twitter para este trabajo dentro del amplio abanico de plataformas se debe a la relación particular de Donald Trump con esta red social, a la cual le daba un uso intenso desde mucho antes de asumir como presidente. Sobre este vínculo se ahonda en los apartados de análisis.

Además, esta elección se respalda en el hecho de que las autoridades de Twitter han remarcado en varias ocasiones que su plataforma apunta a unos términos de servicio claros y a la transparencia en la toma de decisiones.

⁸ Archivo digital que recopila todos los *tweets* de la cuenta @realDonaldTrump. Disponible en: <https://www.thetrumparchive.com/> (última consulta 13/11/2021)

Jack Dorsey, co-fundador de Twitter y CEO al momento de la escritura de esta tesina, expresó en una de sus declaraciones⁹ ante el Senado de los Estados Unidos en octubre de 2020 que era fundamental para él que las personas entendieran sus procesos y que, tanto las reglas de moderación de contenido como sus efectos, “deberían ser explicados de forma sencilla y comprensible para cualquier persona”. A su vez, agregó que empresas como Twitter deberían “publicar sus procesos de moderación” y “ser transparentes acerca de cómo los casos se reportan y se revisan, cómo se toman las decisiones y qué herramientas se utilizan para hacerlas cumplir”.

En noviembre del 2020, en una nueva declaración ante el Senado¹⁰, Dorsey remarcó que la “naturaleza abierta” de Twitter hace que sus acciones sean “claramente visibles para el público, incluso cuando no podemos revelar los detalles privados de las cuentas individuales que han violado nuestras Reglas”. En este mismo discurso, Dorsey reconoció que Twitter podría seguir mejorando para “continuar ganándose la confianza” de sus usuarios y remarcó la importancia de saber reconocer cuándo sus políticas de servicio necesitan ser actualizadas ante “circunstancias nuevas o imprevistas”.

De este modo, se puede ver cómo las ideas de la transparencia y la actualización constante en los términos de servicio -y la posibilidad de que éstos sean accesibles y comprensibles para todos los usuarios- parecerían ser preocupaciones importantes para las autoridades de Twitter.

2.3. Estado del arte

Tal como se ha esbozado en la introducción a este trabajo, las problemáticas de los discursos de odio y la incitación a la violencia no son fenómenos de aparición reciente, pero han recobrado una mayor relevancia en el debate público. Principalmente, se puede observar un creciente interés por comprender cómo estos asuntos se vinculan con la circulación del discurso público en las redes sociales y con sus posibles restricciones.

⁹ Declaración completa de Jack Dorsey ante el Comité de Comercio, Ciencia y Transporte del Senado de los Estados Unidos (en inglés). Disponible en: <https://www.washingtonpost.com/context/twitter-ceo-jack-dorsey-s-statement-to-senate-commerce-committee/c6e7da76-67c9-4136-b846-ce9e7a1f0eaa/> (última consulta 13/11/2021)

¹⁰ Declaración completa de Jack Dorsey ante el Comité Judicial del Senado de los Estados Unidos (en inglés). Disponible en: <https://www.judiciary.senate.gov/download/dorsey-testimony> (última consulta 13/11/2021)

Adicionalmente, el caso de análisis que se trabaja en esta tesina resulta de gran actualidad. Por lo tanto, y si bien se han encontrado numerosos trabajos en artículos de medios especializados, no existen aún tantos trabajos académicos específicamente sobre las cuestiones que aquí se abordan. A continuación, entonces, se recuperan algunos trabajos que son relevantes al haberse centrado en cuestiones relacionadas con el tema en cuestión, aunque con objetivos y enfoques diferentes.

2.3.1. La relación entre Trump, Twitter y los medios masivos de comunicación

Este primer apartado del estado del arte recopila algunos trabajos que analizaron el vínculo entre Donald Trump y las redes sociales, con un particular foco en Twitter. Además, los autores de estos textos trabajan acerca de la importante repercusión que las publicaciones de Trump han tenido en los medios masivos de comunicación de Estados Unidos.

Justel-Vázquez, Fernández-Planells, Victoria-Mas y Lacasa-Mas (2018) analizaron la manera en que Twitter se ha convertido en una herramienta cada vez más utilizada para la comunicación política. En este sentido, los autores destacaron la forma en que los políticos utilizan esta plataforma para comunicar declaraciones y propuestas políticas, complementando así a las formas de difusión más tradicionales.

A su vez, el punto quizás más interesante para este trabajo, tiene que ver con el análisis de los investigadores. Sus conclusiones indican que, principalmente en el caso de Trump y debido a su prolífera actividad en Twitter, los *tweets* de los actores protagónicos de la política se han convertido en fuentes legítimas y recuperadas cada vez con más frecuencia por los medios de comunicación tradicionales a la hora de redactar sus artículos (Justel-Vázquez, *et. al.*, 2018).

En un sentido similar apunta el trabajo de Pérez-Curiel y Limón Naharro (2019). Luego de su investigación, las autoras concluyen que los *tweets* de Donald Trump durante su presidencia han podido “fijar y condicionar la agenda de los medios tradicionales” (Pérez-Curiel y Limón Naharro, 2019:16).

Además, las investigadoras comprenden el perfil de Trump en Twitter como el de un “influencer”. De este modo, entienden que el ex presidente encontró en esa plataforma

un escenario para emitir discursos rápidos, llamativos y efectivos, que captan la atención de sus seguidores a pesar de tener una interacción casi nula con ellos. Al mismo tiempo, las autoras destacan que el ex presidente recurría a discursos basados en el odio y las mentiras para propinar ataques a distintos colectivos, temas que luego serían recuperados en las portadas de los grandes medios de comunicación (Pérez-Curiel y Limón Naharro, 2019).

También, ha sido publicado un trabajo de Stolee y Caton (2018) en el que se analizan de forma más específica los *tweets* de Trump, concluyendo que éste apuntaba discursivamente a una “base” más fiel y activa de seguidores, aun cuando esto significara resignar una posibilidad de alcance mayor. Al igual que en los trabajos anteriores, los autores concluyen que los medios masivos de comunicación luego tomaban estas publicaciones como fuentes de consulta frecuente y las amplificaban hacia otros públicos.

Al mismo tiempo, los investigadores entienden que Twitter, por características inherentes a la plataforma, ha sido uno de los elementos fundamentales para permitir el ascenso al poder de Trump. Esto se debe principalmente a que, a pesar de la supuesta idea de la plataforma de promover el diálogo y el intercambio libre de opiniones, se han formado comunidades –que los autores describen como de ultraderecha y con posiciones cercanas a Trump– que encontraron en Twitter un espacio donde desplegar el abuso y la intimidación para con otros usuarios (Stolee y Caton, 2018).

De este modo, se puede observar que existe una amplia gama de trabajos realizados acerca de la utilización que le daba Donald Trump a su cuenta de Twitter. Estos estudios demostraron aspectos relacionados con la relevancia que estas publicaciones tenían para la agenda mediática de los Estados Unidos y trabajaron también sobre un análisis de los discursos en sí mismos.

Sin embargo, se observa que aún existe la posibilidad de explorar el vínculo de Trump y Twitter desde el punto de vista de la regulación del discurso en las redes sociales. Más aún si se considera el desenlace que dicha relación tuvo, es decir, la limitación de contenidos y la suspensión permanente de la cuenta del ex presidente. A estos objetivos, entre otros que ya han sido mencionados, se apunta en esta tesina.

2.3.2. Aproximaciones al concepto de discurso de odio en redes sociales y a su regulación

Por otra parte, se pueden encontrar trabajos que sirven como antecedentes en otras materias relacionadas con el objeto de esta tesina. Un primer ejemplo es el de Cabo y García (2016), en el que se realiza un acercamiento al concepto de discurso de odio y se elabora un desarrollo sobre el rol que juegan las redes sociales sobre este objeto. El enfoque de este documento está pensado también en las estrategias que existen para combatir estos discursos de odio en la Unión Europea y, específicamente, en España.

Otro artículo interesante que se puede retomar en este punto es el de Bustos Martínez, De Santiago Ortega, Martínez Miró y Rengifo Hidalgo (2019) en el que nuevamente se realiza un acercamiento al esquivo concepto de discurso de odio. Si bien, como ya se planteó, no se trata de un concepto de aparición reciente, sí se ha observado en el último tiempo un resurgimiento de las discusiones sobre esta problemática. Los investigadores se han visto empujados a redefiniciones constantes de este concepto, sobre todo a la luz de la proliferación de los discursos de odio y las manifestaciones de incitación a la violencia en las plataformas de redes sociales.

De este modo, en el citado artículo se recapitulan algunas definiciones que coinciden en que los discursos de odio involucran la agresión o la humillación hacia grupos vulnerables o históricamente discriminados. A su vez, los autores trabajan sobre la afirmación de que la expansión global de Internet en general, y de las redes sociales en particular, ha sido un factor importante para la propagación de este tipo de discursos (Bustos Martínez, *et. al.*, 2019).

Al igual que en el texto de Cabo y García (2016), este artículo hace foco en un repaso de las acciones de prevención de los discursos de odio, tanto en España como en el mundo en general. En el caso del país europeo, se recuperan, por ejemplo, modificaciones del Código Penal que apuntan al castigo para aquellos que incurrieran en la difusión de expresiones racistas o xenófobas en Internet. Además, se recuperan algunas iniciativas de órganos internacionales tales como ACNUR (Alto Comisionado de las Naciones Unidas para los Refugiados) que van también en la dirección del combate contra estos fenómenos, aunque por ahora sin impacto en modificaciones regulatorias (Bustos Martínez, *et. al.*, 2019).

Por otra parte, es muy rico el trabajo acerca de la libertad de expresión en relación con las expresiones de odio, principalmente desde la disciplina del Derecho. Se retoman entonces algunos estudios que hacen referencia a conceptos del derecho a la comunicación y que resultan de importancia para avanzar en esta tesina.

En el trabajo de Pauner Chulvi (2018) hay una reflexión acerca de los peligros que podría acarrear un sistema de censura privada en términos de pluralismo. Además, la autora piensa acerca de los riesgos de la arbitrariedad al momento de la restricción de los contenidos. Finalmente, en este texto se debate en torno a la responsabilidad de las redes sociales en tanto intermediarios, un tema en el cual se profundiza en el marco teórico.

Sobre esta materia, se han expresado también distintas organizaciones de la sociedad civil que trabajan temáticas de libertad de expresión e Internet. En un artículo de Derechos Digitales, Díaz Hernández (2020) trabaja sobre la necesidad de una proporcionalidad en las medidas restrictivas de los discursos de odio y, a su vez, en la dificultad adicional que supone esta regulación en las plataformas en línea.

La autora hace un especial énfasis en los enfoques que se le han dado a esta problemática en América Latina y concluye que, en lugar de mecanismos de prevención y educación que promuevan debates más amplios y plurales, los países se han focalizado en un acercamiento “legislativo y criminalizador”, ya sea a través de sanciones penales (directas o accesorias), o bien prohibiciones que repercuten en medidas “reparatorias” (Díaz Hernández, 2020).

A su vez, *Access Now* (2020) redactó un documento con 26 recomendaciones para la gobernanza de los contenidos en el espacio digital. Para ello, plantearon una primera distinción entre tres tipos posibles de regulación: la regulación estatal, a través de leyes o regulaciones locales, nacionales o regionales; la auto-regulación, aquella que le compete a las propias plataformas y que se apoya en los términos de servicio; y la co-regulación, un mecanismo que sería similar a la auto-regulación, pero con un apoyo y/o monitoreo activo de parte de las autoridades estatales (*Access Now*, 2020).

Para cada una de estas tres categorías, corresponden las recomendaciones que se realizan en ese trabajo. No se realiza aquí un detalle de cada una de ellas, pero a modo de una visión general se puede resumir que estas sugerencias se relacionan con la

necesidad de encontrar mecanismos más transparentes, proporcionales, democráticos, participativos y que concluyan en la menor afectación posible para los derechos humanos que se ponen en juego (*Access Now*, 2020).

En esta recopilación del estado del arte, se han tomado como referencia trabajos con objetivos variados. Por una parte, aquellos que analizaron la figura de Donald Trump, su relación con Twitter y cómo los medios de comunicación utilizaban sus publicaciones como fuentes informativas. Asimismo, se recuperaron documentos que se enfocan en el concepto de los discursos de odio, las facilidades que las redes sociales habrían generado para su propagación y las posibilidades de regulación. Con este fin, se documentaron trabajos académicos del campo de la comunicación, otros del ámbito del Derecho y, finalmente, artículos redactados por organizaciones de la sociedad civil especializadas en derechos digitales.

Así, si bien hay temáticas que se abarcan en esta tesina que se apoyan en conceptos ya trabajados, se puede observar que la relación específica entre el caso de análisis y la problemática de la regulación de los discursos en redes sociales no ha sido aún profundizada. A su vez, resulta un factor diferencial la relevancia de la figura sobre la cual se aplicaron las restricciones en este caso en particular, es decir, el presidente en ejercicio de los Estados Unidos.

3. Marco Teórico

Este capítulo de la tesina de investigación recopila conceptos relevantes para los objetivos de este trabajo y el análisis propuesto. El marco teórico consta de seis secciones que se describen detalladamente en cada uno de los correspondientes apartados.

Las perspectivas teóricas a las que se recurre son el derecho a la información y la economía política de la comunicación, con el fin de comprender el fenómeno contemporáneo de la concentración del discurso público en plataformas globales, con un foco especial en las plataformas de redes sociales, que en su mayoría tienen sus sedes centrales en los Estados Unidos.

Asimismo, se analiza un fenómeno que resulta cada vez más frecuente: la autorregulación que realizan las plataformas, específicamente en cuanto a los discursos de odio y la incitación a la violencia, y las consecuencias que esto tiene en términos de libertad de expresión.

Los conceptos que se desarrollan incluyen: definiciones de los discursos de odio y sus vínculos con la libertad de expresión; legislación y jurisprudencia relevante en los Estados Unidos; el concepto de “plataforma” que utilizan las empresas en cuanto posicionamiento discursivo y estratégico; la responsabilidad que le cabe a los intermediarios por los discursos de sus usuarios; los términos de servicio de las plataformas y su alcance; y una descripción del escenario de concentración en el acceso a la información.

Sobre la base de estas perspectivas, se busca analizar el caso de Donald Trump y el bloqueo de su cuenta de Twitter que tuvo lugar en enero del 2021, luego de los hechos de violencia ocurridos en el Capitolio. Este caso paradigmático pone en cuestión quiénes son los actores que regulan el discurso público –o lo “decible”– incluso cuando el actor involucrado es el primer mandatario de uno de los países más poderosos del planeta. A su vez, permite plantear preguntas acerca del rol que las plataformas de redes sociales tienen para informar sobre temas de interés público y, en ese sentido, cuáles son los límites que la expresión en esas plataformas puede tener dentro de un sistema democrático.

3.1. Discursos de odio y libertad de expresión: definiciones y la perspectiva de los Órganos Internacionales de Derechos Humanos

Para dar comienzo a este apartado, se retoman algunas definiciones de los discursos de odio y, sobre esta base, se propone un repaso por el abordaje que los Órganos Internacionales de Derechos Humanos han realizado sobre la problemática.

En mayo de 2019, la Organización de las Naciones Unidas desarrolló su “Estrategia y Plan de Acción de las Naciones Unidas para la lucha contra el Discurso de Odio”. De este documento¹², surge una definición acerca de este objeto:

Discurso de odio es cualquier forma de comunicación de palabra, por escrito o a través del comportamiento, que sea un ataque o utilice lenguaje peyorativo o discriminatorio en relación con una persona o un grupo sobre la base de quiénes son o, en otras palabras, en razón de su religión, origen étnico, nacionalidad, raza, color, ascendencia, género u otro factor de identidad (ONU, 2019).

En el prefacio de este mismo documento aparece el siguiente párrafo:

Hacer frente al discurso de odio no significa limitar la libertad de expresión ni prohibir su ejercicio, sino impedir que este tipo de discurso degeneren en algo más peligroso, como la incitación a la discriminación, la hostilidad y la violencia, que están prohibidas por el derecho internacional (ONU, 2019).

A propósito de esto, Víctor Abramovich (2021) analiza la aparente contradicción entre la demanda de intervención estatal con el fin de restringir la circulación de los discursos de odio –entendiendo que éstos atentan contra los principios de igualdad y no discriminación– y la prohibición de la censura previa y de otros mecanismos para la regulación de la libertad de expresión.

El autor describe también la evolución del concepto de libertad de expresión y la contraposición central entre dos posturas. Por una parte, una postura clásica que entiende a la libertad de expresión sobre la base de la autonomía y que desemboca en una prohibición de casi toda intervención estatal sobre los discursos. Del otro lado,

¹² Texto completo del documento “Estrategia y Plan de Acción de las Naciones Unidas para la lucha contra el Discurso de Odio”. Disponible en: https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_ES.pdf (última consulta: 22/11/2021)

aparece una idea de la libertad de expresión como facilitadora de un debate público robusto en el que todas las voces tengan una posibilidad igualitaria de expresión. La conclusión a la que se arriba en este artículo apunta a una búsqueda de trabajar en mecanismos de acción colectiva para alcanzar objetivos de igualdad sin poner en riesgo lo central de la libertad de expresión (Abramovich, 2021).

Queda claro entonces como ambos fenómenos –los discursos de odio y la libertad de expresión– están estrechamente relacionados y resulta casi imposible hablar de uno de estos términos sin el otro.

Antes de continuar con los documentos que trabajan este vínculo, corresponde clarificar dos cuestiones con respecto a los Órganos Internacionales y Regionales de Derechos Humanos. En primer lugar, que de acuerdo con la Carta de las Naciones Unidas¹³, las naciones miembro se comprometen a aceptar y cumplir solo aquellas decisiones del Consejo de Seguridad. Si bien el resto de las declaraciones y recomendaciones de la ONU son de indudable relevancia, tienen un peso simbólico en la comunidad internacional y sus principios son receptados en las regulaciones nacionales, no están previstas sanciones hacia los Estados ante posibles incumplimientos en las materias relacionadas con este trabajo.

Por otro lado, cabe aclarar también que Estados Unidos no ha ratificado la Convención Americana sobre Derechos Humanos¹⁴. De todas formas, se citan estos documentos ya que funcionan como un marco de referencia internacional acerca de estas temáticas. Con el fin de complementar la información, en el apartado siguiente se trabaja específicamente con el marco legal y la jurisprudencia de los Estados Unidos sobre la materia para una mejor comprensión de todas las aristas del caso de análisis.

En el artículo 19 de la Declaración Universal de Derechos Humanos de la Organización de las Naciones Unidas de 1948 se establece:

Todo individuo tiene derecho a la libertad de opinión y de expresión; este derecho incluye el de no ser molestado a causa de sus opiniones, el de investigar y recibir informaciones y opiniones, y el de difundirlas, sin limitación de fronteras, por cualquier medio de expresión (ONU, 1948).

¹³ Carta de las Naciones Unidas. Texto completo disponible en: <https://www.un.org/es/about-us/un-charter/full-text> (última consulta: 27/11/2021)

¹⁴ Sitio oficial de la Convención Americana sobre Derechos Humanos: <https://www.oas.org/es/CIDH/jsForm/?File=/es/cidh/mandato/que.asp> (última consulta: 27/11/2021)

Por otro lado, el artículo 13 de la Convención Americana sobre Derechos Humanos, también conocido como el Pacto de San José de Costa Rica¹⁵, indica:

1. Toda persona tiene derecho a la libertad de pensamiento y de expresión. Este derecho comprende la libertad de buscar, recibir y difundir informaciones e ideas de toda índole, sin consideración de fronteras, ya sea oralmente, por escrito o en forma impresa o artística, o por cualquier otro procedimiento de su elección.
2. El ejercicio del derecho previsto en el inciso precedente no puede estar sujeto a previa censura sino a responsabilidades ulteriores, las que deben estar expresamente fijadas por la ley y ser necesarias para asegurar:
 - a) el respeto a los derechos o a la reputación de los demás, o
 - b) la protección de la seguridad nacional, el orden público o la salud o la moral públicas (CADH, 1969).

En palabras de Loreti y Lozano (2021), se puede afirmar que esta perspectiva de la CADH está entre las más “generosas” en cuanto a la prohibición de la censura previa si se la compara con otros sistemas regionales de Derechos Humanos, como por ejemplo el europeo.

Continuando con el artículo 13, el inciso quinto establece:

Estará prohibida por la ley toda propaganda en favor de la guerra y toda apología del odio nacional, racial o religioso que constituyan incitaciones a la violencia o cualquier otra acción ilegal similar contra cualquier persona o grupo de personas, por ningún motivo, inclusive los de raza, color, religión, idioma u origen nacional (CADH, 1969).

De esta manera, los discursos (o apologías) de odio pasarían a ser regulados por la legislación propia de cada país que, siguiendo la doctrina de la CADH, debería prohibir estas expresiones.

En un sentido similar, aparecen los argumentos de la Convención Internacional sobre la Eliminación de todas las Formas de Discriminación Racial de la ONU firmada en 1965. En su artículo 4º se condena como acto punible a:

¹⁵ Texto completo de la Convención Americana sobre Derechos Humanos (Pacto de San José de Costa Rica). Disponible en: https://www.oas.org/dil/esp/tratados_b-32_convencion_americana_sobre_derechos_humanos.htm (última consulta: 28/11/2021)

toda difusión de ideas basadas en la superioridad o en el odio racial, toda incitación a la discriminación racial, así como todo acto de violencia o toda incitación a cometer tales actos contra cualquier raza o grupo de personas de otro color u origen étnico (ONU, 1965).

Otras definiciones pueden encontrarse en un documento¹⁶ del Center for International Media Assistance (CIMA), compuesto en conjunto por miembros de la UNESCO y de la Relatoría Especial para la Libertad de Expresión (RELE) de la Comisión Interamericana de Derechos Humanos (CIDH) en 2017.

Frank La Rue¹⁷ plantea que el discurso de odio debe ser visto como una “excepción” para evitar de este modo “caer en prohibiciones que atentan contra el derecho a la libertad de expresión o en la creación de un contexto en el que las personas dejen de expresarse libremente por miedo a ser penadas por ello” (CIMA, 2017).

En este mismo documento, se establece que si bien no existe una definición universalmente aceptada en torno a este fenómeno, sí hay unanimidad en cuanto a que la prohibición debe ser una “excepción”. Estas excepciones, entonces, son las que correspondería definir y codificar de acuerdo con esta doctrina.

Algunas de las “excepciones” que la propia CIDH establece tienen que ver con aquellos discursos que constituyen una “incitación al genocidio y otras violaciones del derecho internacional” o la “promoción del odio discriminatorio que constituye una incitación a la hostilidad, la discriminación o la violencia”. Se establece también allí una “preocupación” acerca de “la existencia y utilización de leyes que reprimen el derecho a la libertad de expresión en nombre del combate al discurso del odio”.

