

## ENTRE ANALOGÍAS Y METÁFORAS: EL DEBATE SOBRE LA MODERACIÓN DE CONTENIDOS EN LAS REDES SOCIALES

BETWEEN ANALOGIES AND METAPHORS: THE DEBATE ON  
CONTENT MODERATION ON SOCIAL NETWORKS

Pedro José PEÑA JIMÉNEZ  
Letrado de las Cortes Generales  
<https://orcid.org/0000-0003-2722-2685>

Fecha de recepción del artículo: noviembre 2021

Fecha de aceptación y versión final: diciembre 2021

### RESUMEN

*El artículo trata sobre la moderación de contenidos, que es una actividad esencial, pero poco conocida y publicitada, de las redes sociales, y que afecta a valores, bienes públicos y derechos, como la libertad de expresión e información, la privacidad o la seguridad. El marco legal de la misma, que proviene de la última década del siglo pasado, ha hecho posible, con la concesión de un régimen general de amplia inmunidad a los intermediarios que operan en Internet, el nacimiento de las redes sociales y su desarrollo hasta alcanzar la dimensión que ahora tienen. Ese marco está hoy sujeto a diversas iniciativas legislativas de revisión, tanto en Estados Unidos como en la Unión Europea, que probablemente saldrán adelante en el corto plazo. En el origen, formalización y discusión de esas iniciativas ha tenido una influencia innegable un debate, nutrido por aportaciones muy diversas y donde ha sido frecuente el recurso a analogías y metáforas para sustentar distintas posiciones sobre el futuro de la moderación y la regulación de las redes sociales. En el artículo se exponen esas posiciones, que van desde un escepticismo radical sobre las opciones de reforma hasta el mantenimiento de una confianza casi inquebrantable en la capacidad de autorregulación de las empresas, pasando por distintas alternativas de mayor o menor intervención pública, que es precisamente donde se sitúan, en lo fundamental, las iniciativas legislativas en tramitación. La conclusión es que*

*esa orientación es correcta, que el Estado democrático no debe adoptar un papel intervencionista en la gestión de las redes sociales, pero sí ha de asegurar, mediante la aprobación de normas legales ajustadas a la realidad tecnológica y social, que las redes actúan en el mercado como lo que son, empresas, sujetas a las normas y de forma responsable y transparente.*

*Palabras clave: Redes sociales, libertad de expresión, regulación, no discriminación, intervención estatal, proceso democrático, participación, deliberación.*

### ABSTRACT

*The article is about content moderation, which is an essential activity, but little known and publicized of social networks, and which affects values, public goods and rights, such as freedom of expression and information, privacy or security. Its legal framework which comes from the last decade of 20th century, has made possible, with the granting of a general regime of broad immunity to intermediaries operating on the Internet the birth of social media and their development until reaching the dimension they now have. This framework is today subject to various initiatives for revision, both in the United States and in the European Union, which will probably come forward in short term. A debate on the subject has had an undeniable influence on the origin, formalization and discussion of these initiatives, nourished by the very diverse contributions and where the use of analogies and metaphors has been frequent to support different positions on the future of moderation and regulation of social networks. The article sets out these positions, which range from a radical skepticism about the reform options to the maintenance of an almost unshakable confidence in self-regulation of the companies, going through different alternatives of greater or less public intervention, which is precisely where the legislative initiatives in process are located. The conclusion is that this approach is correct, that the democratic state should not adopt an interventionist role in the management of social media, but to ensure, through the approval of legal norms adjusted to the technological and social reality, that social networks act in the market as what they are, companies, subject to the rules and in a responsible and transparent way.*

*Keywords: Social networks, freedom of expression, regulation, non-discrimination, state intervention, democratic process, participation, deliberation.*

SUMARIO: I. INTRODUCCIÓN. II. LA MODERACIÓN DE CONTENIDOS COMO ACTIVIDAD CENTRAL DE LAS REDES SOCIALES. III. EL MARCO LEGAL DE LA MODERACIÓN DE CONTENIDOS EN LOS ESTADOS UNIDOS Y EN LA UNIÓN EUROPEA. 1. *Estados Unidos: La sección 230 de la Ley de Decencia en las Comunicaciones y la Primera Enmienda de la Constitución.* 2. *Unión Europea: la Directiva de comercio electrónico, las normas más estrictas sobre determinados contenidos, la autorregulación y las leyes nacionales.* 3. *Las iniciativas de reforma y sus razones.* IV. EL DEBATE SOBRE LA MODERACIÓN DE CONTENIDOS EN LAS REDES SOCIALES. 1. *Una cuestión previa: la neutralidad.* 2. *Las posiciones en el debate sobre moderación de contenidos.* 2.a. El pesimismo radical. 2.b. La intervención pública en las redes sociales. 2.c. La reforma público-privada de la moderación. 2.d. Los incentivos públicos para la reforma de las empresas. 2.e. Las propuestas participativas y comunitarias. 2.f. Los defensores de la autorregulación. V. CONCLUSIÓN. BIBLIOGRAFÍA.

