

## ¿Libertad para tuitear?<sup>1</sup>

Jason Thacker<sup>1</sup>

Traducido por Ruth Abril

### Moderación de contenidos, libertad religiosa y el espacio público digital

“Si la libertad religiosa se defiende únicamente por razones pragmáticas, puede sacrificarse y se sacrificará a la conveniencia”.<sup>2</sup> Esas palabras, pronunciadas en 1983 por el fallecido teólogo cristiano evangélico Carl F. H. Henry, fueron proféticas. Presagiaban gran parte de lo que vendría después de su muerte en 2003, con la llegada de una nueva era digital marcada por el meteórico ascenso de las redes sociales como uno de los principales canales de comunicación de la sociedad.

Y aunque las amenazas tradicionales a la libertad religiosa y de expresión siguen siendo inquietantemente frecuentes en gran parte del mundo, esta revolución de las redes sociales plantea una nueva y apremiante cuestión, que es cómo podemos preservar las libertades religiosas y de expresión en una sociedad cada vez más digital, en la que los gobiernos han pasado a un segundo plano frente a las empresas tecnológicas transnacionales que ahora ejercen una influencia desmesurada en la plaza pública digital.

Lo que comenzó como una forma de conectar con los demás y compartir información inocua se ha convertido en una de las formas más importantes de comunicarnos y promover el bien común. Al mismo tiempo, nuestra plaza pública se ha desconectado cada vez más de un marco moral trascendente y se ha aliado, en cambio, con la desenfrenada autonomía moral de nuestra era secular. En este nuevo entorno, las personas de fe deben analizar detenidamente cómo se expresan

<sup>1</sup> Publicado originalmente en el número de mayo/junio de 2022 de la revista *Liberty*.

<sup>2</sup> Carl F. H. Henry, *The Christian mindset in a secular society: Promoting evangelical renewal and national righteousness* (Portland, OR: Multnomah Press, 1984), 65.

las convicciones morales en el espacio público actual y trabajar para preservar una base sólida de libertad religiosa para las generaciones futuras.

### **Las grandes tecnologías y el nuevo espacio público**

La preocupación por la enorme influencia de las empresas tecnológicas en nuestro discurso público es uno de los escasos puntos de acuerdo bipartidistas en la sociedad estadounidense actual. Pero hay poco acuerdo sobre los detalles.

Los progresistas abogan tradicionalmente por una mayor moderación de los contenidos, especialmente ante la creciente influencia de la desinformación, las noticias falsas y la incitación al odio a través de internet. Sostienen que hay que hacer más para frenar estos peligros y que si las grandes empresas tecnológicas no están dispuestas a asumir el reto, el gobierno debe intervenir.

Los conservadores, por su parte, han abogado por una menor moderación debido a la noción de que el discurso y los valores conservadores han sido injustamente eliminados o suprimidos. Citan casos de usuarios a los que se ha prohibido el acceso a redes sociales o incluso a plataformas y que se han cerrado por completo, simplemente por la agenda ideológica predominante de Silicon Valley. Asimismo, argumentan que si las grandes tecnológicas no están a la altura de las circunstancias, el gobierno tendrá que intervenir para regular esta influyente industria.

Estos debates se clasifican a menudo bajo el apelativo de “grandes tecnológicas”, con lo que se pretende significar la influencia desproporcionada y la ubicuidad de estas plataformas mediáticas en el espacio público.

Sin embargo, el término no tiene en cuenta a algunas de las mayores empresas de Estados Unidos de Norteamérica como Microsoft, Disney, Comcast, Verizon y otras. Además, se centran en las empresas estadounidenses y se excluye a gigantes de la tecnología y los medios

de comunicación como Tencent y Alibaba de China, que tienen un historial extremadamente preocupante en materia de libertad de expresión y de culto debido a la injerencia del gobierno del Partido Comunista Chino.<sup>3</sup>

El término pretende incluir específicamente a empresas como Meta (Facebook), Alphabet (Google/YouTube) y Amazon, así como también a empresas con bases de usuarios mucho más pequeñas, pero con una enorme influencia en el espacio público digital como lo es Twitter (actualmente “X”). Una de las principales formas en que estas empresas ejercen tal influencia sobre la conversación y la opinión públicas hoy en día es a través de las formas en que deciden moderar el contenido de los usuarios en sus plataformas.