Un ejemplo paradigmático de esta clase de leyes se da en Alemania. Schulz (2018) explica que la regulación conocida como “NetzDG”, que entró en vigencia en 2018, insta a las plataformas de redes sociales a mantener procedimientos “efectivos y transparentes” para gestionar las quejas acerca de contenido ilegal. La ley también establece plazos diferenciales para el manejo de las quejas y la remoción de los contenidos, con algunas diferenciaciones de acuerdo con qué tan “evidente” sea la

¹⁶ Texto completo del documento redactado por el *Center for International Media Assistance* (CIMA). Disponible en: <https://www.corteidh.or.cr/tablas/r37048.pdf> (última consulta: 22/11/2021)

¹⁷ Frank La Rue ocupaba al momento de la redacción del citado texto (2017) el cargo de Subdirector General para la Comunicación e Información de la UNESCO. Anteriormente, había ejercido como Relator de las Naciones Unidas sobre la Protección y Promoción de la Libertad de Opinión y Expresión.

ilegalidad de dichos contenidos. Por otro lado, se establecen sanciones económicas para aquellos proveedores que incurran en infracciones reiteradas sobre estos puntos (Schulz, 2018).

El autor plantea algunas consideraciones respecto de esta normativa. Principalmente, las preocupaciones se relacionan con el corto plazo que se les brinda a las plataformas para dar de baja los contenidos –en algunos casos 24 horas– y, a su vez, con el hecho de que no resulta evidente en todos los escenarios definir si un contenido es ilegal al punto de ameritar su inmediata remoción (Schulz, 2018).

Si bien este trabajo aborda en otro apartado la responsabilidad de los intermediarios específicamente, resultaba relevante en este punto hacer mención a una ley ya vigente que se involucra con el problema que se aborda aquí. Retomando el punto anterior, se puede observar que hay una tensión que está presente constantemente: las “excepciones” versus la menor afectación posible a la libertad de expresión.

A razón de todo lo hasta aquí planteado, y con el objetivo de garantizar que esos límites -o excepciones- a la libertad de expresión sean “admisibles”, la Organización de los Estados Americanos (OEA) ha establecido tres condiciones que las restricciones deberían cumplir y que se conocen como el “test tripartito”:

(1) la limitación debe haber sido definida en forma precisa y clara a través de una ley formal y material, (2) la limitación debe estar orientada al logro de objetivos imperiosos autorizados por la Convención Americana, y (3) la limitación debe ser necesaria en una sociedad democrática para el logro de los fines imperiosos que se buscan; estrictamente proporcionada a la finalidad perseguida; e idónea para lograr el objetivo imperioso que pretende lograr (OEA, 2010).

Aparecen aquí varios factores que son importantes para el desarrollo de este trabajo. En primer lugar, que toda restricción debe cumplir con el principio básico de legalidad, es decir que tiene que basarse en una ley vigente y que aplique en la jurisdicción correspondiente al hecho. A su vez, está presente la idea de la proporcionalidad en las medidas con el fin de lograr la menor afectación posible a la circulación de los discursos. Por último, la necesidad social imperiosa que debe motivar esa restricción, es decir, que esté claro cuál es el fin legítimo que se busca proteger.

En un comunicado de prensa¹⁸ de febrero de 2021 firmado por la Comisión Interamericana de Derechos Humanos (CIDH) y su Relatoría Especial para la Libertad de Expresión (RELE), se describe al momento histórico actual como un “punto de inflexión” en el cual el debate público se ve deteriorado. Además, se alerta acerca de la posibilidad de que exista un “traslado de la violencia online a espacios físicos con capacidad cierta de daño”.

A su vez, en esta misma declaración se plantea que los Derechos Humanos deben ser protegidos de igual manera en entornos digitales y consideran a Internet como una “plataforma y un medio para el ejercicio de los derechos humanos”. En línea con la relación entre la libertad de expresión y las posibles restricciones vinculadas a los discursos de odio y la incitación a la violencia, se puede encontrar el siguiente fragmento: “Mientras el cruce de argumentos y la exposición pública de los desacuerdos enriquecen el debate, la violencia y los discursos que alientan al odio erosionan el sistema democrático” (OEA, 2021).

En el mismo documento se encuentra otro tema esencial para este trabajo: los bloqueos masivos, la remoción de contenidos y la suspensión de cuentas de usuarios. En este sentido, la CIDH y la RELE entienden que estas medidas son “severas” y que deberían ser evaluadas sobre la base de estándares internacionales. Además, subrayan la gravedad de esta situación particularmente cuando se trata de “información de interés público o funcionarios públicos en ejercicio de sus funciones” (OEA, 2021).

Este punto permite reflexionar nuevamente acerca del test tripartito. Si se toma específicamente el tercer criterio, se aclara que las limitaciones deben ser necesarias “en una sociedad democrática para el logro de los fines imperiosos que se buscan” (OEA, 2010).

A modo de aproximación al análisis del caso de este trabajo, cabe plantear la pregunta de hasta qué punto puede ser benéfico para los objetivos de una sociedad democrática el hecho de cortar una vía de comunicación de Donald Trump para con los ciudadanos y, a su vez, restringirle a éstos el acceso a la palabra de su presidente.

¹⁸ Comunicado de Prensa No. 026/21 de la Comisión Interamericana de Derechos Humanos y su Relatoría Especial para la Libertad de Expresión (RELE). Disponible en: <http://www.oas.org/es/cidh/jsForm/?File=/es/cidh/prensa/comunicados/2021/026.asp> (última consulta: 27/11/2021)

En este sentido, se puede plantear también una reflexión acerca de la proporcionalidad de las medidas: es al menos una materia opinable el hecho de que una suspensión permanente sea una respuesta proporcional –o bien “idónea”– ante los hechos, sobre todo si se tiene en cuenta el potencial daño en términos democráticos que puede tener el bloqueo al presidente en ejercicio de los Estados Unidos.

3.1.1. Libertad de expresión y funcionarios públicos durante procesos electorales

En esta sección, se brindan algunas precisiones adicionales con respecto a las implicaciones que puede tener en términos de libertad de expresión el hecho de estar lidiando con funcionarios públicos, especialmente durante un proceso electoral. Estas aclaraciones resultan relevantes a la luz del caso de análisis de esta tesina que, como ya se ha adelantado, refiere a Donald Trump en el momento en que buscaba su reelección como presidente de Estados Unidos.

La Organización de Estados Americanos (OEA) plantea que los funcionarios públicos tienen ciertos “deberes generales” a los que están sujetos en cuanto al ejercicio de su libertad de expresión. El primer punto que se plantea es la necesidad que tienen de pronunciarse ante ciertos asuntos de interés público. Más aún, en ese documento se establece que, en esos escenarios, “el ejercicio de su libertad de expresión no es solamente un derecho, sino un deber” (OEA, 2010).

Este planteo resulta interesante a la luz de los acontecimientos del caso de análisis, ya que el hecho de que las plataformas de redes sociales decidieran suspender permanentemente el acceso a las cuentas del entonces presidente significó, cuanto menos, una dificultad para el ejercicio de este “deber”.

Por otro lado, la OEA también establece la necesidad de que los funcionarios públicos fundamenten sus opiniones en hechos constatables para evitar que los ciudadanos obtengan versiones manipuladas de las situaciones. Además, se hace mención a la importancia de que estas expresiones no puedan incurrir en violaciones a los Derechos Humanos o bien en interferencias para con las autoridades judiciales u otros actores sociales que quieran expresar sus opiniones en el debate público.

En este punto, sí podrían cuestionarse muchos de los dichos de Trump ya que, por ejemplo, no se ha encontrado hasta el momento evidencia de que estuviera ocurriendo un fraude electoral como el que él denunciaba abiertamente. Ante este escenario, cabe plantear el siguiente interrogante: ¿es legítimo limitar el acceso a que un funcionario se exprese ante un asunto de interés público debido a que, justamente, sus expresiones previas no han sido fundamentadas en hechos?

La cuestión se complejiza aún más cuando se trata de procesos electorales. La Relatoría Especial para la Libertad de Expresión (RELE) de la Comisión Interamericana de Derechos Humanos (CIDH) publicó en 2020 una serie de recomendaciones sobre libertad de expresión y elecciones en la era digital dirigidas tanto a los Estados como a “actores no estatales”.

Entre otros puntos, la RELE hace foco en que los Estados deben asegurarse de que cualquier restricción a la libertad de expresión que se realice durante un proceso electoral cumpla estrictamente con el test tripartito citado en el apartado anterior. Dentro de este punto, se resalta la recomendación de no establecer leyes “generales o ambiguas” acerca de la desinformación que tiendan a prohibir la difusión de “falsedades” o “información no objetiva” (RELE, 2020). Así, se intenta que las restricciones sean proporcionales y ajustadas a una necesidad específica.

Al respecto, se destaca una última recomendación que plantea:

Los actores estatales deben asegurarse de que los medios de comunicación gocen de un acceso sin impedimentos a fuentes de información oficial y a los candidatos a ocupar cargos públicos, y de que no encuentren obstáculos indebidos que afecten su posibilidad de difundir dicha información e ideas (RELE, 2020).

Se podría considerar que las redes sociales se han transformado efectivamente en una fuente de información oficial, por lo que entrarían dentro de estas categorías protegidas especialmente durante procesos electorales. Lógicamente, si las cuentas son suspendidas como en el caso de referencia, el acceso a estas fuentes se vería obstaculizado, o bien directamente imposibilitado.

Finalmente, por el lado de los actores no estatales, la RELE plantea que las plataformas deberían adoptar las medidas necesarias para que sus usuarios tengan acceso a información plural y de diversas perspectivas. Para esto, se hace un énfasis especial en

la necesidad de que los algoritmos no “obstaculicen indebidamente” el acceso a contenidos que tengan relación con el proceso electoral. Por otra parte, también se las insta a “abordar la desinformación, la información errónea o manipulada intencionalmente”.

Se presenta aquí un escenario en el que las plataformas deben, por una parte, asegurarles a sus usuarios un acceso a información plural, pero a la vez estar alerta para actuar ante la aparición de información errónea o manipulada. Aquí, vuelve a tornarse relevante la necesidad de aplicar el espíritu del test tripartito para asegurarse de que cualquier restricción que se tome cumpla estrictamente con los principios allí establecidos.

3.2. Legislación y jurisprudencia relevante en Estados Unidos

Como se ha mencionado anteriormente, los Estados Unidos no han ratificado la Convención Americana sobre Derechos Humanos. Por lo tanto, se realiza un recorrido por la legislación y la jurisprudencia sobre la libertad de expresión en dicho país con el fin de alcanzar un mayor entendimiento de las regulaciones que allí rigen.

En este punto, cabe una aclaración con respecto a la importancia que tiene la jurisprudencia para el sistema legal estadounidense. Natalia Bernal Cano (2013) explica que en los sistemas anglosajones –como el de Estados Unidos– la jurisprudencia aparece como resultado de la repetición de los casos concretos y que “obligan normalmente a los jueces a conservar las reglas y los criterios de decisión”. En cambio, en los sistemas romano-germánicos existen normas jurídicas que son de naturaleza legislativa y “cuya aplicación es garantizada por medio de la jurisprudencia” (Bernal Cano, 2013:375).

Las diferencias entre ambos sistemas legales son mucho más complejas, pero una descripción detallada de ellos escapa a los objetivos de este trabajo. La diferenciación citada se considera suficiente para comprender la mayor relevancia que la jurisprudencia tiene como fuente de derecho en Estados Unidos, en contraposición con otros sistemas legales.

Así, un repaso por la legislación y la jurisprudencia vigente en los Estados Unidos sobre la materia es importante para este trabajo debido a que, por supuesto, el caso de análisis

sucedió en ese territorio. Pero a la vez, es fundamental ya que muchas de las plataformas en general, y Twitter en particular, son empresas radicadas en los Estados Unidos y que, por lo tanto, se rigen también por dicha legislación.

En primer lugar, se puede citar la sección 230 de la *Communications Decency Act* (CDA) –traducida como la Ley de Decencia en las Comunicaciones– de 1996, que establece que “ningún proveedor o usuario de un servicio informático interactivo deberá ser tratado o considerado como el editor o portavoz de cualquier información proporcionada por otro proveedor de contenido de información”¹⁹.

En este breve párrafo se explica claramente la manera en que los intermediarios ven limitada su responsabilidad ante el contenido publicado por sus usuarios. Como describe Gillespie (2017), esto implica que los intermediarios no pueden ser responsabilizados por los discursos de sus usuarios, ya que solamente estarían actuando como proveedores de acceso a un servicio, y no como editores del contenido en un sentido legal.

A su vez, otra cuestión interesante que la sección 230 plantea tiene que ver con que estas empresas tampoco perderían su protección legal en el caso de que sí decidieran voluntariamente monitorear y tomar decisiones sobre el contenido publicado en sus plataformas, ya que esta acción no las convierte legalmente en “editoras” de contenido. De este modo, las empresas tienen asegurada la protección, tanto en un escenario en el que decidieran no intervenir ante un contenido cuestionado, como en el caso de que sí escojan, por ejemplo, remover contenidos (Gillespie, 2017).

Otra parte importante de la legislación estadounidense en la materia se encuentra nada menos que en la Primera Enmienda²⁰ a la Constitución de los Estados Unidos. En ella, se establece concretamente que el Congreso no podrá sancionar ninguna ley que limite la libertad de expresión ni de prensa²¹. Se puede observar, entonces, un posicionamiento firme de jerarquía constitucional que evita la posibilidad de establecer restricciones

¹⁹ Artículo redactado por la Electronic Frontier Foundation (EFF) acerca de la Sección 230 de la Ley de Decencia en las Comunicaciones. Disponible en: <https://www.eff.org/es/issues/cda230> (traducción propia, última consulta: 28/11/2021).

²⁰ Las enmiendas a la Constitución de los Estados Unidos son agregados al texto original de 1787 que cuentan con jerarquía constitucional y son fuente fundamental para las decisiones judiciales. Las 10 primeras enmiendas son conocidas como la “Carta de Derechos” y fueron promulgadas en 1791.

²¹ Texto completo de la Constitución de los Estados Unidos. Disponible en: https://www.senate.gov/civics/constitution_item/constitution.html/ (traducción propia, última consulta: 28/11/2021)

legales a la libertad de expresión. Sin embargo, y es relevante hacer la aclaración, esto no significa una prohibición para la posibilidad de regular en la materia.

Cabe aclarar en este punto algunas cuestiones respecto del alcance específico de esta Primera Enmienda. En el texto de Loreti (2021) se recuperan algunas sentencias relevantes de la Corte Suprema de los Estados Unidos en los cuales se establece, por ejemplo, que “los actores no estatales no están sujetos a las obligaciones de la Primera Enmienda”. En este sentido, la jurisprudencia establece que para realizar reclamos a particulares en cuestiones relativas a la libertad de expresión, “es necesario demostrar que existe una acción del Estado involucrada” (Loreti, 2021:7). Comprendiendo esto, se podría pensar que una empresa como Twitter no tendría por qué regirse por esta normativa.

Sin embargo, el hecho de que los actores particulares estén sujetos o no a lo prescripto por la Primera Enmienda es objeto de un debate aún no saldado entre los expertos en libertad de expresión. Por ejemplo, este interrogante se presenta ante aquellos escenarios en los que “las políticas privadas de moderación de contenido de las plataformas violan la Primera Enmienda y regularlas para que no excluyan a las personas del debate público sería una forma de garantizar derechos” (Loreti, 2021:16).

A modo de complemento, se recuperan a continuación otros casos trascendentales de jurisprudencia estadounidense en relación con la libertad de expresión. Cronológicamente, el primero ocurre en 1942 y se conoce como Chaplinsky vs. New Hampshire. Chaplinsky había sido condenado por la justicia local por agredir verbalmente de manera enfática a un oficial de policía de su ciudad.

Basándose en la ya mencionada Primera Enmienda a la Constitución, Chaplinsky apeló esta decisión, pero la Corte terminó manteniendo su condena. El argumento que utilizaron los jueces fue que, si bien la mayoría de las expresiones están protegidas, no se puede reclamar la misma protección para discursos “lascivos, obscenos, profanos, calumniosos, e insultantes”. Específicamente, se hace referencia en este fallo a las *fighting words* que se podrían traducir como aquellas palabras que buscan iniciar una

pelea. O bien, como plantea el juez, aquellas palabras que “por su propia naturaleza cometen un daño o tienden a incitar una inminente ruptura de la paz”²².

Otro de los casos más reconocidos es el de *Brandenburg vs. Ohio* de 1969. Allí, la Corte Suprema resolvió abolir la pena de Clarence Brandenburg quien había sido condenado en instancias anteriores por incitar a la violencia durante una manifestación del Ku Klux Klan (grupo del que él formaba parte).

No se ahondará en los pormenores legales y jurídicos de esta resolución, pero es importante resaltar el argumento principal que el máximo tribunal estadounidense utilizó para tomar su decisión. Los jueces definieron que la expresión de cualquier tipo de discurso, incluso aquellos que contengan expresiones violentas u ofensivas, está protegida por la Primera Enmienda. La única excepción a esto se da cuando se considere que estos discursos puedan incitar “una acción ilegal inminente”²³. Este estándar es el que prima actualmente en la justicia estadounidense y se corresponde con una mirada más dirigida hacia la protección de la libertad de expresión.

Otra decisión importante de la Corte estadounidense en relación con lo planteado se dio en el caso *Hess vs. Indiana* de 1973 en el que se reafirmó el estándar recién mencionado. Gregory Hess había sido condenado por gritar, mientras se reprimía una manifestación, que “volverían a tomar la calle”. Luego, esa decisión fue revocada por la Corte Suprema ya que los dichos de Hess, al tratarse de una expresión hacia el futuro, no implicaban el peligro de una “acción ilegal inminente”²⁴.

Así, si bien los fallos más recientes no implican un descrédito de los anteriores, se puede pensar que hay una tendencia a hacer cada vez más específicas las limitaciones a la libertad de expresión. A su vez, es notorio el hecho de que estas restricciones parecen ceñirse exclusivamente a aquellas ocasiones en las que existe el peligro inminente de una acción y no a la expresión en sí, por más dañina que ésta pueda ser.

²² Artículo completo del Centro de Estudios acerca de la Primera Enmienda de la Universidad Estatal de Tennessee (MTSU) acerca del caso *Chaplinsky*. Disponible en: <https://www.mtsu.edu/first-amendment/article/293/chaplinsky-v-new-hampshire> (traducción propia, última consulta: 28/11/2021)

²³ Artículo completo del Centro de Estudios acerca de la Primera Enmienda de la Universidad Estatal de Tennessee (MTSU) acerca del caso *Brandenburg*. Disponible en: <https://www.mtsu.edu/first-amendment/article/189/brandenburg-v-ohio> (traducción propia, última consulta: 28/11/2021).

²⁴ Artículo completo del Centro de Estudios acerca de la Primera Enmienda de la Universidad Estatal de Tennessee (MTSU) acerca del caso *Hess*. Disponible en: <https://www.mtsu.edu/first-amendment/article/461/hess-v-indiana> (traducción propia, última consulta: 28/11/2021).

3.3. Plataformas como posicionamiento discursivo: ¿empresas tecnológicas o empresas de medios?

Como ya se ha enunciado, el objetivo central de este trabajo es reflexionar acerca de los riesgos para el ejercicio de la libertad de expresión que implica la autorregulación de las plataformas de redes sociales.

Para dar comienzo a este apartado, se retoman algunos trabajos que abordan la dimensión económica y regulatoria de estas plataformas y que analizan también a estos actores. En este sentido, justamente el término “plataforma” brinda una base interesante para poder comenzar el desarrollo.

Tarleton Gillespie (2010) entiende que empresas como Facebook, Twitter y Google, entre otras, se califican a sí mismas como “plataformas” como parte de un posicionamiento discursivo intencional que conlleva una serie de consecuencias.

Uno de los significados que este término admite (y que le transfiere de algún modo a estas empresas) es el de actuar como una “superficie” que eleva a los individuos a un mismo nivel y los empodera para que puedan expresarse. En este sentido, se relacionaría con la clásica retórica del “potencial democratizador de Internet” que, como advirtió Lessig (1998:6), es en sí misma problemática ya que “mientras nosotros celebramos la libertad ‘inherente’ de la red, la arquitectura de la red está cambiando debajo de nosotros mismos”.

Esta idea de las plataformas –y de Internet en general– como un espacio plural y democrático ya estaba puesta en discusión hace más de 20 años. En su ya citado texto, Lessig (1998) hablaba de una transformación desde una “arquitectura de libertad” hacia una “arquitectura de control” que se daba con la intervención de diversos actores:

Está mutando ya sin ningún tipo de intervención gubernamental, a pesar de lo cual el gobierno está rápidamente viendo cómo podría intervenir para acelerar la mutación. Y donde el gobierno está ya interviniendo, lo está haciendo de una forma diseñada para cambiar esta misma arquitectura en una arquitectura de control para hacerla, como ya he dicho, más regulable (Lessig, 1998:6).

Gracias al concepto de “plataformas”, entonces, las empresas evitan posicionarse (desde lo discursivo) como creadoras o moderadoras del contenido, sino tan solo como

facilitadoras para la generación y publicación de estos contenidos por parte de los usuarios (Gillespie, 2010:7-8).

Otra acepción que permite la palabra “plataforma” es aquella que estas empresas utilizan para limitar su responsabilidad. Gillespie (2010:11) estudia el caso de Google y YouTube para plantear el rol de estas compañías:

(...) en ciertas ocasiones trabajan para minimizar su rol y parecer meros intermediarios, limitando así su responsabilidad por la actividad de sus usuarios. (...) En este esfuerzo de habitar el ‘medio’, son recompensados por facilitar la expresión, pero sin responsabilidad por sus excesos. Google y YouTube desarrollaron el término ‘plataforma’ como parte de su estrategia legislativa²⁷.

En los próximos apartados se avanza sobre la responsabilidad legal de estas plataformas en su rol de intermediarios. Sin embargo, resulta relevante presentar este tema desde el aspecto de un posicionamiento discursivo que toman estas compañías, ya que también es parte de su estrategia para dar esos debates regulatorios sobre el tipo de empresas que son.

En relación con esto, se desarrolla la idea de que los únicos trabajos de curación o moderación de contenidos que se dan en estas “plataformas” ocurren por la acción de “algoritmos”, es decir, sin intervención humana directa. A este respecto, Napoli y Caplan (2019:11) consideran que “esta afirmación/percepción acerca de la falta de intervención editorial humana directa es, en varios sentidos, fundamental para la lógica de concebir a estas plataformas como compañías tecnológicas en lugar de medios de comunicación”.

Sin embargo, como se esbozó en el apartado anterior, las empresas tienen la posibilidad de intervenir también sobre el contenido publicado en sus plataformas sin perder su protección legal. De hecho, Gillespie (2017) entiende que las plataformas, aprovechando esta situación, están interviniendo de manera profunda y sistemática, aunque la ley no los obligue a hacerlo.

Esta situación, según el autor, se puede explicar por varios factores. En primer lugar, por una cuestión económica, ya que los anunciantes pueden sentirse más cómodos para invertir en una plataforma que no esté repleta de, por ejemplo, contenidos

²⁷ Traducción propia

discriminatorios o violentos que puedan quedar asociados a sus marcas. A su vez, este ambiente es también beneficioso para no espantar a nuevos usuarios que están ingresando a la plataforma. Finalmente, el autor también entiende que para estas empresas existe también un “compromiso” moral o bien una “obligación pública” de crear una comunidad “saludable” que permita desplegar la creatividad de sus usuarios (Gillespie, 2017).