## I. INTRODUCCIÓN

Aunque apenas haya habido algún año sin sobresaltos para Facebook, 2021 destaca entre los más problemáticos de su corta y asombrosa historia. Dos episodios archiconocidos y de naturaleza bien distinta, protagonizados por Donald J. Trump, el primero, y Frances Haugen, el segundo, sirven para sostener esta afirmación. Para empezar, quien aún entonces era presidente de los Estados Unidos de América vio impedido el 6 de enero su acceso a Facebook e Instagram, donde contaba con treinta y cinco y veinticuatro millones de seguidores, respectivamente, porque algunas de sus declaraciones, videos y manifestaciones relativas al asalto del Capitolio y publicadas en sus cuentas habían violado, según la empresa propietaria, las normas de conducta de ambas plataformas. Al día siguiente, Facebook hizo indefinida la restricción de acceso de Trump, y dos semanas más tarde remitió el caso al Consejo de Supervisión, un órgano colegiado de expertos, instituido y financiado por la propia compañía y popularmente conocido como el «Tribunal Supremo» de Facebook, para su revisión. Este Consejo determinó, posteriormente, que si bien la decisión de restringir a Trump el acceso a su cuenta fue correcta, hacerlo de una manera indeterminada y carente de criterios, mediante una suspensión indefinida, no lo fue. De manera que pidió a Facebook que revisase la cuestión en un plazo de seis meses y justificase una respuesta proporcionada y consistente con las normas aplicadas a otros usuarios en la plataforma<sup>1</sup>. Por lo demás, el Consejo solicitó a Facebook explicaciones sobre la medida en que las decisiones relativas al diseño de las plataformas, incluyendo algoritmos, políticas, procedimientos y características técnicas, habían amplificado los *posts* de Trump después de la elección presidencial de 2020 y si había

---

<sup>1</sup> En concreto el Consejo resolvió, en decisión de 5 de mayo de 2021, que (i) no fue apropiado imponer una suspensión «indefinida», (ii) Facebook no siguió un procedimiento publicado previamente y claro, al aplicar la sanción (iii) siempre resulta útil trazar una línea firme entre líderes políticos y otros usuarios influyentes, ya que otros usuarios con amplias audiencias pueden contribuir a causar daños serios, que el contexto importa cuando se trata de determinar probabilidad e inminencia del daño. Propuso, además, una serie de medidas de más claridad, información y recursos cualificados para evaluar riesgos de daños de cuentas influyentes globalmente. También hizo recomendaciones para poner en práctica políticas claras, necesarias y proporcionadas que promuevan la seguridad pública y el respeto a la libertad de expresión (Case Decision 2021-001-FBR).

llevado a cabo algún análisis interno sobre la posible contribución de tales decisiones a los sucesos del 6 de enero. Facebook no contestó a ninguna de estas dos cuestiones. Lo que sí hizo, por el contrario, fue aclarar que la restricción de acceso de Trump durará al menos dos años, y que en enero de 2023 podrá ser admitido de nuevo, si el interesado así lo pide y los expertos que sean consultados concluyen que el riesgo para la seguridad pública que pueda suponer el retorno del expresidente a la red es mínimo o inexistente<sup>2</sup>.