### **El propósito de la moderación de contenidos**

La moderación de contenidos es un trabajo difícil para cualquier empresa de redes sociales. Cada día se comparten millones de mensajes en estas plataformas. La mayoría son de naturaleza benigna, pero a veces individuos y organizaciones comparten o promueven contenidos abusivos, odiosos o violentos. La mayoría de las empresas de redes sociales esperan que sus usuarios participen en estas plataformas dentro de un cierto conjunto de reglas o normas comunitarias. Estas políticas de contenidos suelen elaborarse con cuidado y reflejan la importancia de la tarea de ofrecer un lugar seguro y apropiado a los usuarios. Se trata de un ejercicio éticamente espinoso, no solo porque las redes sociales desempeñan un papel tan crucial en nuestra sociedad diversa, sino también por la hiperpolitización de muchas de las cuestiones que rodean la expresión en línea.

Durante los últimos años, las prácticas de moderación de contenidos han sido objeto de un intenso escrutinio debido a la amplitud de las propias políticas, así como también a la mala aplicación —o, más exactamente, la aplicación incoherente— de estas normas de

---

<sup>3</sup> Jason Thacker, “Wired for tyranny?,” *Liberty*, septiembre/octubre de 2021, 4.

conducta en línea. Ya se trate de impedir que grupos terroristas como el ISIS o regímenes autoritarios como el Partido Comunista Chino utilicen estas plataformas para la propaganda masiva, o incluso de prohibir el acceso a un expresidente de Estados Unidos, las decisiones de moderación adoptadas por estas empresas influyen en la conversación pública de innumerables maneras, tanto para bien como para mal.<sup>4</sup>

Una de las preguntas más frecuentes que escucho de quienes se preocupan por la moderación de contenidos es si las empresas tecnológicas deberían moderarlos. Algunos argumentan que la moderación es intrínsecamente contraria a la libertad de expresión porque tenemos derecho a expresarnos de la forma que creamos conveniente. Aunque la libertad de expresión es fundamental para nuestro experimento democrático, la cuestión es un poco más compleja de lo que puede parecer a primera vista. Es importante reconocer que hay una diferencia entre censurar un discurso que es desagradable y limitar un discurso que fomenta o glorifica la violencia física, o un discurso que es de naturaleza ilícita, sensual o explotadora, o uno que promueve la delincuencia.

En realidad, las prácticas de moderación de contenidos están fomentadas por el artículo 230 de la Communications Decency Act (Ley en materia de decencia en las comunicaciones) del año 1996, que fue una legislación bipartidista sancionada a mediados de los noventa y diseñada para promover el crecimiento de la incipiente internet. El artículo 230 otorga a las empresas de Internet un escudo de responsabilidad por los contenidos en línea generados por los usuarios, lo que significa que los usuarios, y no las propias plataformas, son responsables del contenido de las publicaciones. La esperanza era que estas empresas adoptaran medidas de “buena fe” para eliminar contenidos censurables con el fin de hacer de Internet un lugar más seguro para nuestra sociedad.<sup>5</sup>

<sup>4</sup> Jason Thacker y Joshua B. Wester, “Understanding Twitter suspensions and the need for consistent policies”, *ERLC*, 17 de febrero de 2021, <https://bit.ly/3sqDkNT>.

<sup>5</sup> Jeff Kosseff, *The twenty-six words that created the Internet* (Ithaca, NY: Cornell University Press, 2019).

Estas medidas de “buena fe” están diseñadas para crear entornos en línea más seguros para todos los usuarios. El debate sobre la moderación de contenidos a menudo se centra en lo que estas medidas deberían implicar exactamente, no en la presencia de las medidas en primer lugar. Una plataforma de Internet o de redes sociales sin ningún tipo de moderación o normas se convertiría rápidamente en un entorno peligroso lleno de desinformación y de un sinfín de contenidos ilegales o sin filtrar. Incluso con normas de moderación de contenidos en vigor, es innegable que las redes sociales se han utilizado de forma que han provocado daños en el mundo real.

Sin embargo, aunque la moderación de contenidos sea clave para mantener un espacio público digital seguro y saludable, también se utiliza cada vez más para suprimir ciertos puntos de vista considerados indignos por el tribunal de la opinión pública, puntos de vista que a menudo tocan temas centrales para la tradición moral cristiana, como son la sexualidad humana y la libertad religiosa. Y estas decisiones de moderación de contenidos a menudo se basan en una interpretación mal definida de lo que constituye incitación al odio.