Teniendo en cuenta estos planteos, es relevante también pensar acerca de cuáles son los usos que tienen las plataformas en cuestión. El *Digital News Report*²⁸ (Reporte de Noticias Digitales) del Instituto Reuters publicado en 2021 arroja varios datos interesantes a tener en cuenta. A nivel internacional, se afirma que las redes sociales continúan consolidándose como una fuente para consumir noticias, sobre todo en las generaciones más jóvenes.

Sin embargo, Estados Unidos es destacado como un caso particular. En primer lugar, por el hecho de que es el país cuyos encuestados demostraron una menor confianza en las noticias a nivel general (29%) y también un menor interés acerca de acceder a noticias (11%).

Este fenómeno, según el reporte, se puede explicar precisamente por los hechos acontecidos durante las elecciones presidenciales de 2020 y los hechos de violencia posteriores en el Capitolio. Esta situación habría generado, entre los seguidores de Donald Trump, un menor interés y una mayor desconfianza para con los medios de comunicación.

Si se deja de lado la cuestión política por un instante, se observa que otro factor importante tuvo que ver con el Coronavirus. En este caso, la pandemia pareciera haber repercutido en un crecimiento significativo de la brecha de credibilidad entre los medios masivos y las redes sociales a favor de los primeros. En otras palabras, los usuarios encontraron en los medios tradicionales –y sobre todo en los más grandes– información más confiable que en las redes sociales.

A pesar de este declive en el consumo, el interés y la confianza que se dio en todos los niveles, un 42% de los usuarios encuestados afirma utilizar las redes sociales como

²⁸ Sitio oficial del Reporte de Noticias Digitales generado por el Instituto Reuters. Disponible en: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021> (última consulta: 28/11/2021)

medio de acceso a las noticias, tan solo por detrás de la televisión (52%) y bastante por delante de la prensa impresa (16%). Si se considera a los medios en línea en su totalidad (incluyendo a las redes sociales), el reporte arroja que un 66% de los estadounidenses encuestados se informa a través de Internet.

Si se toman en cuenta estos resultados, es al menos cuestionable la idea de que estas empresas no tienen relación con los medios de comunicación. Esta distinción justamente es relevante ya que, si se consideran a estas plataformas únicamente como empresas tecnológicas, se estarían dejando de lado “las crecientes dimensiones políticas y culturales de su operación, que se hacen más pronunciadas a medida que estas plataformas se vuelven guardianes [gatekeepers] centrales de noticias e información en el ecosistema de medios contemporáneo” (Napoli y Caplan, 2019:21).

Esta diferencia en cuanto a la clasificación de las empresas también trae a la luz una pregunta fundamental que guía el recorrido del trabajo: ¿tiene Twitter (y otras plataformas) la propiedad y la autoridad editorial sobre el contenido que circula en su plataforma? (Napoli y Caplan, 2019).

3.4. Responsabilidad de los intermediarios: definiciones y principios regulatorios

La responsabilidad que les cabe (o no) a los intermediarios por el contenido que publican y/o comparten sus usuarios, como ya se mencionó, es un tema central para los objetivos de este trabajo.

El concepto mismo de “intermediarios” es amplio y merece una mayor precisión. Según los Principios de Manila Sobre la Responsabilidad de los Intermediarios (un documento redactado en conjunto en el 2015 por una coalición de organizaciones de la sociedad civil) se entiende que “toda comunicación a través de Internet es facilitada por intermediarios, como los proveedores de acceso a Internet, las redes sociales y los motores de búsqueda”³⁰.

³⁰ Texto completo de los Principios De Manila Sobre Responsabilidad De Los Intermediarios. Disponible en: https://www.eff.org/files/2015/06/23/manila_principles_1.0_es.pdf (última consulta: 29/11/2021).

En este sentido, es importante destacar que ya desde hace años se observa que esta “mediación” en Internet se da a través de un grupo concentrado de empresas. Van Dijck (2016) describe las adquisiciones y acuerdos comerciales a los que han llegado Google, Facebook, Apple y Amazon y que han desembocado en la conformación de “grandes cadenas de plataformas –microsistemas verticalmente integrados en virtud de su propiedad, conformación accionaria o por asociación estratégica– que dominan hoy el ecosistema de los medios conectivos” (Van Dijck, 2016:168).

Si se suma a Microsoft a la lista de estas cuatro grandes empresas se terminará de conformar lo que se conoce popularmente como “GAFAM”, un acrónimo conformado por las iniciales de cada una de las empresas que ostentan esta posición dominante. En cualquier caso, resulta relevante marcar una vez más que la concepción de Internet como un espacio puramente democrático y descentralizado está puesta en cuestión a la vista de estos fenómenos de concentración.

Antes de continuar, corresponde aclarar una vez más que este trabajo va a enfocarse principalmente en Twitter al momento del análisis. Si bien no es una de las empresas que ingresa entre las más grandes en el análisis de concentración recién mencionado, su posición es de relevancia para este trabajo, principalmente por el uso destacado que le dan algunos actores centrales de la política, en este caso encarnados en Donald Trump.

Sin embargo, a la hora de trabajar sobre la responsabilidad de los intermediarios, algunas definiciones que abarcan a todo este conjunto de actores puede ser útil para comprender los conceptos que se buscan desarrollar ya que, como se mencionó anteriormente, las redes sociales (y por lo tanto Twitter) entran dentro de esta categoría.

El primero de los principios de Manila establece que “los intermediarios deberían estar protegidos por ley de la responsabilidad por contenido de terceros”. A su vez, en los incisos de ese primer principio se especifica que:

- a. Cualquier disposición que rija la responsabilidad de intermediarios debe ser establecida por leyes, que deben ser precisas, claras y accesibles.
- b. Los intermediarios deberían ser inmunes a la responsabilidad por contenido de terceros en circunstancias en las que no han estado involucrados en la modificación de dicho contenido.
- c. Los intermediarios no deben ser responsabilizados por no restringir contenido lícito.

d. Los intermediarios nunca deben ser responsabilizados conforme los lineamientos del modelo de responsabilidad objetiva por alojar contenido ilícito de terceros, ni deben ser obligados a monitorear contenido proactivamente como parte de un régimen de responsabilidad de intermediarios (Principios de Manila, 2015).

Complementando la posición anterior, desde el punto de vista legal se pueden encontrar diversos regímenes posibles de legislación que afectan a estos actores. Hiram Meléndez Juarbe (2012) establece la existencia de cuatro formas principales en las que el derecho considera la responsabilidad de los intermediarios.

El primer sistema es el de “inmunidad absoluta” para los intermediarios, caso que ocurre por ejemplo en Estados Unidos bajo la protección de la sección 230 de la *Communications Decency Act* (Ley de Decencia en las Comunicaciones), que ya fue descrita en el apartado de legislación estadounidense. El segundo sistema es uno de “responsabilidad estricta” (u objetiva) que obligaría al intermediario a “reparar el daño que se produzca por usuarios del servicio”, independientemente de su diligencia o conocimiento del ilícito. Un tercer sistema que plantea el autor es el de “responsabilidad subjetiva”. En este caso, el intermediario puede estar expuesto solamente a una obligación civil en caso de, por ejemplo, incumplir con su deber de actuar una vez conocido el acto ilícito. Para el cuarto escenario se puede hablar de una “inmunidad condicionada”, que implica el cumplimiento de diversas obligaciones como la bajada de contenido tras alguna notificación que puede o no provenir de un tribunal (Meléndez, 2012: 115-117).

De este modo, se puede observar como la situación de la protección legal que obtienen estos actores depende del régimen legal bajo el cual se encuentren. Al respecto de esto, Álvarez Ugarte y Rabinovich (2013) remarcan la importancia de que cualquiera sea el sistema que impere debería estar “en sintonía con lo que disponen los estándares internacionales sobre libertad de expresión”.

En este sentido, y al respecto de los regímenes anteriormente descriptos, se pueden encontrar nuevamente posicionamientos en los principios de Manila. En el segundo de éstos aparece la idea de que “los intermediarios no deben ser obligados a restringir contenidos a menos que una orden emitida por una autoridad judicial independiente e imparcial haya determinado que el contenido en cuestión es ilícito”. Además, se aclara que “cualquier responsabilidad impuesta sobre un intermediario debe ser proporcional y

directamente correlacionada al comportamiento ilícito del intermediario de no cumplir adecuadamente la orden de restricción del contenido” (Principios de Manila, 2015).

Así, en estos planteos se ve una forma de pensar esta responsabilidad de los intermediarios que se aleja, al menos, de la idea de responsabilidad objetiva. Y esto se relaciona principalmente con el principio de “proporcionalidad” que entiende que cualquier restricción a la libertad de expresión debe ser “idónea, necesaria y estrictamente proporcional al fin legítimo que las justifica (...) interfiriendo en la menor medida posible con el ejercicio legítimo del derecho a la libertad de pensamiento y expresión”. (Álvarez Ugarte y Rabinovich, 2013:111).

Aquí se observa una vez más la esencia del test tripartito que se mencionó en el primer apartado del marco teórico. En primer lugar, porque se apunta la necesidad de cumplir con el principio de legalidad, ya que cada limitación deba ser definida claramente a través de una ley. Por otro lado, ya que aquellas limitaciones que efectivamente ocurran deben ser estrictamente necesarias y proporcionales al fin último que se persigue a través de ellas.

Siguiendo esta línea, Meléndez (2012:116) plantea que el sistema de responsabilidad objetiva plantea un riesgo muy alto de “sobreprotección”, es decir, que en ciertos casos extremos, el costo posible que tendrían que afrontar por la responsabilidad superaría los beneficios de incurrir en la actividad, y esto induciría al intermediario a no proveer ciertos servicios. Llevando esta definición a un ejemplo práctico se podría pensar que, en estos casos, a una plataforma de redes sociales le resultaría menos problemático (y más económico) dar de baja la cuenta de un usuario particular antes que tener que afrontar un proceso judicial por el contenido que éste publique.

El asunto de la responsabilidad de los intermediarios es complejo y se podría ampliar mucho más en algunas de sus definiciones. Sin embargo, el panorama planteado hasta aquí permite tener una perspectiva sobre la importancia que esta discusión tiene y, sobre todo, acerca de los riesgos que pueden conllevar ciertas medidas o decisiones en materia de libertad de expresión.

Con el recorrido desarrollado hasta aquí, pueden plantearse algunos elementos que guían el recorrido de los próximos apartados. Por una parte, el hecho de que ante la masiva cantidad de discursos que circulan en las redes sociales, nace un reclamo –tanto

de parte de usuarios como de las plataformas– para filtrar algunos discursos: los discursos de odio y las noticias falsas (o *fake news*) son ejemplos claros de esto.

Los interrogantes surgen al momento de considerar cómo y quién debería realizar esas moderaciones de contenido. ¿Es ese el rol de los intermediarios? ¿Qué podría ocurrir si limitan contenidos “de más”? ¿Es necesaria una orden judicial para actuar? ¿El contenido que se limita es solamente el considerado “ilegítimo”? ¿La moderación es únicamente un rol de las empresas? Y finalmente, si se considera que las plataformas actúan a nivel global y que existen diferencias entre los distintos países -tanto a nivel legal como cultural- que se ponen en juego, ¿cuál es el rol de los Estados ante esta problemática?

3.5. ¿Qué alcance tienen los términos de servicio?

En los apartados anteriores, se pudo observar que el sector privado tiene un rol importante a la hora de la regulación de los discursos en redes sociales. En este sentido, las plataformas tienen sus posturas establecidas en cuanto a las cuestiones que se vienen desarrollando.

En el caso de Twitter, se puede leer en sus términos de servicio que: “Todo el Contenido es responsabilidad única de la persona que lo produce. No monitorizamos ni controlamos el Contenido publicado por medio de los Servicios, y no podemos hacernos responsables de dicho Contenido”³².

Este planteo parecería condecirse con algunos de los estándares que se mencionaron anteriormente en cuanto a la limitación de su responsabilidad como intermediarios. Sin embargo, estos términos de servicio esconden también algunos otros puntos que vale la pena analizar.

En un documento del año 2020, el Observatorio Latinoamericano de Regulación de Medios y Convergencia (OBSERVACOM) plantea que existe en las plataformas una tendencia a intervenir cada vez más en los contenidos de sus usuarios a través de sus términos de servicio y sus políticas de moderación de contenido. Estos escenarios son

³² Texto completo de los Términos de Servicio de Twitter: <https://twitter.com/es/tos> (última consulta: 29/11/2021)

problemáticos en la medida en que se constituyen como “formas de regulación privada de espacios públicos que hoy son vitales para la deliberación democrática y el ejercicio de derechos fundamentales” (OBSERVACOM, 2020:5).

Nuevamente en este punto se podría recuperar el planteo de Lessig (1998) acerca del proceso de mutación del ciberespacio desde una “arquitectura de libertad” hacia una “arquitectura de control” en la que la intervención, en este caso de las plataformas, es cada vez mayor.

Sobre esta aparente contradicción, se pueden encontrar algunas ideas para intentar pensar en el alcance que tienen (o deberían tener) estos términos de servicio. En el citado documento del OBSERVACOM aparece el siguiente fragmento:

Ninguna plataforma de contenidos debería tener responsabilidad legal por contenidos generados por terceros, en tanto no intervengan modificando o editando esos contenidos. (...) Las grandes plataformas de contenidos deberían ser consideradas responsables por acciones propias si censuran contenidos de interés público protegidos por el derecho a la libertad de expresión y por sus acciones en la promoción activa de expresiones que pudieran afectar derechos de terceros (OBSERVACOM, 2020:19).

De este modo, existe una tendencia marcada a no responsabilizar a los intermediarios por los contenidos que generan terceros. Sin embargo, aquí aparece una arista interesante para este trabajo que implicaría una responsabilidad para estos actores en el caso de cercenar el acceso a algunos contenidos.

Justamente este rol que estas plataformas comienzan a tomar con mayor énfasis es también motivo de análisis y renovadas discusiones. Napoli y Caplan plantean que las plataformas como Twitter tienen un importante margen para participar en la decisión editorial acerca del contenido que allí se produce “sin activar la categorización como editor y, por lo tanto, limitando las responsabilidades legales de gran alcance que atañen a esa figura” (Napoli y Caplan 2019:18).

A propósito de esto, Napoli y Caplan (2019) también explican la manera en que estas plataformas sí toman decisiones sobre el contenido. La “vigilancia” que realizan se da solamente de manera voluntaria, en una modalidad que no difiere mucho de lo que ha realizado históricamente la televisión.

Por otro lado, se puede afirmar que las plataformas también realizan tareas de edición en cuanto a que el contenido que se le muestra a cada usuario está personalizado y adaptado a sus gustos y consumos frecuentes, con el objetivo de que permanezcan más tiempo navegando. Esta intención es también similar a la de los medios tradicionales, pero mucho más eficaz gracias a la cantidad y profundidad de datos acerca de los usuarios con los que cuentan. De este modo, queda aún más desacreditada la noción de que los algoritmos operen de manera completamente neutral y objetiva (Napoli y Caplan, 2019).

En el texto de Becerra y Waisborg (2021) se puede observar como los autores califican a las decisiones editoriales de las plataformas como “problemáticas” debido a su discrecionalidad “en tanto que no son aplicadas estandarizada o sistemáticamente en todos los casos” y a su opacidad, en la medida en que “no suelen abundar en detalles sobre las razones de sus decisiones o intentan involucrar a varios actores públicos en la definición de áreas problemáticas de discurso y respuestas necesarias”.

Lo hasta aquí planteado dispara nuevamente la pregunta de por qué Twitter y el resto de las plataformas de redes sociales tomaron una decisión tan contundente con respecto al caso de Donald Trump teniendo en cuenta que, en principio, no deberían afrontar responsabilidad por los dichos del ex presidente.

Nuevamente, se puede recuperar aquí lo planteado por Gillespie (2017): las plataformas están interviniendo en un grado mayor de lo estrictamente necesario, aprovechando la protección que el sistema regulatorio (principalmente la Sección 230 de la Ley de Decencia en las Comunicaciones) les otorga.

Al respecto de esto, es importante aclarar que no se está planteando aquí un escenario en el que no existan regulaciones y en el que todo discurso pueda circular sin que sus emisores deban afrontar responsabilidades ulteriores por ellos. En este sentido, se puede traer a colación lo planteado por Becerra y Waisborg: “Considerar que cualquier expresión es conducente a buen discurso democrático es ingenuo, idealista (...) y representa una opción facilista y principista que elude los problemas gruesos de una realidad mucho más compleja” (Becerra y Waisborg, 2021:16). Por lo tanto, existe la necesidad de regulación, pero ésta no debería ser unilateralmente establecida por las plataformas, sino que requiere de un enfoque más democrático.

De este modo, en el citado documento del OBSERVACOM (2020) se encuentran algunos estándares y recomendaciones específicamente con respecto a los términos de servicio (TdS) de las plataformas. En primer lugar, haciendo referencia a la mencionada cuestión de la transparencia y la proporcionalidad:

Todas las plataformas de contenidos deberían establecer e implementar TdS que sean transparentes, claros, accesibles y apegados a las normas y principios internacionales en materia de derechos humanos (...) En cualquier situación, los términos de servicio no deberían contener cláusulas abusivas, injustas o desproporcionadas (OBSERVACOM, 2020:11).

A su vez, aparece la necesidad de que estos términos sean concretos en su redacción para evitar de este modo una “interpretación y aplicación arbitraria por parte de las plataformas” (OBSERVACOM, 2020:12), fundamentalmente en aquellos casos en los que las restricciones implican un límite a la libertad de expresión.

Se pueden encontrar reflexiones similares en el texto de Loreti (2021). En el citado documento, se resumen una serie de informes que incluyen recomendaciones para las empresas privadas en materia de Derechos Humanos.

En primer lugar, se encuentra la referencia a los Principios Rectores sobre las Empresas y los Derechos Humanos³³ redactados por las Naciones Unidas, en los cuales se explicita que las empresas privadas, como órganos de la sociedad, deben cumplir con las leyes vigentes en su contexto y con el respeto a los Derechos Humanos. A su vez, deben adoptar políticas internas que apunten en esta dirección y ser capaces de explicar a la sociedad cuando hayan ocurrido violaciones y las medidas que toman para hacer frente a éstas (Loreti, 2021).

Por otro lado, se hace referencia también a un Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión dirigido al Consejo de Derechos Humanos de las Naciones Unidas³⁴. Allí, aparecen algunas

³³ Organización de las Naciones Unidas (ONU, 2011). “Principios Rectores sobre las Empresas y los Derechos Humanos de las Naciones Unidas” HR/PUB/11/04. Disponible en: https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_sp.pdf (última consulta: 2/1/2022).

³⁴ Organización de las Naciones Unidas (ONU, 6 de abril de 2018). “Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión”. Consejo de Derechos Humanos, ONU. A/HRC/38/35. Disponible en: <https://www.refworld.org/es/pdfid/5c6b33774.pdf> última consulta: 2/1/2022).

recomendaciones específicas con respecto a la moderación de contenido y la suspensión de cuentas de las empresas con base en algunas pautas centrales.

En primer lugar, se encuentra la necesidad de una mayor transparencia en las notificaciones cuando las plataformas decidan dar de baja algún contenido para que el usuario pueda conocer el enfoque que adopta la empresa. Un segundo punto se relaciona con la importancia de que existan procesos claros de apelaciones y reparaciones ante este tipo de decisiones, ya que se considera que, actualmente, estas posibilidades son limitadas e insuficientes (Loreti, 2021).

Finalmente, del citado informe también se resalta la recomendación de que las empresas incorporen a sus términos de servicio algunos puntos para asegurarse de que las medidas relacionadas con el contenido estén guiadas por los mismos principios de legalidad, necesidad y legitimidad que aplican cuando se trata la regulación de la expresión de parte de los Estados (Loreti, 2021).

Así, para concluir este apartado, se puede pensar que “resultan aplicables a los actores no estatales, en particular a las empresas de internet, las reglas de aplicación del test tripartito imprescindible para la adopción de restricciones al ejercicio de la libertad de expresión” (Loreti, 2021:24). La legalidad, la transparencia, la proporcionalidad y el hecho de evitar interpretaciones arbitrarias parecen emerger como cuestiones centrales a la hora de pensar en posibilidades de transformación para los términos de servicio existentes en las plataformas de redes sociales.

3.6. El acceso a la información en un escenario de concentración

Una idea que se mencionó en el apartado de las plataformas como posicionamiento discursivo es la tendencia de las compañías de redes sociales, buscadores y otros intermediarios a considerarse como empresas tecnológicas en lugar de empresas de medios. Esta decisión tiene consecuencias desde lo legal y tiende a limitar la responsabilidad de estas empresas por el contenido que comparten sus usuarios.

Sin embargo, otra ramificación importante de este posicionamiento se observa desde el punto de vista económico. Existe una percepción por parte de las comunidades de inversores que tiende a ver un mayor potencial de ingresos en el sector de la tecnología

que en el de los medios de comunicación (Napoli y Caplan, 2019). Por lo tanto, se ve que no solo se trata de una decisión que apunta a un mayor resguardo desde el punto de vista legal, sino que también otorga beneficios financieros.

Más aún, como ya han trabajado diversos autores, existe una marcada concentración en esta industria. Baladron (2018) plantea la manera en que, si bien la naturaleza de Internet permite que muchos actores participen en la producción de contenidos, “los análisis empíricos permiten observar estrategias que tienden a la monopolización a través de la creación artificial de escasez tal como sucede con las industrias de los medios y la comunicación que le precedieron” (Baladron, 2018:4).

Esta concentración trae como consecuencia que sean estas pocas corporaciones las que

han centralizado y concentran el poder de administrar la circulación, intercambio o búsqueda de informaciones y opiniones y ejercen este poder de manera arbitraria y sin ningún mecanismo de *accountability* a las instituciones garantes de derechos. Un riesgo que va creciendo a medida que dicho puñado de empresas refuerza su posición dominante en el mercado y desarrolla tecnologías opacas para la gobernanza de la información (OBSERVACOM, 2020:6).

Aquí se hace presente nuevamente el fenómeno de “GAFAM” (Google, Facebook, Apple, Amazon y Microsoft), las empresas más importantes del sector que representan esta posición dominante a la que se hace referencia y que ponen –al menos– en duda la posibilidad de pensar a Internet como un espacio de acceso libre y democrático a la información.

Se pueden aportar a este punto algunos datos adicionales retomando el *Digital News Report*³⁵ (Reporte de Noticias Digitales) del Instituto Reuters. Según dicho informe, en Norteamérica un 35% de los encuestados afirma haber utilizado Facebook como fuente de acceso a noticias, en comparación con un 25% de Youtube, un 12% de Twitter y un 11% de Instagram. De este modo, se puede pensar que una limitación de contenidos que se tome en cualquiera de estas plataformas tendrá consecuencias importantes a la hora del acceso a la información para estas personas.