Meses después, Frances Haugen, una ingeniera de datos y antigua gerente de productos en Facebook, saltó de repente a la fama mundial a través de una serie de denuncias realizadas en distintos medios de comunicación y en una comparecencia ante el Congreso de los Estados Unidos<sup>3</sup>. En todas ellas Haugen, que ahora se presenta en su página web personal como defensora de la supervisión pública de las redes sociales<sup>4</sup> afirmó, de forma consistente, que Facebook, pese a sus promesas de conseguir un entorno más sano, prima, en realidad, los contenidos extremos y adictivos para asegurarse de que sus usuarios pasen el mayor tiempo posible conectados a su red, que los más de cuarenta mil empleados contratados para moderar contenidos, así como la utilización masiva del aprendizaje automático (*machine learning*) para detectar contenidos inapropiados, apenas habían servido de algo, y no desde luego para frenar la poderosa campaña antivacunación del Covid-19 en los Estados Unidos, y que los ejecutivos de la empresa habían decidido deshacerse de la unidad de integridad cívica, un equipo dedicado a luchar contra la desinformación y otros ataques contra la democracia, poco después de la elección de Joe Biden como presidente, pasando rápidamente la página del propósito de enmienda. Como es notorio, el efecto más directo e inmediato de las revelaciones de Haugen ha sido el de sumir a Facebook en una nueva y profunda crisis de reputación, a la

---

<sup>2</sup> Decisión de Facebook de 4 de junio de 2021 de suspender a Trump por, al menos, 2 años desde los eventos del 6 de enero de 2021.

<sup>3</sup> Las denuncias, anónimas en principio, salieron a la luz en el diario *The Wall Street Journal* en septiembre de este año («Los archivos Facebook»). El 3 de octubre, Frances Haugen fue entrevistada en el programa *60 Minutes* de la cadena CBS y el 5 de octubre compareció en el Senado. Posteriormente, ha realizado otras comparecencias en distintos foros, parlamentarios y públicos.

<sup>4</sup> [www.franceshaugen.com](http://www.franceshaugen.com)

que Mark Zuckerberg, su fundador y presidente, ha respondido de inmediato con un cambio de nombre para la empresa, que ahora es Meta, y con la enésima reformulación de su propósito<sup>5</sup>.

En cualquier caso, los episodios reseñados han generado multitud de comentarios y reflexiones sobre las consecuencias que tienen en la vida real los contenidos que se publican en el «espacio virtual» de las redes sociales, el poder de las grandes empresas privadas que las controlan (e incluso el de quienes las dirigen, como el propio Zuckerberg, Sundar Pichai o Jack Dorsey)<sup>6</sup> para silenciar a responsables políticos elegidos democráticamente, la falta de transparencia, claridad y formalización del proceso de las decisiones de esas empresas sobre los contenidos que aparecen en sus plataformas –o que son retirados de las mismas–, el trato privilegiado que reciben determinados usuarios en función de su posición y cualidades, el régimen de responsabilidad legal de esas empresas y su impacto sobre el ejercicio de la libertad de expresión o la proporción de sinceridad e hipocresía que puede haber en sus planes recurrentes para mejorar la gobernanza de las redes sociales.

El presente artículo trata de la moderación de contenidos en las redes sociales y, en particular, del debate relativo a las implicaciones jurídicas, políticas, económicas y morales de esa actividad. Se va a considerar aquí, pues, si las decisiones de supresión, retirada, cortes, bloqueo de acceso, priorización o democión de contenidos que Facebook, YouTube o Twitter adopten, en su caso, sobre informaciones, textos, fotos, audios, videos, gráficos o enlaces que los usuarios publiquen en sus redes pueden calificarse, en atención a sus circunstancias específicas, como legales, justas, razonables, proporcionadas, eficaces o no. Para lo que habrá que preguntarse, y eso es precisamente lo que hacen quienes contribuyen al debate, si las redes han de intervenir activamente en el control del contenido que los usuarios suben a las

---

<sup>5</sup> Facebook ha pasado, sucesivamente, por distintas etapas en las que se ha presentado como «red social amistosa», que busca conectar amigos, «máquina de compartir», que necesita para funcionar la mayor cantidad de información posible en la red, instrumento para «conectar a la humanidad» o empoderar «a la gente para construir comunidades», o red social que frente a lo político pone un «mayor énfasis en lo privado», «What is Facebook? Just ask Mark Zuckerberg», Chaim Gartenberg, *The Verge*, 8 de marzo de 2018.