### Discurso de odio en línea

Otra pregunta que me hacen a menudo es cómo, exactamente, las políticas de moderación de contenidos limitan la expresión religiosa. Es una pregunta que a menudo se formula a la luz de una autonomía moral desenfrenada que defiende la libre expresión solo para lo que es popular o se considera justo en el sentido secular. Las ideas contrarias a esta ortodoxia secular suelen considerarse odiosas y retrógradas, no aptas para ser expresadas en la plaza pública. El discurso del odio se ha convertido así en un ámbito de debate público de enorme trascendencia en el que la expresión religiosa se considera cada vez más intrínsecamente odiosa y perjudicial para el discurso civil.

Aunque muchas empresas tecnológicas se remiten a las normas internacionales cuando tratan temas controvertidos —incluida la de los derechos humanos—, cabe señalar que el discurso de odio a menudo

se deja sin definir en términos legales debido a la profunda tensión que existe entre el discurso de odio y la libertad de expresión. El propio plan de acción de las Naciones Unidas sobre el discurso de odio de mayo de 2019 lo deja claro al decir: “No existe una definición jurídica internacional del discurso de odio y la caracterización de lo que es ‘odioso’ es controvertida y discutida.” Mientras que la Organización de las Naciones Unidas deja el discurso de odio sin definir, claramente desea protecciones robustas contra el discurso de odio y lo llama “una amenaza para los valores democráticos, la estabilidad social y la paz” que “debe [ser confrontado]...en todo momento”.<sup>6</sup>

Del mismo modo, en los Estados Unidos no existe una definición legal del discurso de odio y la Corte Suprema ha afirmado sistemáticamente que el discurso de odio está protegido por la Primera Enmienda. De acuerdo con la Asociación Americana de Bibliotecas, “según la jurisprudencia actual de la Primera Enmienda, la incitación al odio sólo puede penalizarse cuando incita directamente a una actividad delictiva inminente o consiste en amenazas específicas de violencia dirigidas contra una persona o grupo”.<sup>7</sup>

Definir la incitación al odio es una tarea siempre difícil, que se ha vuelto aún más compleja con el auge de la expresión en línea a través de las plataformas de redes sociales.<sup>8</sup> Hay constantes debates en la sociedad y en el mundo académico sobre qué constituye realmente incitación al odio y si la etiqueta debe limitarse a la expresión que incita a la violencia física o al daño. Muchas empresas como Meta y Twitter (actualmente X) han definido la incitación al odio en sentido amplio, un enfoque que necesariamente infringe la libertad de expresión y la libertad religiosa en relación con algunas de las cuestiones más polémicas de nuestros días, a saber, la sexualidad humana y el matrimonio.

---

<sup>6</sup> United Nations strategy and plan of action on hate speech (ONU, mayo de 2019), <https://bit.ly/3hpjYT8>.

<sup>7</sup> “Hate Speech and Hate Crime”, *Advocacy, Legislation and Issues*, 12 de diciembre de 2017, <https://bit.ly/36Bc86u>.

<sup>8</sup> Jason Thacker, “Where do we draw the line on hate speech?”, *ERLC*, 9 de agosto de 2021, <https://bit.ly/3M78TEg>.

La mayoría de la gente estaría de acuerdo con que muchas de las categorías prohibidas por Twitter —incluidas las amenazas de violencia física, “desear, esperar o pedir un daño grave a una persona o grupo de personas” y “referencias a asesinatos en masa, sucesos violentos o medios específicos de violencia en los que grupos protegidos hayan sido los principales objetivos o víctimas”— están claramente dentro de la moderación de contenidos de buena fe.<sup>9</sup>

Los cristianos, en particular, pueden afirmar estas directrices por su creencia en el valor innato y la dignidad de todas las personas como creadas a imagen de Dios y la libertad de conciencia que se deriva de su comprensión de la *imago Dei* (imagen de Dios) (Gn 1,26-28). Pero cuando el discurso de odio se amplía para incluir el discurso que hace que uno se sienta incómodo o que simplemente no le guste, hemos sentado un precedente peligroso para el discurso público y para el futuro del discurso religioso en línea.

### Libertad religiosa y tecnología digital

En octubre de 2020, comenzó a funcionar la Junta de Supervisión creada por Meta (Facebook). El consejo fue diseñado como un mecanismo de revisión para que los usuarios apelen las decisiones de moderación de contenidos y un lugar para que Meta refiera algunas de sus decisiones de contenido más difíciles. El propósito declarado de la junta es ayudar a Meta a “responder a algunas de las preguntas más difíciles en torno a la libertad de expresión en línea: qué retirar, qué dejar y por qué” a través de un procedimiento independiente.<sup>10</sup> Este Consejo —que puede tener hasta cuarenta miembros de todo el mundo, incluidos cuatro copresidentes— se ha hecho cargo de una serie de casos en los últimos años, especialmente relacionados con la libertad de expresión y la libertad religiosa en el extranjero.