Así, los usuarios parecieran encontrarse con un escenario en el que las corporaciones pueden definir unilateralmente a través de sus términos de servicio ciertas condiciones

³⁵ Sitio oficial del Reporte de Noticias Digitales generado por el Instituto Reuters. Disponible en: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021> (última consulta: 28/11/2021)

que tienen consecuencias en el debate público y, al mismo tiempo, cuentan con un resguardo legal para tomar esas decisiones.

Desde el punto de vista ya citado de las organizaciones civiles y de cierta parte del debate académico, se encuentra sobre este punto una necesidad de transformación a través de un control democrático. Existe una coincidencia en cuanto a que este control no implica la responsabilidad objetiva de los intermediarios sobre el contenido que sus usuarios comparten, pero es más compleja la cuestión a la hora de pensar en las medidas a tomar ante esta problemática.

Una primera aproximación permite comprender que una de las demandas más importantes es la de transparencia en los mecanismos de toma de decisiones (bajas de contenidos, cierres de cuentas, limitaciones de acceso, entre otras). Otra idea que se comparte tiene que ver con la aplicación del test tripartito y la búsqueda de que toda limitación de contenidos cumpla con el principio de legalidad, que sea proporcional, y también idónea para proteger los fines legítimos a los que se apunta. Sin embargo, como se ha visto, se trata de un problema complejo que no tiene una resolución sencilla y que cuenta con diversas posturas de parte de los actores involucrados.

En este punto concluye el marco teórico de esta tesina. En este capítulo, se han repasado cuestiones que son centrales para la temática y relevantes para el análisis posterior. En el primer apartado, se trabajó sobre definiciones de los discursos de odio, su vínculo con la libertad de expresión y la visión de los Órganos Internacionales de Derechos Humanos. Luego, se recopilaron los aspectos más relevantes de la legislación y jurisprudencia de los Estados Unidos, país en donde tuvo lugar el caso de análisis.

El tercer apartado tuvo como tema central a las plataformas y la utilización que las empresas hacen de este concepto como posicionamiento discursivo y estratégico. Más adelante, se trabajó sobre el concepto de intermediarios y las responsabilidades que pueden corresponderles por el contenido que los usuarios publican. En quinto lugar, se analizó la cuestión de los términos de servicio y la regulación privada de parte de las plataformas. Finalmente, el último apartado se centró en la problemática del acceso a la información en un escenario de concentración.

A partir de lo trabajado, se han lanzado diversos interrogantes que se abordan en el capítulo de análisis del caso de estudio elegido. Allí, desde las perspectivas teóricas

propuestas y los conceptos desarrollados, se analiza la relación entre Twitter y Donald Trump, los acontecimientos y las acciones que tuvieron como corolario el bloqueo de la cuenta del ex presidente.

Es parte fundamental de este proceso no perder de vista la relevancia del actor involucrado en el caso y su investidura: Donald Trump, quien era el presidente de los Estados Unidos y estaba en busca de su reelección. Como se ha planteado, el hecho de que la persona en cuestión sea el funcionario público más importante de una de las potencias más grandes del mundo dispara consecuencias imposibles de soslayar a la hora de pensar en los límites legítimos para la libertad de expresión en un sistema democrático, más aún si se considera que el país se encontraba en pleno proceso electoral.

4. El ataque al Capitolio y el “bloqueo” a Donald Trump en Twitter: un estudio de caso

4.1. Presentación y descripción del caso

En este cuarto capítulo se realiza un análisis de caso que incluye un recorrido histórico por los antecedentes de la relación entre Donald Trump y Twitter -respecto del uso político y comunicacional de esta red social y sobre los intentos de regulación que impulsó el ex presidente- y los posicionamientos de ambos actores durante los hechos de violencia en el Capitolio, que concluyeron con el bloqueo permanente de la cuenta personal de Trump. Para ello, se elaboró un corpus con una selección de *tweets* y declaraciones públicas de los actores involucrados, los términos de servicio de la red social Twitter y los cambios introducidos durante el periodo de análisis, para dar cuenta de las regulaciones de las publicaciones y de la cuenta de Trump por parte de la red social.

El caso a analizar se trata de la suspensión permanente de la cuenta de Twitter de Donald Trump (@realDonaldTrump) que ocurrió el 8 de enero de 2021. A continuación, se realiza una recopilación cronológica de los sucesos que precedieron a la mencionada suspensión. Estos antecedentes y su contexto son relevantes para el análisis del caso, ya que dan cuenta de ciertas características de los actores involucrados y de que se trataba de un hecho de interés público, de un mandatario público y respecto de las elecciones presidenciales, cuestión central del funcionamiento del sistema democrático.

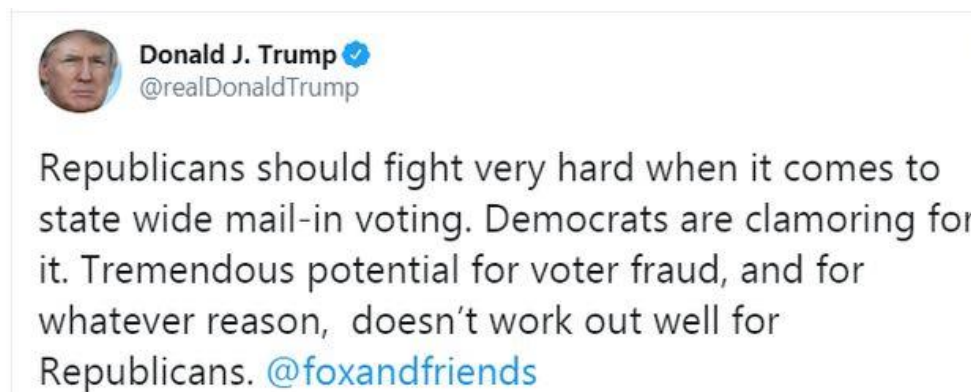
Las elecciones presidenciales de los Estados Unidos tuvieron lugar el martes 3 de noviembre de 2020 y enfrentaron por las dos fuerzas políticas mayoritarias de ese país a Donald Trump (Partido Republicano) -quien estaba en ejercicio de la presidencia desde 2016 y se presentaba para su reelección- y a Joe Biden (Partido Demócrata).

En unas elecciones atravesadas por la pandemia del COVID-19, en diversos Estados se comenzó a expandir la alternativa de que los ciudadanos pudieran emitir su voto por correo (*mail-in*) para evitar aglomeraciones en los centros de votación.

En los Estados Unidos, cada Estado tiene la potestad de establecer (incluso en este contexto) cuál es su metodología de votación para las elecciones presidenciales. Previamente⁴³, la mayoría de los Estados permitía la elección por correo sólo para aquellos electores que tuvieran una justificación para no hacerlo personalmente. Sin embargo, en muchos de ellos el COVID-19 fue agregado como uno de los motivos válidos para emitir votos por correo⁴⁴.

Si bien hubo algunos que no incluyeron esta modalidad⁴⁵, lo más destacable en este punto fue que en el caso de los llamados Estados “pendulares” (*swing states*), es decir, aquellos que no tienen una histórica preferencia por alguno de los dos partidos políticos mayoritarios, todos optaron por esta nueva forma de elección⁴⁶.

Desde el principio, Donald Trump se manifestó en contra de este método bajo el argumento de que existía una importante posibilidad de fraude y que, a su vez, era un mecanismo que “no funciona bien” para los Republicanos.



Fuente: @realDonaldTrump en Twitter⁴⁷

⁴³ Rabinowitz K. y Mayes, B. (25/09/2020). “Al menos 84% de los votantes estadounidenses podrán enviar sus votos por correo en otoño”. *The Washington Post*. Disponible en: <https://www.washingtonpost.com/graphics/2020/politics/vote-by-mail-states/> (en inglés - última consulta: 12/12/2021).

⁴⁴ Davis, E., O’Shea, J. y Grimes, N. (25/08/2020). “Cómo votar por correo en cada uno de los 50 Estados”. *US News*. Disponible en: <https://www.usnews.com/news/elections/articles/how-to-vote-by-mail-in-all-50-states-in-the-2020-election> (en inglés - última consulta: 12/12/2021).

⁴⁵ Texas, Mississippi, Louisiana, Tennessee e Indiana, cinco Estados en los que Trump había vencido en las elecciones 2016.

⁴⁶ Brownstein, R. (11/04/2020). “Los Estados más importantes para 2020 ya tienen voto por correo”. *The Atlantic*. Disponible en: <https://www.theatlantic.com/politics/archive/2020/04/voting-mail-2020-race-between-biden-and-trump/609799/> (en inglés - última consulta: 16/12/2021).

⁴⁷ Traducción propia del tweet completo: “Los republicanos deberían luchar muy duramente cuando se trata de la votación por correo en todo el estado. Los demócratas están clamando por él. Tremendo potencial para el fraude electoral y, por la razón que sea, no funciona bien para los Republicanos”.

De acuerdo con un estudio del Pew Research Center⁴⁸, un 46% de los votantes totales emitió su voto por correo, pero tan solo el 33% de los votantes de Trump indicó haber utilizado esta modalidad por lo que se podría inferir que se verificaron los planteos del candidato republicano, ya que coincidieron con la conducta posterior de sus votantes.

En un nuevo estudio, el Pew Research Center⁴⁹ resaltó que en la elección presidencial del 2020 casi dos de cada tres ciudadanos estadounidenses habilitados para participar de los comicios (66,2%) emitieron su voto (por cualquiera de los métodos posibles). Esto significó un aumento de 7 puntos porcentuales con respecto a la elección anterior de 2016, así como también el porcentaje de votantes más alto desde, al menos, 1980. Cabe aclarar en este punto que el voto en los Estados Unidos no es obligatorio, lo cual puede resaltar de alguna manera el notable crecimiento en la participación para la elección de 2020.

No se ahonda en las especificidades de la votación por correo ni en los argumentos que se utilizaron para defenderlo o criticarlo ya que escapa a los objetivos de este trabajo. Sin embargo, resulta relevante marcar la disconformidad de Trump con este asunto ya que se trata de un punto sobre el cual Twitter toma intervención posteriormente.

Debido a esta modalidad de votación, este particular proceso eleccionario contó también con importantes demoras en el conteo de votos a lo largo de distintos Estados⁵⁰, lo cual contribuyó a alimentar los rumores y las denuncias de fraude por parte de Trump.

El entonces presidente canalizó sus denuncias a través de sus redes sociales en diversas ocasiones y, a raíz de esto, Twitter recurrió también muchas veces a cuestionar sus publicaciones y a etiquetarlas como contenido engañoso. Se profundiza sobre este aspecto en el apartado de análisis del corpus.

Un dato curioso y que cabe remarcar aquí es que las redes sociales también fueron un factor que dio lugar a la polémica durante las elecciones de 2016 que proclamaron a

⁴⁸ Pew Research Center (2010). “La experiencia de votación en 2020”. Disponible en: <https://www.pewresearch.org/politics/2020/11/20/the-voting-experience-in-2020/> (en inglés – última consulta: 19/12/2021).

⁴⁹ Desilver, D. (28/01/2021) “La participación se disparó en 2020 con casi dos tercios de los votantes habilitados en Estados Unidos que emitieron sus votos para presidente”. Disponible en: <https://www.pewresearch.org/fact-tank/2021/01/28/turnout-soared-in-2020-as-nearly-two-thirds-of-eligible-u-s-voters-cast-ballots-for-president/> (en inglés - última consulta: 16/12/2021).

⁵⁰ Redacción BBC (4/11/2020). “Elecciones en Estados Unidos 2020: ¿cuándo conoceremos el resultado?”. BBC. Disponible en: <https://www.bbc.com/news/election-us-2020-54096399> (en inglés - última consulta: 16/12/2021).

Donald Trump como presidente, cuando venció por un estrecho margen a la candidata demócrata Hillary Clinton. Por un lado, debido a las múltiples denuncias de *fake news*⁵¹ que habrían jugado a favor del candidato republicano durante esa campaña electoral. Por otra parte, por los escándalos desatados a través de la supuesta injerencia de servicios de inteligencia rusos⁵² que habrían intervenido con acciones de desinformación a través de Facebook para favorecer la llegada de Trump al poder. Finalmente, también se puede mencionar el recordado caso de Cambridge Analytica⁵³ en el cual se acusó a la empresa de utilizar información privada de usuarios de Facebook con el fin de influenciar su decisión en favor de Donald Trump.

Si bien muchas de estas situaciones están ligadas a la plataforma de Facebook, y más allá de que no tengan vinculaciones específicas con el tema que se analiza a continuación, es relevante mencionar la manera en que las redes sociales ya venían siendo un factor de importante influencia en los procesos electorales de los Estados Unidos.

Finalmente, respecto de las elecciones de 2020, Joe Biden fue proclamado como presidente de los Estados Unidos por sobre Donald Trump⁵⁴. La reticencia de Trump a convalidar el resultado electoral y las acciones legales⁵⁵ iniciadas para revertir el resultado de los comicios llevó a un período de transición entre mandatos particularmente tensa. Durante este tiempo, Trump realizó publicaciones en su cuenta de Twitter en las que aseguró, por ejemplo, que las elecciones habían sido arregladas, que él había sido el ganador o que no aceptaría el resultado.

⁵¹ Blake, A. (3/4/2018). “Un nuevo estudio sugiere que las ‘fake news’ pueden haberle ganado la elección de 2016 a Donald Trump”. *The Washington Post*. Disponible en:

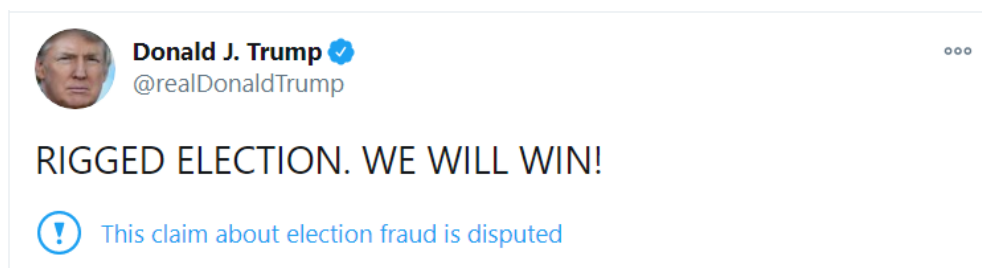
<https://www.washingtonpost.com/news/the-fix/wp/2018/04/03/a-new-study-suggests-fake-news-might-have-won-donald-trump-the-2016-election/> (en inglés – última consulta: 2/1/2022)

⁵² Mars, A. (24/2/2018). “¿Cómo influyó la desinformación en las elecciones presidenciales?”. *El País*. Disponible en: https://elpais.com/internacional/2018/02/24/estados_unidos/1519484655_450950.html (última consulta: 2/1/2022).

⁵³ Redacción BBC Mundo (20/3/2018). “5 claves para entender el escándalo de Cambridge Analytica que hizo que Facebook perdiera US\$37.000 millones en un día”. *BBC*. Disponible en: <https://www.bbc.com/mundo/noticias-43472797> (última consulta: 2/1/2022).

⁵⁴ Lemire, J., Miller, Z. y Weissert, W. (8/11/2020). “Biden derrota a Trump por la Casa Blanca y dice que es ‘tiempo de curar’”. *AP*. Disponible en: <https://apnews.com/article/joe-biden-wins-white-house-ap-fd58df73aa677acb74fce2a69adb71f9> (en inglés - última consulta: 16/12/2021).

⁵⁵ Sherman, M. (4/11/2020). “Trump demanda en tres Estados, preparando el terreno para impugnar los resultados”. *AP*. Disponible en: <https://apnews.com/article/donald-trump-seeks-voting-stop-25762f69b27dfbccc4fd8077fb5f5dc91> (en inglés - última consulta: 16/12/2021).



Fuente: @realDonaldTrump en Twitter⁵⁶

Algunas semanas después, más precisamente el 6 de enero de 2021, fue la fecha designada para que las dos cámaras del Congreso de los Estados Unidos se reunieran en el Capitolio para el conteo final de los votos del Colegio Electoral y para la posterior proclamación del nuevo presidente.

Ese mismo día, Trump brindó un discurso⁵⁷ ante sus seguidores en el que los instaba a “luchar” y a dirigirse al Capitolio en donde se estaba realizando el conteo para “convencer” a los congresistas de interrumpirlo. Estas palabras fueron tomadas por sus opositores como una “incitación a la insurrección” que luego terminaría en un juicio político⁵⁸.

La “insurrección” a la que se hace referencia tuvo lugar ese mismo día. Un grupo de manifestantes que defendían las posiciones de Trump empezó a protestar en las inmediaciones del Capitolio. Esas protestas comenzaron a escalar en su intensidad mientras las fuerzas de seguridad iniciaban acciones represivas. Finalmente, un importante número de manifestantes irrumpió en el Capitolio⁵⁹ y generó destrozos en el interior del edificio, en un evento inédito que recorrió los medios de comunicación de todo el mundo⁶⁰.

⁵⁶ Traducción propia del tweet completo: “Elección arreglada. ¡Vamos a ganar!”. Traducción propia de la etiqueta agregada por Twitter: “Esta afirmación acerca de fraude electoral es controversial”.

⁵⁷ Naylor, B. (10/2/2021). “Lectura del discurso de Trump del 6 de enero, una parte clave del juicio político”. *National Public Radio*. Disponible en: <https://www.npr.org/2021/02/10/966396848/read-trumps-jan-6-speech-a-key-part-of-impeachment-trial> (en inglés - última consulta: 16/12/2021).

⁵⁸ Fandos, N. (13/01/2021). “Trump a juicio político por incitar la insurrección”. *The New York Times*. Disponible en: <https://www.nytimes.com/2021/01/13/us/politics/trump-impeached.html> (en inglés - última consulta: 16/12/2021).

⁵⁹ Wise, L., Lucey, C. y Restuccia, A. (6/1/2021). “‘Los manifestantes están en el edificio’. Dentro del Capitolio tomado por una multitud pro-Trump”. *The Wall Street Journal*. Disponible en: <https://www.wsj.com/articles/the-protesters-are-in-the-building-inside-the-capitol-stormed-by-a-pro-trump-mob-11609984654> (en inglés - última consulta: 16/12/2021).

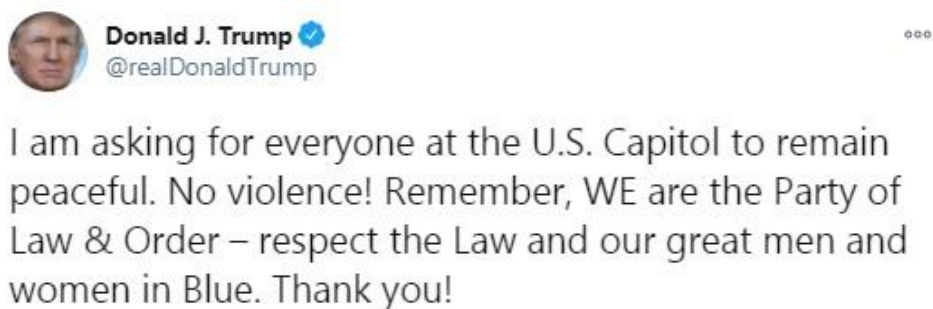
⁶⁰ Allegretti, A. (7/1/2021). “En imágenes: caos por manifestantes pro-Trump que rompen la barricada y atraviesan la seguridad del Capitolio de los Estados Unidos”. *Sky News*. Disponible en:

La actividad de Trump en Twitter durante ese día también fue intensa. En primer lugar, realizó distintas publicaciones insistiendo en el “fraude” ocurrido en las elecciones y pidiéndoles a los representantes de su partido en el Capitolio que detuvieran el conteo de los votos. Especialmente, el pedido era para su vicepresidente, Mike Pence, quien ocupaba también la Presidencia del Senado.



Fuente: @realDonaldTrump en Twitter⁶¹

Luego de iniciados los serios incidentes en el Capitolio, Trump realizó algunas publicaciones llamando a sus seguidores a mantener la paz y, principalmente, a no enfrentarse con las fuerzas de seguridad. Sin embargo, los incidentes se extendieron durante esa jornada y dejaron un saldo de cuatro personas fallecidas.



Fuente: @realDonaldTrump en Twitter⁶²

<https://news.sky.com/story/in-pictures-turmoil-as-pro-trump-protesters-break-through-us-capitol-security-12181006> (en inglés - última consulta: 16/12/2021).

⁶¹ Traducción propia del tweet completo: “Los Estados quieren corregir sus votos, los cuales ahora saben que estaban basados en irregularidades y fraude, además de un proceso corrupto que nunca recibió aprobación legislativa. Todo lo que Mike Pence tiene que hacer es mandarlos de vuelta a los Estados, Y GANAMOS. ¡Hazlo Mike, este es el momento de tener un coraje extremo!”. Traducción propia de la etiqueta agregada por Twitter: “Esta afirmación acerca de fraude electoral es controversial”.

Con respecto a los objetivos de análisis de este trabajo, cabe destacar en este punto que ese mismo día ocurrió por primera vez el pedido de eliminación⁶³ de contenido publicado por la cuenta oficial de Trump, específicamente de tres de sus *tweets*.

Finalmente, el último de los momentos centrales de esta recopilación llegó dos días después. El 8 de enero Trump realizó sus dos últimas publicaciones en Twitter. En la primera de ellas, les manifestaba a sus votantes que no permitiría que les faltaran el respeto o que los trataran de manera injusta. En el segundo y último de esos *tweets*, hizo referencia específicamente a su decisión de no asistir a la asunción presidencial de Biden que tendría lugar el 20 de enero.

Ese mismo día, y tomando en consideración también estas dos publicaciones, Twitter decidió suspender de manera permanente la cuenta de Donald Trump⁶⁴. Nuevamente, cabe aclarar aquí que Twitter no fue la única plataforma⁶⁵ que tomó decisiones de este calibre, pero sí se trata de la que se analiza en este trabajo debido al recorte escogido.

En este punto finaliza la primera recopilación de los hechos sucedidos en torno al caso de análisis. En los apartados posteriores se presenta un análisis de las publicaciones de Trump y los argumentos esgrimidos públicamente por las autoridades de Twitter para llevar adelante estas decisiones, desde la perspectiva del derecho a la comunicación y la economía política de la comunicación.

⁶² Traducción propia del *tweet* completo: “Le pido a todos en el Capitolio de los Estados Unidos que mantengan la calma. ¡Sin violencia! Recuerden, NOSOTROS somos el Partido de la Ley y el Orden – respeten la Ley y a nuestros grandes hombres y mujeres de Azul. ¡Gracias!”.

⁶³ Tweet de una cuenta oficial de Twitter (@TwitterSafety) (6/1/2021). Traducción propia del *tweet* completo: “Como resultado de la situación violenta sin precedentes y en curso en Washington, D.C., hemos pedido la remoción de tres Tweets de @realDonaldTrump que se publicaron hoy por violaciones graves y repetidas de nuestra política de integridad cívica”. Disponible en: <https://twitter.com/TwitterSafety/status/1346970430062485505> (última consulta: 16/12/2021).