<sup>6</sup> Sundar Pichai es el consejero delegado de Google, Jack Dorsey es el consejero delegado de Twitter.

mismas o han de ser meros conductos neutrales para su publicación, si ofrecen suficiente información y garantías sobre la forma en la que moderan, o si el régimen de responsabilidad legal de las empresas es el adecuado, entre otras cosas. La moderación de contenidos por las redes es, como puede apreciarse a primera vista, un tema amplio y complejo. Primero, porque afecta y remite a valores, como la libertad, la igualdad o el pluralismo político, a derechos de libre expresión e información, privacidad, seguridad, o no discriminación, y a bienes públicos, como la participación, la deliberación o la inclusión social. Segundo, porque se relaciona con aspectos negativos, cuando no patológicos, de Internet, como la proliferación de contenidos banales, estúpidos, ofensivos, dañinos o directamente ilegales, y se puede emparentar con actividades tan antipáticas como la censura, la vigilancia y el control sobre los individuos. Tercero, en la medida en que resulta inevitablemente heterogéneo, habida cuenta de los diferentes sistemas jurídicos, tipos de Estados, culturas, valores, e incluso modelos de Internet que cabe encontrar en el mundo, lo que vale decir que aquello que en un sitio se reputa dañino o contrario a la ley, pueda no serlo en otro sitio<sup>7</sup>. Cuarto, al expresar de manera aguda y poliédrica la tensión entre lo privado y lo público, o entre la empresa y el Estado, ya que las grandes tecnológicas (*Big Tech*), como Facebook, Google (YouTube) o Twitter, por citar tan solo a las más puramente digitales, superan en poder y riqueza a muchos Estados<sup>8</sup>, suplantando o confundiendo, en ocasiones, funciones propias de los Estados, toman decisiones que afectan a derechos individuales, sin que sus titulares puedan, muchas veces, recurrir a la administración o a los tribunales del Estado, y a la vez se inspiran en normas internacionales de derechos humanos y en normas sustantivas y procesales de ordenamientos jurídicos democráticos para incluirlas en sus condiciones del servicio y políticas comunitarias. Y, en fin, porque manifiesta una segunda tensión, interna en este caso, entre el imperativo implacable de un modelo de negocio basado en ofrecer servicios gratuitos a cambio recoger masivamente

<sup>7</sup> Sobre diversidad de sistemas jurídicos y políticos y moderación de contenidos, Daskal (2019), y sobre modelos de Internet: Shira Ovide, «The Internet is Splintering», *The New York Times*, 17 de febrero de 2021.

<sup>8</sup> Rebecca MacKinnon y Lori Andrews creen que hay que pensar en Facebook o Google como «naciones-estado» o «soberanos del ciberespacio».

datos personales de los usuarios, para venderlos luego a anunciantes y permitir, así, una publicidad potenciada al extremo, y la retórica de rectos principios, prácticas honestas y elevadas aspiraciones comunitarias que, según la propaganda y el *lobby* de las empresas, guían sus operaciones<sup>9</sup>.

Por ello, para no desviarme en las ramificaciones de estas cuestiones y ceñirme al objeto principal del artículo, voy a analizar brevemente, antes de entrar en el mismo, qué hacen las redes sociales cuando moderan contenidos, y qué leyes, en Estados Unidos y en la Unión Europea<sup>10</sup>, les permiten realizar tal actividad.

## II. LA MODERACIÓN DE CONTENIDOS COMO ACTIVIDAD CENTRAL DE LAS REDES SOCIALES

Las redes sociales, que de manera convencional pueden definirse como plataformas *online* que facilitan la interacción entre sus usuarios para comunicarse y compartir contenidos que ellos mismos generan<sup>11</sup>, han sido causa decisiva, junto al despliegue de las infraestructuras de banda ancha y la popularización de los *smartphones*, del crecimiento exponencial de Internet en la primera década de este siglo. Tienen, como es sabido, una enorme popularidad, ya que de los 4.800 millones de personas que acceden a Internet en el mundo, 4.500 son usuarios de redes sociales<sup>12</sup>, acaparan la mayor parte del tiempo que se pasa en Internet, que supone algo más de 200.000 millones de horas contando solo a Facebook, Youtube y Twitter<sup>13</sup>, y se usan para múltiples propósitos, ya sean familiares, personales, comerciales, informativos, recreativos o políticos. No es de extrañar, pues, que las más exitosas de entre ellas, como Facebook o YouTube (Google), formen parte del selecto grupo de *Big Tech*, las empresas que han

---

<sup>9</sup> Sobre el modelo de negocio de las grandes empresas tecnológicas, ver entre otros Zuboff (2019), Cohen (2019) y Galloway (2017).

<sup>10</sup> No voy a considerar, pues, el caso de países de enorme importancia geopolítica, que tienen redes sociales nacionales enormemente populares, como China o Rusia, y un sistema político y jurídico, así como una regulación de Internet, bien distintos a los occidentales.