<sup>9</sup> X Help Center, “Hateful conduct”, <https://bit.ly/3ItQkLe>.

<sup>10</sup> Obtenga más información sobre el Consejo de Supervisión en <https://oversightboard.com/>.

El número desmesurado de casos relacionados con la libertad religiosa no solo habla de la gravedad de la toma de decisiones en materia de moderación hoy en día, sino también de la complejidad de navegar por la expresión religiosa en línea en un entorno transnacional. Aunque algunas empresas tienen un historial mejor —aunque no perfecto— que otras a la hora de proteger la expresión religiosa, otras, como Twitter (actualmente X), han impedido sistemáticamente que los creyentes expresen libremente sus creencias, especialmente en cuestiones de sexualidad humana y género.<sup>11</sup>

La mayoría de las políticas tecnológicas de moderación de contenidos no mencionan explícitamente la libertad religiosa como objetivo de la libertad de expresión, aunque hablen de un sólido compromiso con la diversidad y la inclusión. A menudo, la religión solamente se menciona en términos de discurso de odio basado en la afiliación religiosa de alguien, pero la expresión religiosa no se considera un valor fundamental de muchas empresas. Sin embargo, la verdadera inclusión y la diversidad debe ser para todas las personas, no solo para las que tienen opiniones populares o ampliamente aceptadas sobre algunas de las cuestiones más importantes de la actualidad.

La libertad religiosa no es solo la libertad de creer o rendir culto, sino la libertad de vivir de acuerdo con esas creencias en todos los ámbitos de la vida. La limitación de la expresión religiosa está cada vez más normalizada en nuestra sociedad secular. Las personas de fe deben ser conscientes del cambiante diálogo en torno a la expresión religiosa, especialmente en lo que se refiere a nuestro entorno digital, donde gran parte de la comunicación y las conversaciones tienen lugar en plataformas sociales privadas.

La fe por naturaleza es pública y, como escribió el difunto Richard John Neuhaus, no existe una verdadera “plaza pública desnuda”,<sup>12</sup> desprovista de matices religiosos y compromisos ideológicos, incluso por

---

<sup>11</sup> Jason Thacker, “Is content moderation stifling public discourse?”, *ERLC*, 5 de mayo de 2021, <https://bit.ly/3BUNnxX>.

<sup>12</sup> Julio Agustín Sánchez León, “Religión, sociedad y vida pública en Richard John Neuhaus”, *Excerpta e Dissertationibus in Sacra Theologia* 68 (2019): 411.



parte de sus miembros más laicos. Todo el mundo lleva sus creencias al espacio público y suprimir la expresión religiosa es contrario a la naturaleza de la libertad de expresión y la dignidad humana. No importa lo que digan algunos, las personas de fe no pueden dejar sus creencias en la puerta ni actuar como si su fe no influyera en su vida pública.

Algunas personas de fe han sido criticadas con razón por adoptar un enfoque excluyente de la libertad religiosa, un enfoque que dice que “la libertad religiosa [es] para mí, pero no para ti”. Una crítica similar puede dirigirse a quienes valoran la libertad de expresión para las ideas socialmente aceptables, pero no para las que se consideran indignas de expresión, como la ética sexual cristiana tradicional.

Carl F. H. Henry dijo una vez: “No es el papel del gobierno juzgar entre sistemas rivales de metafísica y legislar uno sobre otros; más bien su papel es proteger y preservar un curso libre para sus garantías constitucionales”.<sup>13</sup> Si bien hay complejidades en la aplicación de este enfoque en la plaza pública digital, que está gobernada por empresas privadas, los principios de la libertad de expresión y la libertad religiosa siguen siendo vitales para una plaza pública robusta y saludable. Con la enorme influencia de las grandes empresas tecnológicas en el discurso público, especialmente en un momento de tanta polarización y división, necesitamos un enfoque verdaderamente inclusivo, un enfoque que no suprima los puntos de vista religiosos en un irónico llamamiento a la tolerancia, la diversidad y la inclusión.

---

<sup>13</sup> Henry, *The Christian mindset in a secular society*, 80.