⁶⁴ Artículo del blog oficial de Twitter (8/1/2021). “Suspensión permanente de @realDonaldTrump”. Publicado el 8/1/2021. Disponible en: https://blog.twitter.com/en_us/topics/company/2020/suspension (en inglés - última consulta: 16/12/2021).

⁶⁵ Byers, D. (14/1/2021). “Cómo Facebook y Twitter decidieron bajar las cuentas de Trump”. *NBC News*. Disponible en: <https://www.nbcnews.com/tech/tech-news/how-facebook-twitter-decided-take-down-trump-s-accounts-n1254317> (en inglés - última consulta: 16/12/2021).

4.2. La relación Twitter - Donald Trump y el sistema de “etiquetado”

Para comprender de mejor manera los hechos previamente descritos, resulta preciso recuperar ciertos antecedentes. En primera instancia, se encuentra la figura de Donald Trump y la importancia que le brindaba a su cuenta personal de Twitter desde antes de asumir la presidencia de los Estados Unidos.

En 2017, cuando Trump comenzó a cursar su mandato, heredó también la cuenta oficial de la Presidencia de los Estados Unidos (@POTUS, por sus siglas en inglés). Sin embargo, manifestó desde el principio que no dejaría de utilizar su cuenta personal⁶⁶ que ya en ese momento contaba con algo más de 20 millones de seguidores.

Si bien @POTUS, una cuenta de una investidura más formal, tuvo actividad durante la presidencia de Trump, el ex presidente mismo se encargó de aclarar que no sería él quien utilizaría dicho usuario. En cambio, sí se mantuvo con gran actividad en su cuenta personal. En una entrevista con el periódico *The Times* de Londres de enero de 2017⁶⁷, Trump señaló que sufría una cobertura “deshonesta” de parte de la prensa y que, por este motivo, utilizaba Twitter ya que le servía como una herramienta para que los medios publicaran rápidamente lo que él compartía.

Si bien en este trabajo no se hace foco en la relación de Trump con las “noticias falsas” ya que escapa a los objetivos propuestos, cabe destacar que el ex presidente denunció frecuentemente a los medios de comunicación por publicar *fake news*⁶⁸. De aquí se puede comprender también la importancia que tenía para Trump la utilización de sus redes sociales personales en términos de comunicación política.

Sin embargo, la relación de Trump con Twitter comenzó a atravesar importantes turbulencias. Uno de los motivos centrales fue la decisión de la plataforma de estrenar una herramienta frente a lo que consideraban “información engañosa”.

⁶⁶Matyszczyk, C. (16/01/2017). “Donald Trump dice que mantendrá su cuenta personal de Twitter”. *CNET*. Disponible en: <https://www.cnet.com/news/donald-trump-will-keep-his-personal-twitter-account-report-says/> (en inglés - última consulta: 17/12/2021).

⁶⁷Redacción The Times (16/01/2017). “Transcripción completa de la entrevista con Donald Trump”. *The Times*. Disponible en: <https://www.thetimes.co.uk/article/full-transcript-of-interview-with-donald-trump-5d39sr09d?region=global> (en inglés - última consulta: 17/12/2021).

⁶⁸Redacción BBC Mundo (18/01/2018). “Los polémicos ‘premios a las noticias falsas’ otorgados por Donald Trump a algunos medios de comunicación críticos de su gobierno”. Disponible en: <https://www.bbc.com/mundo/noticias-internacional-42727725> (última consulta: 17/12/2021).

En un artículo⁶⁹ publicado en su blog oficial en febrero de 2020, Twitter manifestó que los usuarios deben tener la posibilidad de acceder a “información confiable” en su plataforma. Por este motivo, en el mismo artículo se anunció que entraría en vigencia un sistema de etiquetado (“*label*”) para aquellos *tweets* que contaran con contenido que hubiera sido “alterado o fabricado para engañar”.

A su vez, la plataforma puso a disposición de los usuarios tres criterios en forma de preguntas a través de las cuales analizaría los contenidos publicados: a) ¿El contenido fue alterado o fabricado significativamente de manera engañosa?; b) ¿El contenido fue compartido de manera engañosa?; c) ¿Es posible que el contenido tenga un impacto en la seguridad pública o que cause un daño serio?

Is the content significantly and deceptively altered or fabricated?	Is the content shared in a deceptive manner?	Is the content likely to impact public safety or cause serious harm?	
			Content may be labeled.
			Content may be labeled.
			Content is likely to be labeled, or may be removed.*
			Content is likely to be labeled.
			Content is likely to be removed.

Fuente: Centro de Ayuda de Twitter

Como se ve en la imagen precedente, dependiendo de los criterios con los que cumpla el contenido en cuestión, la medida a tomar por parte de la plataforma puede variar: desde la mera posibilidad de ser etiquetado [*may be labeled*] hasta una alta probabilidad de ser eliminado [*is likely to be removed*].

En este punto, se observa la intervención de Twitter respecto de los contenidos publicados por terceros en su red social. Por lo tanto, se pone en evidencia el hecho de que las plataformas, que se posicionan discursivamente como tales para evitar ser catalogadas como editoras o moderadoras de contenido (Gillespie, 2010), en la práctica terminan tomando decisiones sobre el contenido y realizando prácticas de vigilancia “voluntaria” (Napoli y Caplan, 2019).

⁶⁹ Artículo del blog oficial de Twitter (4/2/2020). “Construyendo reglas en público: nuestro acercamiento ante contenidos multimedia falsos y alterados”. Disponible en: https://blog.twitter.com/en_us/topics/company/2020/new-approach-to-synthetic-and-manipulated-media (en inglés -última consulta: 17/12/2021).

Así, aprovechando la protección legal con la que cuentan, las plataformas están interviniendo en el contenido sistemáticamente, más allá de no estar obligadas a hacerlo (Gillespie, 2017). De este modo, las empresas terminan funcionando como “gatekeepers” o guardianas del contenido (Napoli y Caplan, 2019) a través de las decisiones corporativas que toman.

Otro aspecto relevante tiene que ver con los riesgos que implica para la libertad de expresión el hecho de que se trate de formas de regulación privada, fundadas en los términos de servicio de las empresas y en sus políticas de moderación, las que establecen las reglas de juego para intervenir en los contenidos de los usuarios (OBSERVACOM, 2020). A su vez, estas restricciones afrontan el riesgo de no cumplir con los criterios de legalidad y proporcionalidad establecidos por estándares internacionales, tal como es el caso del mencionado “test tripartito”.

Es importante resaltar que Twitter también compartió una descripción más detallada de los criterios previamente mencionados. Con respecto al contenido “alterado”, se observa que el foco está puesto en aquellos que cuenten con una edición del material a tal punto que llevara a una modificación “sustancial” de su composición. Otro punto que se tiene en cuenta aquí es el agregado o la remoción de “información visual o auditiva” al material original (fragmentos de video, doblaje de audio o modificación de subtítulos). Finalmente, también se considera bajo este criterio la “representación o fabricación simulada” de una persona.

Un aspecto importante que se menciona hacia el final de este punto es que la determinación de si un contenido fue alterado de manera significativa se realizará a través de “tecnología propia o reportes obtenidos gracias a una asociación con terceras partes”, sin una profundización mayor acerca de este punto.

Nuevamente, se hace presente la problemática de la opacidad en la toma de estas decisiones y la necesidad de que los términos de servicio de las plataformas sean “transparentes, claros, accesibles y apegados a las normas y principios internacionales en materia de derechos humanos” (OBSERVACOM, 2020:11). Esta cuestión se torna aún más relevante cuando las regulaciones privadas están apuntando a posibles limitaciones en el ejercicio de la libertad de expresión de los usuarios, entrando así en conflicto con regulaciones locales y estándares internacionales vigentes.

El segundo criterio se relaciona con la manera de compartir el contenido en cuestión. Aquí la atención está puesta sobre el contexto en el que cierto material se difunde y la posibilidad de que lleve a “confusiones” o bien la existencia de “intentos deliberados para engañar a las personas”.

Finalmente, el tercer criterio apunta a los contenidos que pueden generar “daños” específicos tales como: amenazas a la seguridad física de una persona o grupo; riesgo de violencia en masa o disturbios; amenazas a la privacidad o a la posibilidad de una persona de expresarse libremente o participar de eventos cívicos.

Una vez más, la cuestión de que estos criterios sean unilateralmente establecidos resulta cuestionable, sobre todo cuando se trata de conceptos esquivos o sin una definición concreta como el “riesgo”. Cabe preguntarse aquí quién y de qué manera realizaría estas “evaluaciones” del riesgo, ya que la conclusión de ese análisis determinará una posible restricción a la libertad de expresión.

Se puede retomar aquí también la recomendación de que los términos de servicio deberían apuntar a ser específicos y concretos en su redacción, para evitar darle el espacio a las plataformas para realizar interpretaciones y aplicaciones arbitrarias que escapen al principio de proporcionalidad (OBSERVACOM, 2020).

Ante el escenario de un contenido que, según estos criterios, pudiera ser “etiquetado” las medidas específicas que la plataforma plantea tomar son:

- A) Aplicar una etiqueta al contenido
- B) Mostrar una advertencia a las personas antes de que compartan dicho contenido
- C) Reducir la visibilidad de ese contenido
- D) Proveer un enlace a explicaciones o información adicional
- E) Desactivar las respuestas y las opciones de “me gusta” o *retweet*⁷⁰

En cambio, aquellos contenidos que cumplan con los tres criterios están directamente sujetos a ser eliminados y, a su vez, aquellas cuentas que reincidieran en este comportamiento pueden ser “suspendidas permanentemente”⁷¹.

⁷⁰ Un concepto propio de Twitter que implica compartir la publicación de una persona con los propios seguidores del usuario que “retuitea”, ampliando así su alcance.

En este último aspecto se observan algunos de los peligros de la aplicación de la censura privada. Tal como se explicita en las medidas propuestas, un usuario que incurra en una conducta que no cumpla con los criterios de la plataforma será pasible de sanciones y recibirá la explícita advertencia de una suspensión permanente de su cuenta. Por lo tanto, la posibilidad de una autocensura ante este escenario resulta evidente. Es por esto que, nuevamente, se necesita que aquellas medidas restrictivas que se tomen –si es que se toman– sean específicas para cumplir con un fin legítimo y concreto y no habiliten la posibilidad de incurrir en actos de censura previa. Otro punto a tener en cuenta, y que no suele estar presente en la regulación de contenido que realizan las plataformas, es el debido proceso, la posibilidad de “defensa” de la persona afectada y la inclusión de reglas y procedimientos claros.

En un segundo artículo⁷² del blog oficial de Twitter, que hace referencia a los mismos objetivos aunque con el foco puesto en la información engañosa relacionada con el COVID-19, se señala que las “etiquetas” colocadas en el contenido tendrían un enlace hacia “páginas curadas por Twitter”⁷³ o bien “fuentes externas de confianza” que contengan información adicional sobre las afirmaciones que se realicen en el *tweet* cuestionado.

De este modo, se observa que no solamente es el criterio de la plataforma el que prima a la hora de tener en cuenta qué contenidos son cuestionables o no, sino que también se convierte en la autoridad que determina dónde puede obtenerse la información fidedigna.

El trabajo de análisis de este caso continúa en el próximo apartado con los contenidos específicos publicados por Trump que fueron etiquetados por la plataforma y en las consecuencias regulatorias que estos hechos pudieron generar.

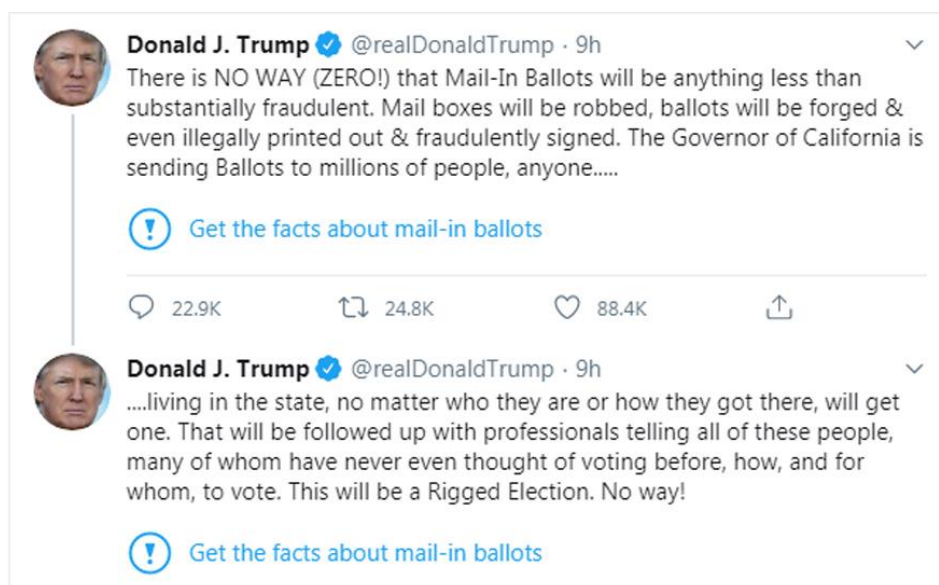
⁷¹ Artículo del Centro de Ayuda de Twitter (s.f.). “Política relativa a los medios contenidos multimedia falsos y alterados”. Disponible en: <https://help.twitter.com/es/rules-and-policies/manipulated-media> (última consulta: 17/12/2021).

⁷² Artículo del blog oficial de Twitter (11/05/2020). “Actualizando nuestro enfoque ante la información engañosa”. Disponible en: https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information (en inglés - última consulta: 17/12/2021).

⁷³ Artículo del Centro de Ayuda de Twitter (s.f.). “Directrices y principios de los Momentos de Twitter”. Disponible en: <https://help.twitter.com/es/rules-and-policies/twitter-moments-guidelines-and-principles> (última consulta: 17/12/2021).

4.3. Contenidos etiquetados por Twitter y propuesta de regulación

Respecto del caso de análisis, Donald Trump publicó el 26 de mayo del 2020 una serie de *tweets* anticipando que los votos emitidos por correo serían “sustancialmente fraudulentos” y acusando al gobernador del estado de California de enviarle boletas a cualquier persona que viviera en dicho estado “sin importar quiénes son o cómo llegaron allí”.



Fuente: @realDonaldTrump en Twitter⁷⁴

Estos *tweets*, tal como se observa en las imágenes anteriores, fueron los primeros en recibir una advertencia por parte de la plataforma que se puede traducir como “obtén información acerca de las elecciones por correo”. Esa etiqueta contenía un enlace⁷⁵ hacia una de las mencionadas “páginas curadas” por Twitter –conocidas como

⁷⁴ Traducción propia de los *tweets* completos: “No hay forma (¡CERO!) de que los votos por correo sean algo menos que sustancialmente fraudulentos. Los buzones serán robados, las boletas serán falsificadas e incluso ilegalmente impresas y fraudulentamente firmadas. El Gobernador de California está enviando boletas a millones de personas. Cualquiera que viva en el Estado, sin importar quién es ni cómo llegó allí obtendrá una. Esto continuará con profesionales diciéndoles a todas esas personas, muchos de los cuales nunca habían incluso pensado en votar anteriormente, cómo y por quién deberían votar. Esta será una elección arreglada. ¡De ninguna manera!”. Traducción propia de las etiquetas agregadas por Twitter: “Obtén los hechos acerca de los votos por correo”.

⁷⁵ “Evento” de Twitter (26/05/2020). “Trump afirma sin fundamentos que la votación por correo conducirá al fraude electoral”. Disponible en: <https://twitter.com/i/events/1265330601034256384> (en inglés - última consulta: 17/12/2021)

“eventos”– que se titulaba “Trump afirma sin fundamentos que la votación por correo conducirá al fraude electoral”.

En dicho enlace, además, se pueden ver *tweets* de cuentas oficiales de Twitter y enlaces adicionales a medios de comunicación como CNN y *The Washington Post* con información acerca de las elecciones por correo. Por otro lado, se observa también una sección titulada “Lo que necesitas saber” en donde se informa que los “verificadores de hechos” (*fact-checkers*) dicen que no existe evidencia sobre lo denunciado por Trump y, a su vez, se refuta el reclamo acerca de la situación de California.

De acuerdo con un vocero de Twitter⁷⁶, esta decisión se tomó dado que las afirmaciones de Trump contenían información “potencialmente engañosa” acerca de los procesos electorarios y que, por este motivo, se había brindado información adicional de contexto a través del sistema de etiquetado.

A raíz de esto, el mismo día Trump publicó una nueva serie de *tweets* en los cuales acusó a la plataforma de interferir en la venidera elección presidencial. Además, denunció que Twitter estaba “sofocando completamente la libertad de expresión” y aseguró que no permitiría que eso sucediera.



Fuente: @realDonaldTrump en Twitter⁷⁷

⁷⁶ Mangan, D. y Breuninger, K. (26/5/2020). “Twitter ‘verifica los hechos’ de Trump y coloca etiquetas de advertencia en sus *tweets* sobre las elecciones por correo”. *CNBC*. Disponible en: <https://www.cnn.com/2020/05/26/twitter-fact-checks-trump-slaps-warning-labels-on-his-tweets-about-mail-in-ballots.html> (en inglés - última consulta: 19/12/2021).

⁷⁷ Traducción propia de los *tweets* completos: “@Twitter está interfiriendo en la elección presidencial de 2020. Están diciendo que mi afirmación acerca de que la votación por correo llevará a una corrupción masiva y a un fraude es incorrecta, basándose en un chequeo de datos realizado por las ‘fake news’ de

En este sentido, Trump redactó dos días después una orden ejecutiva⁷⁸ que se tituló “Previniendo la Censura En Línea”. En ese documento se pueden leer fragmentos como: “No podemos permitir que un número limitado de plataformas en línea escojan los discursos a los que los estadounidenses pueden acceder y transmitir en Internet”. O bien: “Cuando las grandes y poderosas empresas de redes sociales censuran opiniones con las que no están de acuerdo, ejercen un poder peligroso. Dejan de funcionar como tableros de anuncios pasivos y deben ser vistos y tratados como creadores de contenido”.

A nivel regulatorio, esta orden ejecutiva se centra en la modificación de la ya mencionada sección 230 de la *Communications Decency Act* (CDA) [Ley de Decencia en las Comunicaciones], que garantiza la inmunidad de las plataformas frente al contenido publicado por sus usuarios. Frente a este escenario, en la orden se observa el siguiente párrafo:

La inmunidad no debe extenderse más allá de su texto y propósito para brindar protección a aquellos que pretenden brindar a los usuarios un foro de expresión libre y abierta, pero en realidad utilizan su poder sobre un medio vital de comunicación para participar en acciones engañosas o que se usan como pretexto para asfixiar el debate libre y abierto censurando ciertos puntos de vista.

Tal como se explicitó en el apartado regulatorio del marco teórico, la sección 230 establece que los proveedores de servicios informáticos no pueden ser considerados como editores de la información o los contenidos que produzcan otros actores. Sin embargo, según explicó Gillespie (2017), esta regulación implica también una segunda faceta: los intermediarios pueden conservar su protección aún en esos casos en los que decidan monitorear e intervenir sobre el contenido que se publica en sus plataformas, ya que estas acciones tampoco las convertirían en editoras de contenido.

Los fragmentos citados de las modificaciones regulatorias propuestas por esta orden ejecutiva permiten comprender con claridad la intención manifiesta de limitar la protección con la que cuentan las plataformas de redes sociales en los Estados Unidos, tanto en aquellos escenarios en los que decidan intervenir como en los que no.

CNN y por el Washington Post de Amazon. ¡Twitter está sofocando completamente la LIBERTAD DE EXPRESIÓN y yo, como Presidente, no voy a permitir que suceda!”

⁷⁸ Registro Federal de los Estados Unidos (28/05/2020). “Previniendo la censura en línea”. Disponible en: <https://www.federalregister.gov/documents/2020/06/02/2020-12030/preventing-online-censorship> (en inglés - última consulta: 19/12/2021).

Como es lógico, este proyecto no fue bien recibido por parte de Twitter. En una publicación de una de sus cuentas oficiales, se responde a la orden ejecutiva acusándola de “reaccionaria y politizada”. Además, se defiende el funcionamiento de la Sección 230 y su valor para la “innovación y la libertad de expresión”.



Fuente: @Policy en Twitter⁸⁰

Resulta destacable de este intercambio el hecho de que ambas partes manifiestan una clara intención de defender la libertad de expresión como un valor primordial, aun cuando sus posiciones ante la misma regulación sean completamente contradictorias.

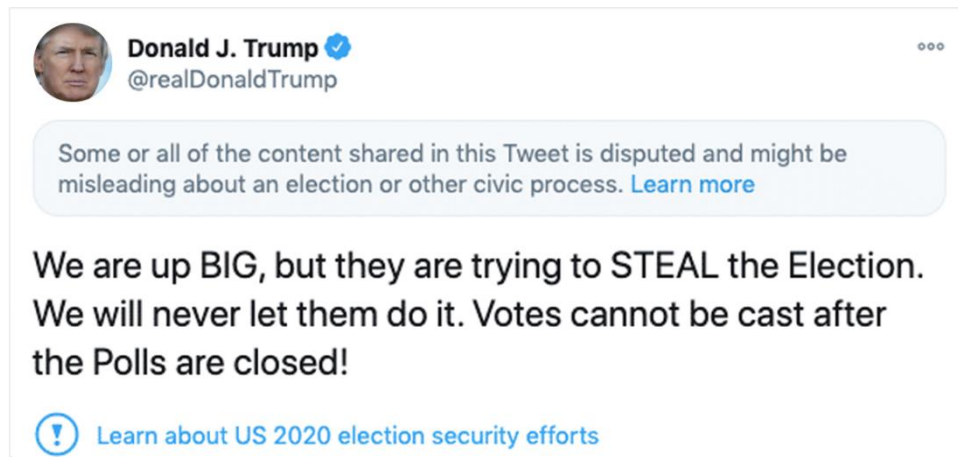
En términos prácticos, la orden ejecutiva emitida por Trump no tuvo mayores efectos y, de hecho, fue revocada por su sucesor Joe Biden a través de una nueva orden ejecutiva⁸¹ del 14 de mayo de 2021. Sin embargo, Twitter no se detuvo en su “etiquetado” de los *tweets* del entonces presidente.

Como se mencionó anteriormente, las elecciones presidenciales tuvieron una importante demora en el conteo de sus votos y eso permitió que surgieran diversas denuncias de fraude desde el espacio republicano. Cabe aclarar también que, al momento de la escritura de esta tesina, ninguna de estas denuncias tuvo fundamentos suficientes como para generar algún efecto.

⁸⁰ Traducción propia del *tweet* completo: “Esta Orden Ejecutiva es un enfoque reaccionario y politizado ante una ley de referencia. La #Sección230 protege la innovación y la libertad de expresión de Estados Unidos y se apoya en valores democráticos. Los intentos por erosionarla unilateralmente son una amenaza para el futuro de la libertad en el discurso en línea y en Internet”. Disponible en: <https://twitter.com/Policy/status/1266170586197262337> (última consulta - 19/12/2021).