<sup>11</sup> UK Competition Markets Authority (2019), «Online platforms and digital advertising market study», Glossary.

<sup>12</sup> DataReportal (2021).

<sup>13</sup> Zyro, 2021.

alcanzado las mayores capitalizaciones en bolsa de la historia (superiores al millón de millones de dólares en varios casos) y disponen de recursos y capacidad de influencia casi sin límites<sup>14</sup>.

Que las redes sociales se nutran, principalmente, de contenido generado por sus usuarios en cantidades masivas les ha obligado a aplicar, de forma concomitante, sus propias condiciones de uso y normas comunitarias, así como las leyes aplicables de los países en donde operan, porque el contenido inapropiado puede ser una fuente importante de responsabilidad y afectar, además, a su imagen y reputación (Roberts, 2017). Así, Facebook pide a sus usuarios el cumplimiento de una serie de compromisos, como el de no usar sus productos para realizar acciones o compartir contenidos que «infrinjan las Condiciones del Servicio, las Normas comunitarias u otras condiciones y políticas que rijan el uso de Facebook», o que sean «ilegales, engañosos, discriminatorios o fraudulentos» o que «infrinjan los derechos de otras personas, como los relativos a la propiedad intelectual», de manera que si esto no ocurre, Facebook «puede eliminar o bloquear el contenido que incumpla lo dispuesto por las Condiciones del Servicio», informando de ello al usuario y explicándole, en principio, las opciones que tiene de revisión<sup>15</sup>. Twitter, por su parte, se reserva el derecho de retirar contenido que incumpla el Contrato con el Usuario, incluyendo, por ejemplo, violaciones de propiedad intelectual o derecho de marca, falsedades, conductas ilegales o acoso, y «pone a disposición de los usuarios información sobre políticas específicas y procesos para reportar o para apelar violaciones del Contrato»<sup>16</sup>. En cuanto a YouTube, advierte que puede «usar sistemas automatizados para analizar el contenido, con el fin de ayudar a detectar infracciones o usos inadecuados como el spam,

---

<sup>14</sup> No todas las grandes empresas tecnológicas son redes sociales. Por ejemplo, Amazon o Apple no lo son, aunque Microsoft es propietaria de una red social profesional, LinkedIn. Las redes sociales son plataformas *online*, como también lo son los buscadores (Google Search), las empresas de comercio electrónico (Amazon) o las tiendas de aplicaciones (Apple Store). Pero no todas las plataformas *online* son redes sociales: es el caso de Uber o Airbnb, por ejemplo. En este artículo habrá referencias a las redes sociales como plataformas, que lo son, cuando ese sea el término utilizado por la norma, disposición o autor de que se trate.

<sup>15</sup> Condiciones del Servicio de Facebook (2021).

<sup>16</sup> Contrato con el Usuario de Twitter (2021).

el software malicioso o el contenido ilegal» y que si tiene motivos razonables para creer que el contenido incumple el contrato o puede generar daños, puede, también, retirarlo o inhabilitarlo, ya sea en parte o en su totalidad<sup>17</sup>.

Desde esta perspectiva, la moderación de contenidos aparece, según Sarah T. Roberts, como «una práctica esencial en el ciclo de producción de las redes sociales», una actividad crucial que protege la marca (reforzando la adhesión del usuario a las normas y directrices de la red), asegura el cumplimiento de leyes y reglamentos que resulten de aplicación y contribuye positivamente a mantener una audiencia en la red (Roberts, 2019: 33-34). Y abundando en esta idea, Tarleton Gillespie considera que la moderación es algo central, y de ninguna manera periférico, en la actividad de las redes sociales, hasta el punto de identificarla como el producto (*commodity*) que, en realidad, ofrecen (2018). No obstante, las redes sociales no hacen ninguna ostentación de su actividad moderadora y prefieren evitar, por lo general, la publicidad y la discusión sobre la misma, tal vez porque no casa bien con el paradigma fundacional de Internet como lugar libre de cualquier tipo de censura<sup>18</sup>, o con la relación personalizada que promueven las redes sociales, o porque a estas les resulta incómodo reconocer, como apunta la propia Sarah T. Roberts, que las decisiones de moderación las toman, básicamente, «ejércitos de trabajadores desconocidos [...], ocultos a la visión y percepción de los usuarios, mal pagados, poco considerados y frecuentemente externalizados»<sup>19</sup>.

La revisión del contenido en que consiste la moderación puede tener lugar antes o después de su publicación, e iniciarse en virtud de

---

<sup>17</sup> Términos y Condiciones del Servicio de YouTube (2021).

<sup>18</sup> Son famosas las posiciones de pioneros del libertarismo característico de los primeros años de Internet, como John Perry Barlow, autor de la Declaración de Independencia del Ciberespacio (1996), con su rechazo frontal a las formas tradicionales de control estatal sobre la red, o John Gilmore, para quien la arquitectura de Internet resulta estructuralmente inmune a cualquier tipo de censura de la información que viaje a través de sus nodos interconectados.

<sup>19</sup> Roberts, (2019: 71 y ss.). En este sentido, Shoshana Zuboff considera la moderación de contenidos como una función de paria del capitalismo de vigilancia y se refiere a las «banalidades diarias de los flujos de trabajo racionalizado, donde hay que decidir sobre los horrores y el odio del mundo, a un ritmo y frente a tal volumen, que solamente ofrece un breve momento [a quien modera] para elevar el pulgar hacia arriba o hacia abajo» (2019: 508-509).

una decisión de un moderador, o de un administrador de la red, o de una decisión automatizada, o de la queja de un tercero, o la denuncia de un usuario que señala (*flagging*) algo que estima problemático. Y el resultado de esa revisión, en caso de que el contenido no se considere apropiado, puede consistir, entre otras posibles opciones, en excluirlo, en aplicación de sus políticas de retirada de contenidos, reducir su visibilidad, mediante la aplicación de sus sistemas de clasificación (*ranking*), etiquetarlo de una forma determinada, adjuntar al mismo un aviso sobre sus cualidades (*warning*), impedir o deshabilitar comentarios o anuncios sobre el mismo, eliminarlo en ciertas regiones geográficas o jurisdicciones, o hacerlo solo accesible a aquellos usuarios que proactivamente hayan manifestado previamente querer verlo (*opt-in*) (Keller, 2019).

Por último, conviene destacar entre las notas distintivas de las condiciones de servicio (o contrato con los usuarios o términos de uso) de las redes sociales que suelen ser, en cuanto a la evaluación del contenido que se considera admisible, más restrictivas que las legislaciones estatales (por ejemplo, en el tratamiento de los cuerpos humanos desnudos), han ido ganando en pragmatismo y concreción (Klonick, 2018) y, por último, presentan bastantes notas comunes con las propias de otros servicios, como las telecomunicaciones o la electricidad. Siendo cierto que, como señala Nicolas Suzor, se trata de documentos relevantes que amparan decisiones con efectos reales sobre la vida de las personas, confieren un amplio poder a las empresas y están redactados para salvaguardar los propios intereses empresariales, nada de ello resulta exclusivo de las redes sociales, sino más bien lo contrario (basta pensar, por ejemplo, en un corte de suministro eléctrico o de energía) (2019).

### III. EL MARCO LEGAL DE LA MODERACIÓN DE CONTENIDOS EN LOS ESTADOS UNIDOS Y EN LA UNIÓN EUROPEA

El marco legal que aún gobierna la moderación en Internet proviene de una época en la que aún no se había producido el desarrollo incremental de las tecnologías que soportan la *world wide web*, y que permiten al usuario contribuir directamente a la producción de la información, en vez de ser receptor pasivo de la misma (web 2.0), y cuando ni siquiera Facebook, YouTube o Twitter existían: la década

de los noventa del siglo pasado. En esos años previos a la explosión comercial de Internet, el objetivo del Congreso norteamericano era «preservar el mercado libre, vibrante y competitivo, que existe en la actualidad en Internet, libre de las restricciones de la regulación federal o de los estados» y «promover su desarrollo continuo»<sup>20</sup>, mientras que la Comisión Europea no pretendía otra cosa que

alentar el desarrollo del comercio electrónico en la Unión [...], dejando claro que las empresas no tenían la obligación de asegurarse sobre la legalidad de todo el contenido que alojaban, conseguir un equilibrio entre los derechos fundamentales de las partes interesadas (así, privacidad y libertad de expresión, libertad de empresa y derecho de propiedad), y fortalecer el Mercado Único Digital mediante la adopción de un criterio común sobre las exenciones de responsabilidad<sup>21</sup>.