⁸¹ Registro Federal de los Estados Unidos (14/05/2021). “Revocación de determinadas acciones presidenciales y modificación técnica”. Disponible en: <https://www.federalregister.gov/documents/2021/05/19/2021-10691/revocation-of-certain-presidential-actions-and-technical-amendment> (última consulta - 19/12/2021).

El 4 de noviembre de 2020, cuando todavía los resultados oficiales no estaban publicados, Trump compartió un *tweet* en el que anunciaba, tal como se ve en la imagen que sigue, que se encontraba por delante en el recuento de votos y que estaban intentando “robar la elección”. A su vez, denunciaba que había votos que se habrían enviado luego del cierre de los comicios.



Fuente: @realDonaldTrump en Twitter⁸²

Frente a esta situación, Twitter tomó la decisión de colocar una advertencia en la publicación de Trump (que también se puede observar en la imagen precedente) en la que aparecen dos cuestiones. En la parte inferior, tal como se vio en casos anteriores, un enlace con información de contexto, en este caso acerca de los “esfuerzos de seguridad durante la elección 2020”. Además, en este caso particular se agregó una segunda advertencia por encima del *tweet* propiamente dicho en la que se puede ver la siguiente frase: “Parte o todo el contenido de este *tweet* es controversial y podría resultar engañoso para una elección u otro proceso cívico”.

Posteriormente, una de las cuentas oficiales de Twitter (@TwitterSafety) compartió un *tweet* en el que anunciaba la advertencia previamente señalada y comunicaba que esta decisión estaba alineada con su “Política de Integridad Cívica”.

⁸² Traducción propia del *tweet* completo: “Estamos ganando POR MUCHO, pero están tratando de ROBAR la elección. Nunca los dejaremos hacerlo. ¡Los votos no se pueden emitir una vez que cierran los comicios!” Traducción propia de la etiqueta superior: “Parte o todo el contenido de este *tweet* es controversial y podría resultar engañoso para una elección u otro proceso cívico”. Traducción propia de la etiqueta inferior: “Conoce acerca de los esfuerzos de seguridad para las elecciones 2020”.



Fuente: @TwitterSafety en Twitter⁸⁴

Además, se podía observar allí un enlace⁸⁵ que redirigía al usuario hacia el texto de la mencionada Política. En ese documento se informa, por ejemplo, que Twitter puede “etiquetar y reducir la visibilidad de los *tweets* que contengan información falsa o engañosa acerca de los procesos cívicos”. Así, por más de que no se haya realizado explícitamente en esta oportunidad, se observa que además de la posibilidad de etiquetado existe también la alternativa de conseguir que ciertos contenidos sean menos visibles para los usuarios.

Aquí cabe recuperar lo planteado por Napoli y Caplan (2019) acerca de la imposibilidad de hablar de algoritmos que trabajan de manera “neutral y objetiva”. Como se explicita en el documento de Twitter, es evidente que las plataformas tienen un poder claro a la hora de decidir qué contenidos pueden ver los usuarios y con qué frecuencia.

Finalmente, otro de los puntos interesantes para analizar que se observan en esta política corporativa tiene que ver con el sistema de penalidades en el caso de incumplir con ella que se canaliza a través de un sistema de *strikes*⁸⁶.

En primer lugar, ante violaciones de “alta severidad” se menciona la posible eliminación de *tweets* o de otra información dentro del perfil de un usuario que incluya

⁸⁴ Traducción propia del *tweet* completo: “Hemos colocado una advertencia en un *tweet* de @realDonaldTrump por realizar una afirmación potencialmente engañosa acerca de una elección. Esta acción está en línea con nuestra Política de Integridad Cívica. Más aquí:”.

⁸⁵ Centro de Ayuda de Twitter (Octubre 2021). “Política de integridad cívica”. Disponible en: <https://help.twitter.com/es/rules-and-policies/election-integrity-policy> (última consulta - 19/12/2021).

⁸⁶ Este sistema tiene su origen en el béisbol. Consiste en que un jugador que cometa tres “*strikes*”, es decir que no pueda golpear la pelota con el bate en el momento oportuno, queda fuera de juego.

información engañosa sobre la participación democrática o contenidos intimidatorios. Un *tweet* eliminado implica la acumulación de dos *strikes*.

La segunda situación que se plantea en este documento es el ya mencionado sistema de “etiquetado”. A las particularidades ya descritas de esta herramienta de etiquetado implementada por Twitter cabe agregar que, bajo este nuevo sistema, esta violación de los términos implica un *strike*.

Finalmente, la última de las posibles sanciones tiene que ver con el bloqueo o la suspensión permanente de la cuenta de un usuario a las que se llegaría luego de “severas o repetidas violaciones de esta política”. Hacia el final del artículo, se encuentra la siguiente clasificación de las medidas punitivas:

- A) 1 *strike*: no se toman acciones sobre la cuenta
- B) 2-3 *strikes*: suspensión de la cuenta por 12 horas
- C) 4 *strikes*: suspensión de la cuenta por 7 días
- D) 5 o más *strikes*: suspensión permanente de la cuenta

Resulta relevante recuperar este sistema de penalidades de Twitter ya que luego se ve reflejado en el caso de Trump. En los apartados que siguen, el análisis continúa específicamente con los contenidos eliminados y la definitiva suspensión de la cuenta de Donald Trump.

4.4. *Tweets* eliminados a Donald Trump: las reglas de Twitter y la relación con el interés público

El 6 de enero de 2021, día en que sucedieron los recordados incidentes en el interior del Capitolio, fue también la primera ocasión en que Twitter decidió que algunas publicaciones de Donald Trump debían ser eliminadas. Precisamente, fueron tres los *tweets* borrados y, a continuación, se profundiza en estas decisiones.

En el primero de los *tweets* que luego serían eliminados, Trump criticaba directamente a su vicepresidente Pence, afirmando que éste “no tuvo el coraje para hacer lo que debió haberse hecho para proteger al país”, haciendo referencia a la certificación de los votos que Trump consideró “fraudulentos o inexactos”.



Fuente: @realDonaldTrump en Twitter⁸⁷

Tomando en cuenta los hechos ocurridos, y antes de decidirse por la remoción de esta publicación, la plataforma había colocado una advertencia en el *tweet* tal como se puede ver en la imagen anterior. Allí, se observa que la afirmación de fraude electoral fue catalogada como “controversial” y que se había restringido la posibilidad de responder, “retuitear” o señalar con “me gusta” a dicho *tweet* debido a un “riesgo de violencia”.

La segunda de las publicaciones que luego sería eliminada era un video⁸⁹ grabado por Trump en el que se dirigía directamente a sus seguidores que estaban en el Capitolio al momento de los hechos. En dicho documento, se podía escuchar al entonces presidente plantear nuevamente que se había tratado de una “elección fraudulenta”, o bien que le habían “robado una elección”. Sin embargo, también había un llamado a sus seguidores a “volver a casa” y a conservar la paz y el orden para evitar más heridos. Al igual que en el caso anterior, este video fue en principio catalogado como “controversial” y se impidió la posibilidad de compartirlo antes de ser eliminado de manera permanente.

Finalmente, en el tercero de los *tweets* eliminados de esa jornada del 6 de enero, Trump escribió: “Estas son las cosas y los eventos que ocurren cuando una victoria electoral aplastante es brutal y bruscamente despojada de los grandes patriotas que han sido tratados mal e injustamente por tanto tiempo. Váyanse a casa con amor y en paz. ¡Recuerden este día para siempre!”.

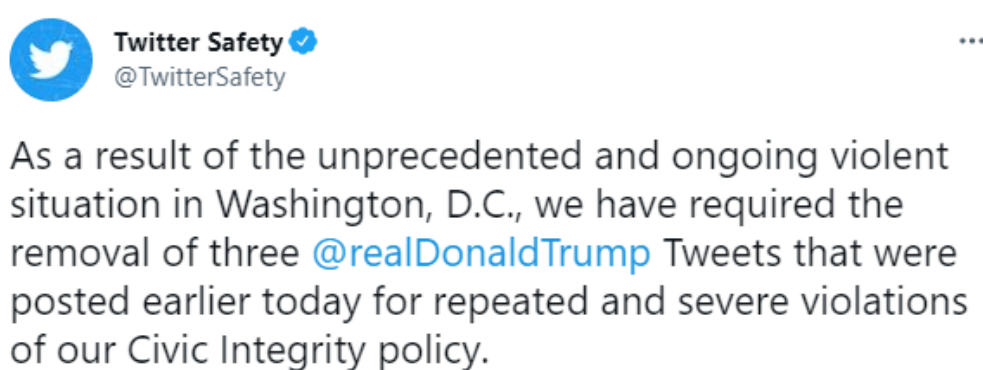
⁸⁷ Traducción propia del *tweet* completo: “Mike Pence no tuvo el coraje para hacer lo que debió haberse hecho para proteger a nuestro país y nuestra Constitución, dándole a los Estados una posibilidad de certificar los datos corregidos, no aquellos fraudulentos o imprecisos que se solicitaron certificar anteriormente. ¡Estados Unidos demanda la verdad!” Traducción propia de la etiqueta: “Esta afirmación de fraude electoral es controversial, y este *tweet* no puede ser respondido, *retwitteado* ni se le puede indicar ‘me gusta’ debido a un riesgo de violencia”.

⁸⁹ Moreau, J. (6/1/2021). “El mensaje tranquilizador del presidente Trump para la multitud violenta que ingresó al Capitolio: ‘Vayan a casa. Los amamos, son muy especiales.’”. *Variety*. Disponible en: <https://variety.com/2021/politics/news/trump-protesters-riot-capitol-video-1234879939/> (en inglés - última consulta: 19/12/2021).



Fuente: @realDonaldTrump en Twitter⁹⁰

Esta publicación, que recibió la misma advertencia que las anteriores, fue la última que publicó Trump durante ese día. Luego, como ya se ha adelantado, Twitter tomó la decisión de que estas tres publicaciones debían borrarse y comunicó esa determinación a sus usuarios a través de una nueva publicación de la cuenta oficial @TwitterSafety.



Fuente: @TwitterSafety en Twitter⁹²

En este hilo⁹³, desde la plataforma se explicó que, como resultado de la “situación violenta en curso y sin precedentes en Washington”, se requería la eliminación de los mencionados tres *tweets* debido a sus “repetidas y severas violaciones” de la Política de Integridad Cívica (la misma que fue analizada en el apartado anterior de este trabajo).

⁹⁰ Traducción propia de la etiqueta: “Esta afirmación de fraude electoral es controversial, y este *tweet* no puede ser respondido, *retwitteado* ni se le puede indicar ‘me gusta’ debido a un riesgo de violencia”.

⁹² Traducción propia del *tweet* original: “Como resultado de la situación violenta en curso y sin precedentes en Washington D.C., hemos pedido la eliminación de tres *tweets* de @realDonaldTrump que fueron publicados hoy por repetidas y severas violaciones a nuestra Política de Integridad Cívica”.

Disponible en: <https://twitter.com/TwitterSafety/status/1346970430062485505> (última consulta: 19/12/2021).

⁹³ Una sucesión encadenada de *tweets* que hacen referencia al mismo tema.

Además, en las publicaciones siguientes del hilo se aclaró que la cuenta de Trump se mantendría bloqueada por doce horas y, en el caso de que éste no borrara los *tweets*, la cuenta continuaría bloqueada. Por otro lado, se advirtió que en el caso de futuras violaciones de las “Reglas de Twitter” la cuenta sería suspendida de manera permanente.

Las mencionadas “Reglas de Twitter”⁹⁴, si bien no estaban enlazadas directamente al hilo de *tweets*, se pueden encontrar con facilidad en el centro de ayuda oficial. Este documento presenta una breve descripción de cada una de las políticas y otras reglas que delimitan qué contenidos son pasibles de ser cuestionados por las autoridades de la plataforma.

Algunas de ellas, como la Política de Integridad Cívica o aquella que hace referencia a la manipulación de contenido multimedia, ya han sido mencionadas anteriormente en este trabajo a medida que fueron utilizadas por la plataforma en los acontecimientos relacionados con el caso de análisis.

No resulta relevante para los objetivos de este trabajo realizar una descripción detallada de todas las reglas de Twitter ya que, en su mayoría, no están relacionadas directamente con el caso. En cambio, y tal como se fue realizando hasta el momento, se recuperan aquellas políticas que esgrimió la propia plataforma como argumentos a la hora de tomar sus decisiones.

Finalmente, otra información relevante que agrega esta serie de publicaciones oficiales de Twitter es un *tweet* en el que se puede leer lo siguiente: “Nuestra política de interés público, que ha guiado nuestras acciones de cumplimiento en esta área durante años, termina allí donde creemos que el riesgo de daño es mayor y/o más severo”⁹⁵.

⁹⁴ Centro de Ayuda de Twitter (s.f.). “Las Reglas de Twitter”. Disponible en: <https://help.twitter.com/es/rules-and-policies/twitter-rules> (última consulta: 19/12/2021).

⁹⁵ Traducción propia



Twitter Safety
@TwitterSafety

...

En respuesta a @TwitterSafety

Our public interest policy — which has guided our enforcement action in this area for years — ends where we believe the risk of harm is higher and/or more severe.

Fuente: @TwitterSafety en Twitter

Adicionalmente, se encuentra un enlace hacia una entrada del blog oficial de Twitter⁹⁷ que lleva el título de “Líderes Mundiales en Twitter: principios y enfoque”.

En este artículo, que tiene fecha de octubre del 2019, la plataforma desarrolló una explicación sobre la manera en que analiza las denuncias que realizan los usuarios sobre los *tweets* de los “líderes mundiales”. En este sentido, plantean que estos *tweets* reportados se analizan a la luz de las ya citadas “Reglas de Twitter” y, a continuación, se encuentra el siguiente párrafo:

Sin embargo, si el Tweet de un líder mundial viola las Reglas de Twitter, pero existe un claro valor de interés público para mantener el Tweet en el servicio, podemos mostrarlo detrás de un aviso que proporcione contexto sobre la violación en cuestión y permita que las personas que desean ver el contenido hagan clic en el mismo.

En el párrafo siguiente, sin embargo, se puede leer que estas cuentas pertenecientes a líderes mundiales “no están completamente por encima” de las políticas de Twitter. Así, ante ciertos casos, cualquier cuenta queda sujeta a acciones disciplinarias “sin tener en cuenta el posible valor de interés público al permitir que el *Tweet* permanezca visible detrás de un aviso”.

Los “casos” planteados por Twitter para incurrir en restricciones son los siguientes: promoción del terrorismo; amenazas de violencia claras y directas en contra de un individuo; publicar información privada; publicar o compartir videos íntimos sin

⁹⁷ Blog oficial de Twitter (15/10/2019). “Líderes mundiales en Twitter: principios y enfoque”. Disponible en: https://blog.twitter.com/es_la/topics/company/2019/lideres-mundiales-en-twitter-principios-y-enfoque (última consulta: 19/12/2021).

consentimiento; participar en comportamiento relacionados a la explotación sexual infantil; y fomentar o promover la autolesión.

De este análisis se podría concluir que resultaría complicado encasillar a las cuestionadas publicaciones de Trump dentro de alguna de estas categorías, dando por descontado que el presidente en ejercicio de los Estados Unidos calificaría como un “líder mundial”.

Por este motivo es que resulta más importante aún la publicación de la cuenta *Twitter Safety* recuperada previamente. Es la propia voz de la plataforma la que indica que, en este caso, deja de lado la política que tenía ante estos escenarios allí donde cree que “el riesgo de daño es mayor y/o más severo”.

Para complementar esta parte del análisis, cabe retomar algunas de las recomendaciones específicas de la Organización de Estados Americanos (OEA) y la Relatoría Especial para la Libertad de Expresión (RELE) de la Comisión Interamericana de Derechos Humanos (CIDH) para funcionarios públicos y para actores estatales y no estatales durante procesos electorales.

En el marco teórico, se había recuperado el planteo de la OEA (2010) que indicaba, por un lado, que los funcionarios públicos tienen la responsabilidad de expresarse ante ciertos hechos de relevancia pública, y que deberían hacerlo con el sustento de hechos constatables para evitar que los ciudadanos reciban versiones manipuladas.

Por otra parte, la RELE (2020) había planteado que, durante los procesos electorales, los actores estatales deberían asegurarse de que los medios obtuvieran un acceso sin impedimentos a las fuentes de información oficial y a los candidatos. A su vez, la RELE también recomendaba a las plataformas que adoptaran las medidas necesarias para garantizarles a sus usuarios el acceso a información plural, pero manteniéndose alerta ante la posible aparición de información manipulada o errónea.

Si bien, como ya se ha planteado, los Estados Unidos no han adherido a la Convención Americana sobre Derechos Humanos, estas recomendaciones forman parte de un marco de referencia internacional que puede ser tomado en cuenta ante estos escenarios.

Así, se puede observar que Twitter reconoce la importancia que puede tener que un contenido de interés público circule. Sin embargo, como explícitamente aclara la

plataforma, estos discursos no están por encima de sus políticas y, ante ciertos escenarios, las restricciones ocurrirán más allá de ese reconocido interés público.

Si se contrapone esto con las recomendaciones de los Órganos Internacionales, se puede observar que, sobre todo en el contexto de una elección, es necesario que los funcionarios se expresen ante hechos de relevancia pública. A su vez, es importante que las plataformas trabajen en pos de la pluralidad en la información para sus usuarios.

Al mismo tiempo, está presente el pedido para las plataformas de prestar atención ante las informaciones erróneas o manipuladas, por lo que se puede inferir de aquí la demanda de intervenir ante estos escenarios. La colisión entre ambas situaciones presenta un escenario de debate que no parece estar saldado, pero que sí dispara la preocupante consecuencia de depender del criterio de un actor privado a la hora de definir qué discursos pueden circular y cuáles no.

Esta investigación no pretende ni se propone evaluar si efectivamente existía un riesgo mayor para la continuidad de los serios incidentes en el Capitolio si la publicación de Trump hubiera permanecido en línea. Sin embargo, resulta interesante remarcar el modo en que una norma que rige las condiciones bajo las cuales se publica en la plataforma puede adaptarse de acuerdo con las circunstancias. Cabe recordar que esa “norma” no es una ley, sino que son las reglas de un actor no estatal, cuyos criterios de aplicación y evaluación no son transparentes y que no cuenta con un procedimiento de “queja” o “defensa” de la parte afectada.

4.5. La suspensión permanente de @realDonaldTrump: motivos, argumentos y reflexiones

Finalmente, Trump decidió eliminar los *tweets* que habían sido objetados por la plataforma y, luego del plazo de suspensión, recuperó acceso a su cuenta⁹⁸. La única publicación que realizó durante el 7 de enero fue un enlace a un video⁹⁹ en el que se lo

⁹⁸ Klar, R. (7/1/2021). “Trump borró los *tweets* que llevaron al bloqueo de su cuenta de Twitter”. *The Hill*. Disponible en: <https://thehill.com/policy/technology/533114-trump-deleted-tweets-that-led-to-his-twitter-lockout?rl=1> (en inglés - última consulta: 19/12/2021).

⁹⁹ Feiner, L. (7/1/2021). “Trump reconoce el resultado de las elecciones en el primer *tweet* después de que se levantara la suspensión”. *CNBC*. Disponible en: <https://www.cnn.com/2021/01/07/trump-sends-first-tweet-after-twitter-removes-lock-from-his-account.html> (en inglés - última consulta: 19/12/2021).

escucha criticar los hechos del Capitolio y anunciar que intentaría realizar una transición ordenada para la asunción del nuevo gobierno.

A pesar de esto, el día posterior sería el último en el que la cuenta @realDonaldTrump estaría habilitada. Las dos únicas (y últimas) publicaciones de Trump durante ese 8 de enero pueden leerse con un tono diferente al del video recientemente citado. En la primera de ellas, se observa que Trump alienta nuevamente a sus seguidores indicando que tendrán una “voz gigante” en el futuro y que no permitirá que “les falten el respeto ni los traten injustamente”.

En la segunda publicación, el mensaje es el siguiente: “Para todos aquellos que han preguntado, no voy a asistir a la Inauguración el 20 de enero”. Este mensaje, lógicamente, hace referencia a la ceremonia de asunción de Joe Biden como nuevo presidente de los Estados Unidos.



Donald J. Trump @realDonaldTrump · 2h

The 75,000,000 great American Patriots who voted for me, AMERICA FIRST, and MAKE AMERICA GREAT AGAIN, will have a GIANT VOICE long into the future. They will not be disrespected or treated unfairly in any way, shape or form!!!

105K 99.3K 313K



Donald J. Trump @realDonaldTrump · 1h

To all of those who have asked, I will not be going to the Inauguration on January 20th.

164K 101K 251K

Fuente: @realDonaldTrump en Twitter¹⁰⁰

Ante esto, la reacción por parte de Twitter fue contundente: la cuenta de Trump fue suspendida de manera permanente. En una entrada del blog oficial de la plataforma¹⁰¹,

¹⁰⁰ Traducción propia de los tweets completos: “¡Los 75.000.000 grandes Patriotas Estadounidenses que votaron por mí, por poner a ESTADOS UNIDOS PRIMERO, y HACER A LOS ESTADOS UNIDOS GRANDES NUEVAMENTE, tendrán una VOZ GIGANTE en el futuro. No les faltarán el respeto ni los tratarán injustamente de ninguna forma o manera! Para todos aquellos que han preguntado, no voy a asistir a la Inauguración el 20 de enero”.

se puede encontrar el análisis completo que se realizó para tomar la decisión. Sin embargo, y a los fines de este trabajo, se rescatan algunos fragmentos. El primero es un párrafo que hace referencia a los dos *tweets* recién mencionados:

Debido a las tensiones en curso en Estados Unidos, y a un repunte en la conversación global con respecto de las personas que irrumpieron violentamente en el Capitolio estadounidense el 6 de enero de 2021, estos dos Tweets deben leerse en el contexto de eventos mayores en el país y en las formas en que las declaraciones del presidente pueden ser movilizadas por diferentes audiencias, incluso para incitar a la violencia, así como en el contexto del patrón de comportamiento de esta cuenta en las últimas semanas. Después de evaluar el lenguaje de estos Tweets bajo nuestra política de Glorificación de la Violencia, hemos determinado que estos Tweets infringen la política de Glorificación de la Violencia y el usuario @realDonaldTrump debe ser suspendido de manera inmediata y de forma permanente del servicio.

De estas líneas se pueden desprender algunas cuestiones. En primer lugar, que el contexto tanto del país como de las publicaciones anteriores hechas por Trump fueron un factor esencial a la hora de tomar la decisión. Por otra parte, y quizás aún más interesante, resulta el hecho de que se interpretó que las publicaciones podrían movilizar a ciertas personas a incitar a la violencia. Se podría entender, entonces, que no existiría una incitación a la violencia en sí misma, sino una posibilidad indirecta de llegar a estas nuevas escenas de violencia.

Finalmente, se observa también que aparece una nueva política que habría sido infringida por Trump, aquella que hace referencia a la “glorificación de la violencia”. Esta política¹⁰², según se explica en el sitio oficial, prohíbe “hacer amenazas de violencia contra una persona o un grupo de personas”, así como también la “glorificación de la violencia”. El concepto de “glorificación”, para Twitter, incluye celebrar o elogiar acciones violentas entendiendo que esto “podría inspirar a otros a participar en actos de violencia similares”.

¹⁰¹ Blog oficial de Twitter (8/1/2021). “Suspensión permanente a @realDonaldTrump”. Disponible en: https://blog.twitter.com/es_la/topics/company/2020/suspension-permanente-a-realdonaldtrump (última consulta: 19/12/2021).

¹⁰² Centro de Ayuda de Twitter (Marzo, 2019). “Política relativa a la glorificación de la violencia”. Disponible en: <https://help.twitter.com/es/rules-and-policies/glorification-of-violence> (última consulta: 19/12/2021).

Acerca de esto, Twitter realizó un análisis más detallado de las frases particulares de los *tweets* que incumplirían con esta política en la misma entrada del blog citada recientemente.

En primera medida, entienden que la negativa de Trump a asistir a la asunción de Biden “está siendo recibida por varios de sus partidarios como una confirmación más de que la elección no fue legítima”. Nuevamente, se observa el hecho de que el foco está puesto en la interpretación que terceras personas podrían hacer acerca de una afirmación. Además, en ese artículo se interpreta que la ausencia de Trump podría significar un incentivo para aquellos que quisieran realizar actos violentos.

Por otra parte, y en referencia al primero de los *tweets*, el hecho de que Trump se haya referido a sus partidarios como “patriotas” puede interpretarse, para Twitter, como un apoyo a quienes cometieron actos violentos en el Capitolio. A su vez, la afirmación de que sus seguidores tendrían una “voz gigante” y que no se les faltaría el respeto en el futuro fue interpretado “como una indicación más de que el presidente Trump no planea facilitar una ‘transición ordenada’”.

Como ya se mencionó, este trabajo no se propone evaluar si las publicaciones de Trump en Twitter debían ser o no eliminadas, sino los procedimientos y argumentaciones que se llevaron adelante para hacerlo. Por lo tanto, se puede indicar que las justificaciones están mayormente basadas en interpretaciones acerca de cómo terceras personas podrían a su vez interpretar las publicaciones. Como es evidente, todas estas situaciones se analizan exclusivamente a la luz de políticas y normas diseñadas por la propia plataforma.

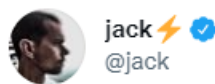
A partir de ese 8 de enero, la cuenta de Trump permaneció suspendida. Al momento de escribir esta tesina, si se intentaba ingresar al perfil¹⁰³ se podía observar un perfil que no tenía foto, biografía ni ninguna otra información más allá del nombre de la cuenta (@realdonaldtrump). Por debajo, aparecía una leyenda que indicaba: “Cuenta suspendida. Twitter suspende las cuentas que incumplen las Reglas de Twitter” y un enlace que dirigía al usuario a las ya citadas “Reglas de Twitter”.

¹⁰³ Enlace a la cuenta oficial de Donald Trump, suspendida al momento de la escritura de esta tesina. Disponible en: <https://twitter.com/realdonaldtrump> (última consulta: 19/12/2021).



Fuente: @realDonaldTrump en Twitter

Cinco días después de la suspensión, Jack Dorsey, co-fundador y CEO de Twitter, se expresó acerca de este hecho a través de un extenso hilo¹⁰⁴ de *tweets* que publicó en su cuenta oficial (@jack). Se analizan a continuación algunos de los fragmentos más interesantes de su exposición.



I do not celebrate or feel pride in our having to ban @realDonaldTrump from Twitter, or how we got here. After a clear warning we'd take this action, we made a decision with the best information we had based on threats to physical safety both on and off Twitter. Was this correct?

Fuente: @jack en Twitter¹⁰⁵

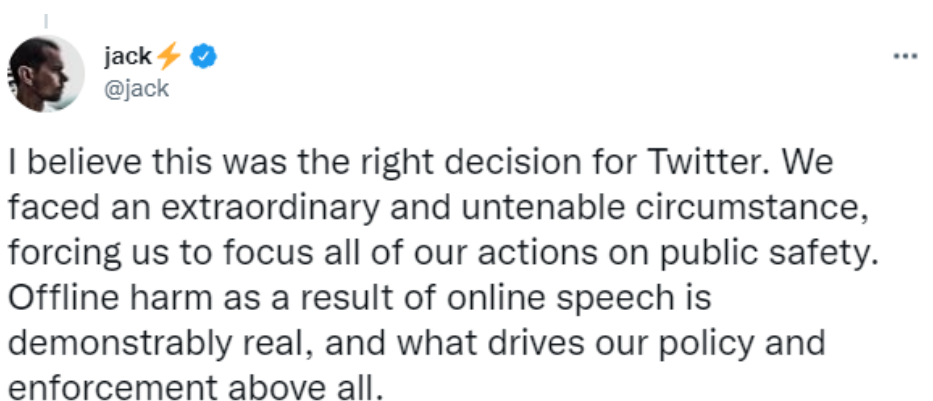
En esta primera publicación, Dorsey afirmó que no celebraba ni se sentía orgulloso de esta decisión que, sin embargo, fue tomada “luego de una clara advertencia”. A su vez,

¹⁰⁴ Hilo de *tweets* redactados por Jack Dorsey, co-fundador y, en ese momento, CEO de Twitter. (13/01/2021). Disponible en: <https://twitter.com/jack/status/1349510769268850690> (en inglés - última consulta: 19/12/2021).

¹⁰⁵ Traducción propia del *tweet* completo: “No celebro ni siento orgullo en haber tenido que suspender a @realDonaldTrump de Twitter ni de cómo llegamos aquí. Luego de una clara advertencia de que tomaríamos esta acción, tomamos una decisión con la mejor información que teníamos basada en las amenazas a la seguridad física tanto dentro como fuera de Twitter. ¿Fue la decisión correcta?”

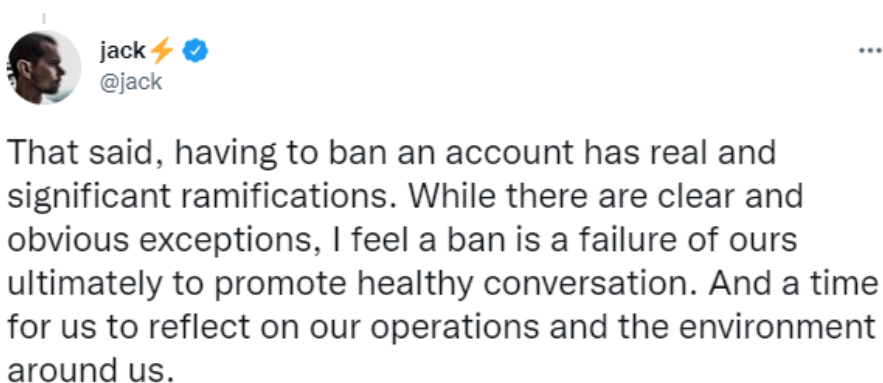
explicó que la decisión se tomó sobre la base de la “mejor información” que tenían acerca de “amenazas a la seguridad física tanto dentro como fuera de Twitter”.

Dorsey continúa su descargo planteando que la suspensión de la cuenta de Trump fue la “decisión correcta” para Twitter, entendiendo que “el daño fuera de las redes como resultado del discurso en línea es real de manera demostrable” y que ese hecho es el que guía en última medida sus decisiones. Esta argumentación es muy similar a la que se encuentra en el comunicado oficial que informaba la suspensión de la cuenta.



Fuente: @jack en Twitter¹⁰⁷

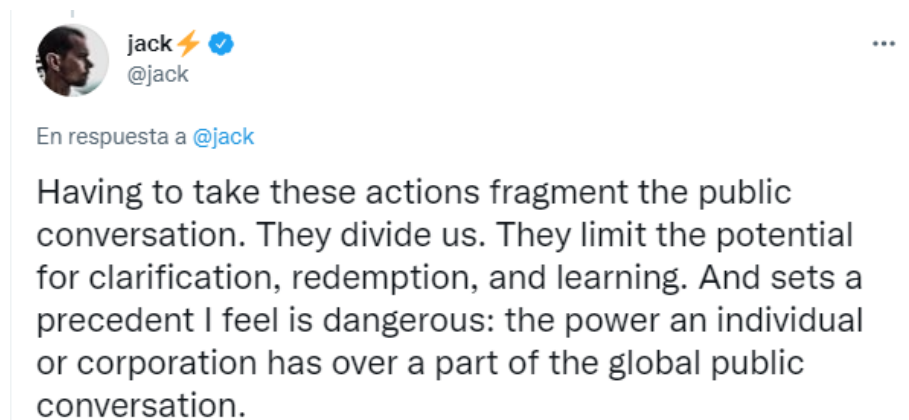
Sin embargo, más adelante en su texto Dorsey plantea que el hecho de haber tenido que llegar a esta decisión fue “un fracaso” de la intención de la plataforma de facilitar una “conversación sana” y que era una acción que “fragmenta la conversación pública”.



Fuente: @jack en Twitter¹⁰⁸

¹⁰⁷ Traducción propia del *tweet* completo: “Creo que fue la decisión correcta para Twitter. Enfrentamos una circunstancia extraordinaria e insostenible, que nos forzó a concentrar todas nuestras acciones en la seguridad pública. El daño fuera de línea como consecuencia del discurso en línea es real de manera demostrable, y eso es lo que guía nuestras políticas y nuestras decisiones por sobre todas las cosas”.

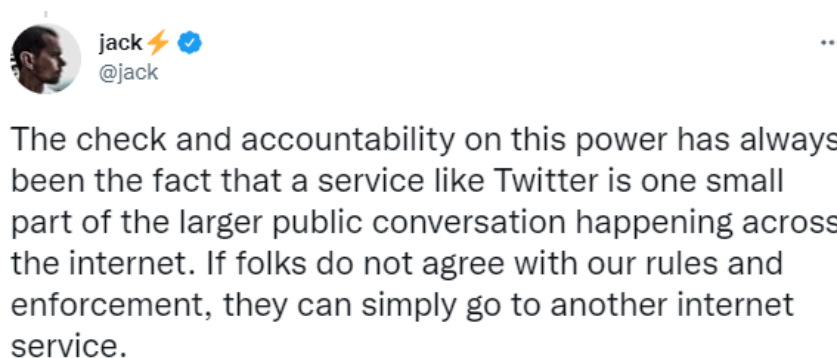
Más interesante aún resulta el hecho de que el CEO de Twitter describe esto como un “precedente” peligroso haciendo referencia al “poder que tiene un individuo o una empresa sobre parte de la conversación pública global”.



Fuente: @jack en Twitter¹⁰⁹

A continuación, Dorsey menciona otro punto interesante para los objetivos del presente trabajo y que tiene que ver con el control de este “poder” que se trajo a colación en el párrafo anterior.

El control y la responsabilidad de este poder siempre ha sido el hecho de que un servicio como Twitter es una pequeña parte de la conversación pública más amplia que tiene lugar en Internet. Si la gente no está de acuerdo con nuestras reglas y su cumplimiento, simplemente puede ir a otro servicio de Internet.



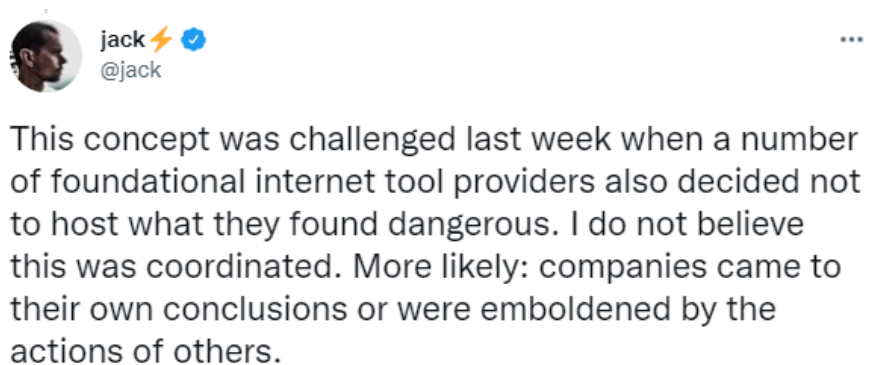
Fuente: @jack en Twitter

¹⁰⁸ Traducción propia del *tweet* completo: “Dicho esto, tener que suspender una cuenta tiene ramificaciones reales e importantes. A pesar de que existen excepciones claras y obvias, siento que una suspensión es un fracaso nuestro en última instancia en promover una conversación sana. Y un momento para que reflexionemos sobre nuestras operaciones y sobre el ambiente que nos rodea”.

¹⁰⁹ Traducción propia del *tweet* completo: “Tener que tomar estas acciones fragmenta la conversación pública. Nos dividen. Limitan el potencial de clarificación, redención y aprendizaje. Y establecen un precedente que creo peligroso: el poder que un individuo o una empresa tiene sobre parte de la conversación pública global”.

Así, de alguna manera, se puede interpretar que serían los mismos usuarios quienes harían que este poder no fuera tan absoluto, ya que tendrían el poder de mudarse de plataforma si así lo desearan.

Sin embargo, Dorsey reconoce que este sistema se vio desafiado en este caso ya que varios proveedores de servicios similares “también decidieron no alojar lo que consideraban peligroso”, aunque aclara que no se trató de una acción coordinada entre las distintas plataformas.



Fuente: @jack en Twitter¹¹¹

Más allá de que Dorsey afirme que las restricciones a través de las distintas plataformas no han sido coordinadas, es importante recalcar que la existente –y marcada– concentración empresarial que existe en Internet hace que sean estas grandes cadenas de plataformas las que dominan con sus decisiones el ecosistema de los medios en línea (Van Dijck, 2016).

De este modo, por más que realmente no haya sido una acción en conjunto, continúan siendo pocas empresas las que toman las decisiones y definen lo que está permitido o no en las plataformas. Como se ha visto en el apartado de concentración, las redes sociales se han convertido en importantes fuentes de acceso a la información para los usuarios. El hecho de que la gran mayoría de estas empresas, de manera coordinada o no, decidan actuar de manera idéntica ante un hecho de tamaño magnitud representa un daño evidente para las opciones de los usuarios a acceder a información de interés público de forma plural y democrática.

¹¹¹ Traducción propia del *tweet* completo: “Este concepto se vio desafiado la semana pasada cuando muchas de los proveedores de herramientas fundacionales de internet también decidieron no alojar aquello que encontraban peligroso. No creo que esto haya sido coordinado. Lo más probable: las compañías llegaron a sus propias conclusiones o fueron incentivadas por las acciones de otros”.

Finalmente, la última reflexión que se destaca de la máxima autoridad de Twitter sobre este asunto es la siguiente: “Una compañía que toma la decisión empresarial de moderarse a sí misma es diferente a un gobierno que elimina el acceso, pero puede sentirse muy parecido”. Así, si bien en todo momento Dorsey afirma que Twitter ha tomado la decisión adecuada para el contexto en el que se encontraba, abre una ventana de reflexión acerca de la moderación de contenidos y plantea explícitamente que se necesita “mayor transparencia” en ese ámbito para garantizar un Internet global “libre y abierto”.



Fuente: @jack en Twitter¹¹³

Acerca de este último planteo que se fue construyendo sobre la base del hilo de Dorsey se profundiza en el próximo apartado de este trabajo.

¹¹³ Traducción propia de los *tweets* completos: “Este momento puede haber necesitado de esta dinámica, pero a largo plazo será destructivo para los propósitos nobles y los ideales de Internet. Una compañía que toma la decisión empresarial de moderarse a sí misma es diferente a un gobierno que elimina el acceso, pero puede sentirse muy parecido” y “Sí, todos nosotros necesitamos darle una mirada crítica a las inconsistencias de nuestra políticas y nuestras acciones de cumplimiento. Sí, necesitamos revisar cómo nuestro servicio puede incentivar la distracción y el daño. Sí, necesitamos más transparencia en nuestras operaciones de moderación. Todo esto no puede erosionar un Internet libre y abierto”.

4.6. Términos de servicio, reglas y políticas: propuestas y relación con otros modos de regulación

Tal como se estableció hacia el final del apartado anterior, se retoma aquí el análisis partiendo de la base de los planteos realizados por el CEO de Twitter, Jack Dorsey, respecto de la suspensión de la cuenta de Donald Trump en su red social. En primer lugar, aparece rápidamente una concepción que los propios propietarios, en este caso encarnados en Dorsey, tienen acerca de las “plataformas” como un espacio potencialmente abierto y libre de conversación.

Este concepto se relaciona estrechamente con lo planteado por Gillespie (2010) en cuanto al “potencial democratizador” que la utilización de este término encarnaría. Sin embargo, y a pesar de las promesas, las “plataformas” finalmente terminan siendo “más como los medios tradicionales de lo que pueden admitir” ya que terminan volviendo a poner sobre la mesa clásicas discusiones acerca del libre discurso y la expresión pública (Gillespie, 2010:13).

En relación con esto, se puede recuperar nuevamente la discusión acerca de cómo operan las plataformas en términos de la moderación de contenido. Tal como se planteó en el capítulo anterior, y sobre todo en los Estados Unidos, es baja la carga de responsabilidad legal que puede caberles a las empresas de redes sociales por el contenido que sus usuarios publican.

Por lo tanto, podría resultar contradictorio en primera instancia el hecho de que se hayan tomado decisiones tan fuertes en la moderación del contenido publicado por una persona pública de la relevancia de Donald Trump.

Sobre este punto se puede recuperar una vez más a Napoli y Caplan (2019). Los autores plantean que la moderación de contenidos que las plataformas realizan es “voluntaria” y se relaciona con la doctrina del “buen samaritano”. Así, se trataría de una modalidad de regulación más cercana a medios tradicionales como la televisión.

De este modo, se observa un escenario en el que son las propias plataformas las que pueden decidir (con sus criterios propios) qué contenido quieren regular o no, ya que no existe una obligación de limitar ciertos contenidos, ni tampoco una consecuencia negativa en el caso de no limitarlos.

Como ya se ha visto, Gillespie (2017) precisó también que los intermediarios cuentan con una protección legal para evitar ser responsabilizados por los discursos de sus usuarios, ya que son catalogados como proveedores de acceso a un servicio y no como editores de contenido. Esta protección se extiende tanto para cuando las plataformas deciden no intervenir ante un contenido cuestionado, como para aquellos escenarios en los que sí elijan actuar, por ejemplo, a través de la vigilancia y remoción de contenido.

Recuperando también lo planteado por Lessig (1998), se puede interpretar que hay un escenario que está más cerca de una “arquitectura de control” en la que los gobiernos tienen una menor capacidad de intervención y los mecanismos de dicho control caen del lado de las propias compañías y se materializan en los términos de servicio.

Cabe recordar aquí lo planteado por Becerra y Waisborg (2021) con respecto a la discrecionalidad y la opacidad de las decisiones editoriales de las plataformas. Los autores hacen énfasis en que las empresas no brindan suficientes detalles sobre las medidas que toman, no involucran a actores públicos en sus decisiones, ni toman las mismas acciones o criterios necesariamente ante casos similares.

Si bien en este caso de análisis en particular se podría decir que existieron explicaciones acerca de los motivos que llevaron a la decisión, está claro que en ninguna medida fueron consultadas con otros actores, mucho menos del ámbito público. Es pertinente agregar que las declaraciones se vinculan con la relevancia del caso, pero no sucede tal cosa ante la cantidad de etiquetados, eliminaciones de *tweets* y cuentas que enfrentan otros usuarios.

Esta situación se agrava aún más en un contexto que, tal como se ha visto en apartados anteriores, tiende a la concentración. Más aún, como expresó el propio Jack Dorsey, en este caso en particular se tomaron decisiones muy similares a lo largo de las distintas plataformas.

Se comprenden, entonces, cuáles son algunas de las problemáticas de la regulación privada y concentrada en Internet. Ahora bien, a la hora de preguntarse por las alternativas a esta situación aparecen también algunas propuestas. En el ya citado documento del OBSERVACOM, por ejemplo, se plantea que la regulación pública de las plataformas “debería adoptar medidas antimonopolio para enfrentar el escenario de concentración y ausencia de competencia” (OBSERVACOM, 2020:7).

Otra de las posibles propuestas de modificación que se vislumbran son aquellas que se relacionan con la transparencia a la hora de reducir los ya planteados inconvenientes de opacidad y discrecionalidad en la toma de decisiones. En el texto de Becerra y Waisborg (2021:5) se plantea que una “creciente demanda de transparencia y de normas básicas que reduzcan la discrecionalidad editorial de las plataformas coloca en un lugar central la orfandad de estándares adecuados para su funcionamiento”.

Es importante repasar el hecho de que aquí la transparencia no solamente se relaciona con la comunicación de las decisiones, sino también con el involucramiento de actores de distintas esferas a la hora de la construcción de los términos de servicio que llevan a esas decisiones.

Con respecto al caso analizado en este trabajo, se observa que una propuesta de regulación vino de la mano justamente de Donald Trump. Tal como se trabajó en este capítulo, el diagnóstico de la situación era similar al que se está haciendo referencia y al que el mismo Dorsey comunicó. La orden ejecutiva del entonces presidente remarcaba que las plataformas tienen un poder “inmenso” y “sin precedentes” a la hora de “censurar, borrar o desaparecer información; y controlar lo que la gente ve o no ve”¹¹⁴.

Su propuesta, entonces, era “prevenir la ‘censura en línea’” a través de un cuestionamiento sobre la inmunidad que tienen estas empresas para realizar acciones de moderación de contenidos, que se funda en la ya explicada Sección 230 de la *Communications Decency Act*. En ese mismo documento, Trump hace foco en algunos aspectos que ya se han planteado: “Las plataformas en línea invocan justificaciones inconsistentes, irracionales e infundadas para censurar o restringir el discurso de los estadounidenses”. Nuevamente, se pone en juego la cuestión de la discrecionalidad de los términos de servicio.

Dentro de este análisis ya se ha planteado la particular relación que existía entre Twitter y Trump. Desde antes de asumir la presidencia y durante los primeros años de su mandato, Trump se apoyó en su cuenta oficial de Twitter como medio fundamental para comunicar sus opiniones y actos de gobierno. Así, hasta el momento en que comenzó el

¹¹⁴ Registro Federal de los Estados Unidos (28/05/2020). “Previniendo la censura en línea”. Disponible en: <https://www.federalregister.gov/documents/2020/06/02/2020-12030/preventing-online-censorship> (en inglés - última consulta: 19/12/2021).

conflicto relacionado con el proceso electoral, Trump no se había manifestado en contra del “poder” con el que contaban las plataformas ni de la ya citada sección 230.

Por su parte, Twitter tampoco había planteado ninguna objeción contra los contenidos publicados por Trump antes de esta situación, más allá de que estos hubieran incurrido en contenidos racistas¹¹⁵ o misóginos, que también podrían haber sido interpretados como incitaciones a la violencia.

Como ya se ha planteado, esta orden ejecutiva fue derogada por Joe Biden. Sin embargo, la intención de Trump de limitar la “inmunidad” de estas plataformas continuó a través de la presentación de demandas a las mismas plataformas y a sus propietarios¹¹⁶. Si bien hasta el momento de la escritura de este trabajo no se han producido resultados respecto de estas acciones legales, su repercusión podría ser una cuestión de análisis a futuro.

Es importante aclarar nuevamente que tampoco la solución a este dilema parece encontrarse en una libertad de expresión completamente irrestricta y libre de consecuencias. O bien, como se explicó en el texto de Becerra y Waisborg: “Considerar que cualquier expresión es conducente a buen discurso democrático es ingenuo, idealista (...) y representa una opción facilista y principista que elude los problemas gruesos de una realidad mucho más compleja” (Becerra y Waisborg, 2021:16).

En este punto, se puede retomar una vez más la importancia de que aquellas medidas que efectivamente se tomen, e impliquen una restricción a la libertad de expresión, se basen en estándares democráticos.

Tal como estableció Loreti (2021), se puede considerar que los principios del test tripartito –fundamentales a la hora de pensar en la adopción de restricciones al ejercicio de la libertad de expresión– resultan aplicables también a actores no estatales, principalmente en el caso de las plataformas de contenidos en Internet.

De este modo, las restricciones deberían estar definidas claramente a través de una ley, estar orientadas a objetivos específicos, y ser necesarias, idóneas y proporcionales para

¹¹⁵ Redacción BBC (15/07/2019). “El escándalo por los tuits ‘racistas’ de Trump contra 4 congresistas demócratas”. BBC. Disponible en: <https://www.bbc.com/mundo/noticias-internacional-48983811> (última consulta: 26/12/2021).

¹¹⁶ Colvin, J. y O’Brien, M. (7/7/2021). “Trump demanda a Facebook, Twitter y YouTube”. AP News. Disponible en: <https://apnews.com/article/lawsuits-business-government-and-politics-c7e26858dcb553f92d98706d12ad510c> (en inglés - última consulta: 19/12/2021).

alcanzar estos objetivos. Si estos tres pasos se cumplen, seguramente se tratará de restricciones planteadas de un modo mucho más democrático que aquellas que son definidas exclusivamente a través de los criterios de una empresa privada.

El gran inconveniente parecería estar en el momento en que estas limitaciones (y las consecuencias para aquellos que incurren en faltas) son definidas a través de los términos de servicio de un conjunto reducido de empresas en un contexto de concentración económica. Más aún, con decisiones que, como se vio en el caso de análisis, son muy similares a lo largo de este pequeño espectro de corporaciones.

Aparece, entonces, la necesidad de repensar estos mecanismos. En análisis posteriores seguramente se plantee el interrogante acerca de qué acciones concretas podrán realizarse para pensar en estos modos de regulación más democráticos y transparentes para las plataformas y los discursos que pueden restringirse o no dentro de ellas. Otra cuestión relevante a desarrollar es la forma de implementación de tales reglas, dado el volumen de publicaciones e información que circula por las redes sociales.

Las reflexiones de Jack Dorsey que apuntaban a una mayor transparencia en los mecanismos de toma de decisiones pueden pensarse como un punto de partida. Si bien puede resultar demasiado optimista pensar a Internet como ese espacio global, libre, abierto y de acceso democrático e igualitario, no sería imposible pensar en un espacio en donde lo “decible” no quede a merced de los criterios de una empresa privada. Los riesgos para el ejercicio de la libertad de expresión que implica la autorregulación de las plataformas podrían mitigarse a través de la adopción de mecanismos más democráticos, que involucren a más actores y que dejen de lado la opacidad y la discrecionalidad.

5. Conclusiones

A lo largo del presente trabajo, se ha trabajado sobre el objetivo central de realizar una reflexión acerca de los riesgos para el ejercicio de la libertad de expresión que implica la autorregulación de las plataformas de redes sociales a partir del análisis del bloqueo a la cuenta de Twitter de Donald Trump en enero de 2021.

Para alcanzar este objetivo central, se ha realizado un amplio recorrido a lo largo de diferentes elementos de análisis. En principio, se trabajó con las distintas formas de regulación existentes para los discursos, con el foco puesto en las plataformas de redes sociales en general y en Twitter en particular. A su vez, se analizaron también algunas propuestas relevantes de modificaciones regulatorias que, en el caso de haber prosperado, podrían haber repercutido en cambios importantes.

Dentro de estas regulaciones existentes, la cuestión de la autorregulación por parte de las plataformas se convirtió en un factor fundamental. Por ello, se trabajó específicamente sobre los términos de servicio de las plataformas con el fin de entender su alcance y qué estándares podrían aplicarle si se pensara en un control más democrático de ellos.

En este sentido, también resultó relevante poder comprender hasta qué punto los Estados tienen poder de intervención sobre una cuestión tan sensible como la libertad de expresión. En países como Estados Unidos, se observa que la posibilidad estatal de limitar la expresión y, sobre todo, de actuar frente a actores privados en esta materia resulta escasa.

También, se estudió a lo largo del trabajo la responsabilidad que les puede caber a las plataformas de redes sociales por el contenido que sus usuarios publican en tanto actúan como intermediarias de un servicio. Acerca de esto, se concluyó que, según la normativa y jurisprudencia estadounidense, no solamente cuentan con una protección legal para evitar actuar ante contenidos que puedan ser dañinos, sino que también tienen cierta libertad para restringir los discursos que consideren necesarios sin caer en la categoría de “editores” y, por lo tanto, sin consecuencias legales.

Por último, el trabajo realizado en esta tesina desplegó una sistematización y un análisis de los argumentos esgrimidos por la plataforma de Twitter a la hora de efectivamente

tomar decisiones sobre el contenido. En particular, se focalizó sobre el caso de la limitación de contenidos y la suspensión de la cuenta de Donald Trump con una descripción y una lectura detallada de la argumentación de la empresa con el fin de comprender cómo se justificaban estas decisiones y qué contradicciones existían con los sistemas de regulación vigentes.

De lo trabajado se puede obtener como una primera conclusión que la autorregulación privada de los discursos que realizan las plataformas de redes sociales efectivamente implica riesgos para la libertad de expresión. Algunos de estos peligros se relacionan con la discrecionalidad y la falta de transparencia en la aplicación de los criterios restrictivos, pero también con el hecho de que los términos de servicio –que rigen el funcionamiento en las plataformas– se redacten con una nula participación de actores ajenos a las empresas.

Si bien se podría pensar que la falta de intervención de otros actores de la sociedad civil es lógica al tratarse de una empresa privada –y más aún bajo un marco legal poco restrictivo ante la libertad de expresión como el de los Estados Unidos– esto resulta cuestionable ante ciertos escenarios.

Tomando como ejemplo paradigmático el caso analizado en esta tesina, el hecho de que la persona involucrada en una restricción sea un actor de indudable relevancia pública como lo fue el presidente en ejercicio de los Estados Unidos no es un factor que pueda soslayarse. Cuando se trata de estos actores, y más aún en el contexto de un disputado proceso electoral, el hecho de que una empresa privada pueda unilateralmente prohibirles el acceso a la expresión significa un indudable peligro para el acceso a la información de interés público.

Más allá de tratarse de la restricción de una vía para que el presidente en ejercicio pueda comunicarse con los ciudadanos de su país en un contexto de crisis, una decisión de este calibre implica también una dificultad en el acceso para los ciudadanos mismos que utilizaban las redes sociales como fuente de acceso a la información de interés público.

Sin embargo, otra cuestión que surgió del análisis realizado es que estos planteos no significan que la solución a este dilema se encuentre en permitir la libre circulación –sin consecuencias– de discursos que pueden ser calificados como de odio o de incitación a hechos de violencia. Resulta evidente que este escenario también implicaría numerosos

riesgos para el funcionamiento democrático de una sociedad y que, por lo tanto, el hecho de que existan consecuencias para quienes emiten estos discursos es importante.

Así, es necesario que el espacio para estas consecuencias exista, pero los mecanismos en que éstas se plantean y efectivamente ocurren deberían ser los que se transformen hacia modos más democráticos. Basándose en los citados estándares internacionales, surge la necesidad de que toda posible restricción a la libertad de expresión sea previamente estipulada a través de una ley; que las posibles restricciones se realicen con un fin legítimo; y que las medidas que se tomen sean idóneas, necesarias y proporcionales a este fin establecido.

Con estos principios, se propone reducir los actos de censura previa o de limitación discrecional de contenidos por parte de las plataformas. A su vez, esto tampoco implicaría adosarles una responsabilidad objetiva a las empresas por el contenido que sus usuarios publican, sino una mayor apertura y transparencia en sus mecanismos de toma de decisiones, sobre todo ante casos como los descriptos.

En futuros estudios se podrá reflexionar acerca de las maneras concretas en que estas modificaciones pueden tener lugar, qué actores deberían estar involucrados en estas transformaciones y cómo pueden ser llevadas a cabo en los distintos lugares del mundo ya que, como es lógico, los marcos regulatorios varían en cada geografía, como así también la cultura y los contextos a considerar en la implementación.

Con respecto al caso de Trump, como se ha visto, la continuidad de sus reclamos para con las plataformas se dio por la vía judicial. Si bien hasta el momento de la presentación de este trabajo no han surgido conclusiones de estos procesos, resultará interesante también a futuro analizar los argumentos que se esgriman para las decisiones que se tomen.

Post scriptum

Más allá de la vía judicial, Donald Trump también tomó la decisión de lanzar su propia red social¹¹⁸ llamada “*Truth Social*” (que podría traducirse como “red social de la verdad”, o bien como “verdad social”). Según se presenta en la página oficial de la red (truthsocial.com), esta plataforma “alienta una conversación global, abierta, libre y honesta, sin discriminaciones por ideología política”¹¹⁹. Al momento de la escritura de este trabajo, la plataforma se encontraba aún en etapa de lanzamiento. Sin embargo, su apertura al público en general podría disparar también consecuencias interesantes para analizar.

Resulta interesante una última reflexión acerca de los términos de servicio de esta nueva red social. Por ejemplo, una lectura de ellos permite darse cuenta de que la plataforma se reserva el derecho de: “a nuestro exclusivo criterio y sin limitación, rechazar, restringir el acceso, limitar la disponibilidad o deshabilitar (en la medida que sea tecnológicamente factible) su acceso o cualquiera de sus contribuciones”¹²⁰.

Por otra parte, *Truth Social* también se resguarda en las protecciones de la citada Sección 230 de *Communications Decency Act* (CDA) [Ley de Decencia en las Comunicaciones] ya que aclaran que no se consideran responsables por ningún contenido publicado por terceros en la plataforma.

Así, basándose en cómo están estructurados sus términos de servicio, se puede aventurar que la nueva “conversación” que propone la plataforma no diferiría demasiado de lo que se encuentra en otras plataformas. Sin embargo, sí permitiría el retorno a las redes sociales de Donald Trump, lo cual podría aportar nuevos materiales de análisis en relación con esta problemática.

¹¹⁸ Furlong, S. (25/10/2021). “«Follow the truth»: expectativas y límites de la red social de Trump”. *El Grito del Sur*. Disponible en: <https://elgritodelsur.com.ar/2021/10/follow-the-truth-expectativas-limites-red-social-trump.html> (última consulta: 4/1/2022).

¹¹⁹ Traducción propia del material disponible en el sitio web oficial de *Truth Social*: <https://truthsocial.com/> (última consulta: 4/1/2022).

¹²⁰ Traducción propia del punto 11.3 de los términos de servicio de *Truth Social*. Texto completo disponible en: <https://truthsocial.com/terms-of-service/> (en inglés - última consulta: 4/1/2022).

6. Bibliografía

Abramovich, V (2021). *Dilemas jurídicos en la restricción de los discursos de odio*. En: Abramovich, V.; Guembe, M. J.; Capurro Robles, M. (coordinación). *El límite democrático de las expresiones de odio*. Editorial Teseo. Buenos Aires. Disponible en: <https://www.teseopress.com/ellimitedemocraticodelasexpresionesdeodio> (última consulta: 27/11/2021)

Access Now (2020). *26 recommendations on content governance. A guide for lawmakers, regulators and company policy makers*. Disponible en: <https://www.accessnow.org/cms/assets/uploads/2020/03/Recommendations-On-Content-Governance-digital.pdf> (última consulta: 22/11/2021)

Álvarez Ugarte, R. y Rabinovich, E. (2013). *Los intermediarios y los desafíos para la libertad de expresión en Internet*, en Cuestión de derechos nº4, pp. 115-124. Disponible en: <https://www.apc.org/sites/default/files/ADC%20-%20Cuestion%20de%20derechos%20-%20Revista-numero4%20-%202013.pdf> (última consulta: 20/10/2021)

Baladron, M. (2018). *Infraestructura y plataformas de internet: concentración en el ecosistema digital*. REVCOM; no. 6. Red de Carreras de Comunicación Social y Periodismo (REDCOM). Páginas 32-44. Recuperado de: <http://sedici.unlp.edu.ar/handle/10915/69355> (última consulta: 24/11/2021)

Baladron, M. y Califano, B. (2011). *¿Quién controla Internet? Gobernanza, políticas y desafíos para el futuro de la red de redes*. AVATARES de la comunicación y la cultura, Nº 2. ISSN 1853-5925. Agosto 2011.

Becerra, M. y Waisbord, S. (2021). *La necesidad de repensar la ortodoxia de la libertad de expresión en la comunicación digital*. Desarrollo Económico. Revista de Ciencias Sociales | Vol. 60 - Nº 232 - pp. 295-313 | ISSN 1853-8185

Bernal Cano, N. (2013). *Algunas reflexiones sobre el valor de la jurisprudencia como fuente creadora de derecho*. Cuestiones Constitucionales. Revista Mexicana de Derecho Constitucional. Núm 28, enero-junio 2013. Ciudad de México. ISSN 1405-9193 Disponible en: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-91932013000100012 (última consulta: 28/11/2021).

Bustos Martínez, L.; De Santiago Ortega, P.; Martínez Miró, M.A. y Rengifo Hidalgo, M. (2019). *Discursos de odio: una epidemia que se propaga en la red. Estado de la cuestión sobre el racismo y la xenofobia en las redes sociales*. En Mediaciones Sociales, (18), 25-42. Disponible en: <http://dx.doi.org/10.5209/MESO.64527> (última consulta: 20/11/2021).

Cortés, C. (2014). *La gobernanza de Internet. La trampa de las formas*. Centro de Estudios en Libertad de Expresión y Acceso a la Información (CELE). Universidad de Palermo. Recuperado de:

https://www.palermo.edu/Archivos_content/2014/derecho/agosto/internet/CELE_GobernanzaDeInternet.pdf (última consulta: 10/09/2021).

Cabo Isasi, A. y García Juanatey, A. (2016). *El discurso del odio en las redes sociales: Un estado de la cuestión*. Barcelona: Ajuntament de Barcelona. Disponible en: http://ajuntament.barcelona.cat/bcnvsodi/wp-content/uploads/2017/02/Informe_discurso-del-odio_ES.pdf (última consulta: 28/11/2021).

Díaz Hernández, M. (2020). *Discurso de odio en América Latina*. Derechos Digitales, Febrero 2020. Disponible en: <https://www.derechosdigitales.org/wp-content/uploads/discurso-de-odio-latam.pdf> (última consulta: 22/11/2021).

Gillespie, T. (2017). *Governance of and by platforms*. SAGE Handbook of Social Media, edited by Jean Burgess, Thomas Poell, and Alice Marwick Sage. Disponible en: <https://culturedigitally.org/wp-content/uploads/2016/06/Gillespie-Governance-ofby-Platforms-PREPRINT.pdf> (última consulta: 28/11/2021).

Gillespie, T. (2010). *The Politics of 'Platforms'*. New Media & Society 12, no. 3 <http://nms.sagepub.com/cgi/content/abstract/12/3/347> (última consulta: 8/11/2021).

Gottfried, J. y Shearer, E. (2016). *News use across social media platforms 2016*. Pew Research Center. Recuperado de <http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/> (última consulta: 8/11/2021).

Justel-Vázquez, S.; Fernández-Planells, A.; Victoria-Mas, M. y Lacasa-Mas, I. (2018). *Twitter e información política en la prensa digital: la red social como fuente de declaraciones en la era Trump*. El profesional de la información, 2018, Septiembre-Octubre, v. 27, n. 5. eISSN: 1699-2407. Disponible en: <http://repositori.uic.es/handle/20.500.12328/1362> (última consulta: 6/11/2021).

Kurbalija, J. (2016). *Introducción a la gobernanza de Internet*. DiploFoundation, Malta.

Lessig, L. (1998). *Las leyes del ciberespacio*, conferencia Taiwan Net'98, mimeo, Taipei.

Loreti, D. (2021). *Redes sociales, plataformas digitales y libertad de expresión: obligaciones de los actores no estatales*. Observatorio Latinoamericano de Regulación de Medios y Convergencia (OBSERVACOM). Diciembre de 2021. Recuperado de: <https://www.observacom.org/wp-content/uploads/2021/12/Redes-sociales-plataformas-digitales-libertad-de-expresion.-Obligaciones-de-los-actores-no-estatales.-D.-Loreti.pdf> (última consulta: 29/12/2021).

- Loreti, D. y Lozano, L. (2021). *La tensión entre la libertad de expresión y la protección contra la discriminación*. En: Abramovich, V.; Guembe, M. J.; Capurro Robles, M. (coordinación). *El límite democrático de las expresiones de odio*. Editorial Teseo. Buenos Aires. Disponible en: <https://www.teseopress.com/ellimitedemocraticodelasexpresionesdeodio> (última consulta: 27/11/2021)
- Marradi, A.; Archenti, N. y Piovani, J. I. (2007). *Metodología de las Ciencias Sociales*. Buenos Aires: Emecé.
- Meléndez Juarbe, H. (2012). *Intermediarios y libertad de expresión: apuntes para una conversación*. En “Hacia una Internet libre de censura. Propuestas para América Latina”, Eduardo Bertoni, Comp. Ed. Universidad de Palermo. Pp. 110-124. Disponible en: https://www.palermo.edu/cele/pdf/internet_libre_de_censura_libro.pdf (última consulta: 9/12/2021).
- Napoli, P. y Caplan, R. (2019). *Por qué las empresas de medios insisten en que no son empresas de medios, por qué están equivocadas y por qué es importante* (Traducción de Mariela Baladrón, Ana Bizberge, Bernadette Califano, Cecilia Fariña y Guillermo Mastrini). Revista Hipertextos 12(7), pp. 15-46. DOI: <https://doi.org/10.24215/23143924e002> (última consulta: 19/12/2021).
- Pauner Chulvi, C. (2018). *Noticias falsas y libertad de expresión e información. El control de los contenidos informativos en la red*. Teoría y Realidad Constitucional, 41, 297-318
- Pérez-Curiel, C. y Limón Naharro, P. (2019). *Influencers de la Política. Estudio de la marca personal de Donald Trump en Twitter y efectos en medios y usuarios*. ISSN 2386-7876 – Communication & Society, 32(1), 57-76.
- Principios de Manila Sobre Responsabilidad de los Intermediarios (2015). Version 1.0, 24 de marzo de 2015. Disponible en: https://www.eff.org/files/2015/06/23/manila_principles_1.0_es.pdf (última consulta: 4/11/2021).
- Schulz, W. (2018). *Regulating Intermediaries to Protect Privacy Online – the Case of the German NetzDG*. HIIG Discussion Paper Series 2018-01. Disponible en: <https://www.hiig.de/wp-content/uploads/2018/07/SSRN-id3216572.pdf> (última consulta: 23/11/2021)
- Stake, R. E. (1999). *Investigación con estudio de casos* (2.a ed.). Madrid: Morata.
- Stolee, G. y Caton, S. (2018). *Twitter, Trump, and the Base: A Shift to a New Form of Presidential Talk?* Signs and Society, vol. 6, no. 1. Semiosis Research Center at Hankuk University of Foreign Studies. Disponible en:

- <https://www.journals.uchicago.edu/doi/full/10.1086/694755> (última consulta: 22/11/2021)
- Van Dijck, J. (2016). *La cultura de la conectividad: una historia crítica de las redes sociales*. Buenos Aires: Siglo XXI.
- Yin, R. (1994). *Case Study Research: Design and Methods*. Sage Publications, Thousand Oaks, CA.
- Center for International Media Assistance (2017). *Estándares internacionales de libertad de expresión: Guía básica para operadores de justicia en América Latina*. Recuperado de: <https://www.corteidh.or.cr/tablas/r37048.pdf> (última consulta: 2/10/2021)
- Convención Americana sobre Derechos Humanos (1969). Artículo 13. Recuperado de: <http://www.oas.org/es/cidh/expresion/showarticle.asp?artID=25&IID=2> (última consulta: 5/10/2021)
- Observatorio Latinoamericano de Regulación de Medios y Convergencia – OBSERVACOM (2020). *Estándares para una regulación democrática de las grandes plataformas que garantice la libertad de expresión en línea y una Internet libre y abierta*. Julio de 2020. Recuperado de: <https://www.observacom.org/wp-content/uploads/2020/09/Estandares-para-una-regulacion-democratica-de-las-grandes-plataformas.pdf> (última consulta: 3/1/2022)
- Organización de Estados Americanos (2010). *Marco Jurídico Interamericano sobre el Derecho a la Libertad de Expresión*. Relatoría Especial para la Libertad de Expresión, Comisión Interamericana de Derechos Humanos. Disponible en: http://www.oas.org/es/cidh/expresion/docs/cd/sistema_interamericano_de_derechos_humanos/index_MJIAS.html (última consulta: 24/11/2021)
- Organización de Estados Americanos (2011). *Declaración Conjunta sobre Libertad de Expresión e Internet*. Recuperado de: <http://www.oas.org/es/cidh/expresion/showarticle.asp?artID=849> (última consulta: 22/11/2021)
- Organización de Estados Americanos (2021). *La CIDH advierte un punto de inflexión de la libertad de expresión en internet y convoca a diálogo en la región*. Recuperado de: <http://www.oas.org/es/cidh/jsForm/?File=/es/cidh/prensa/comunicados/2021/026.asp> (última consulta: 14/10/2021)
- Organización de las Naciones Unidas (1948). *Declaración Universal de Derechos Humanos*. Asamblea General de las Naciones Unidas, 26 de junio de 1945. Recuperada de: <http://www.ohchr.org/Documents/Publications/ABCan nexessp.pdf> (última consulta: 14/10/2021)

Organización de las Naciones Unidas (2017). *Plan de acción para Líderes y Actores Religiosos para la Prevención de la Incitación a la Violencia que Podría Dar Lugar a Crímenes Atroces*. Recuperado de:

https://www.un.org/en/genocideprevention/documents/publications-and-resources/UN%20Plan%20of%20Action_ES.pdf (última consulta: 15/10/2021)

Organización de las Naciones Unidas (2019). *La estrategia y plan de acción de las Naciones Unidas para la lucha contra el discurso de odio*. Recuperado de:

https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_ES.pdf (última consulta: 3/1/2022)

Relatoría Especial para la Libertad de Expresión (2020). *Declaración Conjunta sobre Libertad de Expresión y Elecciones en la Era Digital*. Recuperado de:

<https://www.oas.org/es/cidh/expresion/showarticle.asp?artID=1174&IID=2> (última consulta: 5/12/2021)