Singularmente elocuente, en este punto, resulta la afirmación del Tribunal Supremo estadounidense en *Reno v. American Civil Liberties Union* (1997), de que «el vasto foro democrático de Internet» no había estado tradicionalmente sujeto al tipo de regulación y supervisión gubernamental de la industria de la radiodifusión» y que Internet no era «tan invasivo como la radio o la televisión». Eran otros tiempos, qué duda cabe. En perspectiva, que ese marco siga aún vigente expresa bien a las claras como las políticas públicas y el derecho suelen ir, en cuanto a su creación y aplicación, a la zaga de las tecnologías, y ofrece razones tanto a quienes apoyan sus normas, para justificar su adecuación a la realidad, y su contribución a la innovación y al crecimiento de la industria, como a quienes las critican, para probar que su mantenimiento es fruto de la inercia, la incapacidad del derecho para responder con prontitud a los retos de la tecnología y la influencia del mercado sobre el Estado.

---

<sup>20</sup> Preámbulo de la Ley de Telecomunicaciones (Telecommunications Act) de 1996.

<sup>21</sup> Comisión Europea, «Memorando explicativo del proyecto de la Comisión de una directiva sobre ciertos aspectos legales del comercio electrónico en el mercado único» (1998).

1. *Estados Unidos: La sección 230 de la Ley de Decencia en las Comunicaciones y la Primera Enmienda de la Constitución*

La capacidad legal de las redes sociales para moderar contenidos en Estados Unidos procede de la sección 230 de la Ley de Decencia en las Comunicaciones (Communications Decency Act, CDA)<sup>22</sup>, que otorga a las empresas una amplia inmunidad sobre la responsabilidad en que podrían incurrir por el contenido generado por sus usuarios y publicado en su web o en sus aplicaciones. Según esta norma «ningún proveedor o usuario de un servicio informático interactivo será tratado como editor o creador de cualquier información proporcionada por otro proveedor de contenidos de información». Y añade, además, que

ningún proveedor o usuario de un servicio informático interactivo será considerado responsable de cualquier acción de buena fe para restringir el acceso o la disponibilidad de material que el proveedor o el usuario considere que es obsceno, lujurioso, lascivo, indecente, excesivamente violento, acosador u objetable, con independencia de que el material esté protegido constitucionalmente o no.

El propósito de esta concesión de inmunidad, que de todas formas no es completa<sup>23</sup>, fue evitar problemas de censura colateral que pudiesen afectar a la libertad de expresión, y que surgirían con facilidad si los prestadores de servicios de Internet (que incluyen hoy a redes sociales y todo tipo de plataformas *online*) fuesen considerados editores o creadores y, en consecuencia, responsables del contenido creado por los usuarios, y al tiempo incentivar a los prestadores de servicios a que jugasen el papel de «buen samaritano», contribuyendo de manera activa a la retirada de aquellos contenidos que estimasen

---

<sup>22</sup> Formalmente codificada como 47 U.S.C. § 230, fue incorporada a la Ley de Decencia de las Telecomunicaciones (Communications Decency Act, CDA), que fue el nombre que se dio al Título V de la Ley de Telecomunicaciones (Telecommunications Act) de 1996. La Ley de Decencia de las Comunicaciones fue declarada inconstitucional por el Tribunal Supremo en *Reno v. American Civil Liberties Union* (1997), si bien dejó indemne a la sección 230, cuya conformidad con la Constitución ha sido posteriormente validada.

<sup>23</sup> Así, los prestadores de servicios de Internet pueden ser denunciados y procesados por infracción de normas penales, como sucedió con Backpage.com, por facilitar ilícitamente la prostitución.

inapropiados u ofensivos. Y, por supuesto, resolver la situación creada por una sentencia del Tribunal Superior de Nueva York que había considerado a Prodigy, un prestador de servicios de Internet que alojaba y moderaba un tablón de anuncios digital, «editor» de las difamaciones publicadas por un tercero y, por ende, responsable en un proceso civil por libelo<sup>24</sup>.

Resulta difícil exagerar la importancia de la sección 230, «las veintiséis palabras que crearon Internet» según el título de la monografía de Jeff Kosseff sobre la historia de esta disposición, que ha sido el catalizador del crecimiento de un sector (el de las redes sociales) basado en el contenido generado por terceros (2019), y que ha permitido prosperar a empresas como Google o Facebook que, según Jack Balkin, «podrían no haber sobrevivido a la serie de pleitos por difamación que hubieran tenido que afrontar de haber sido tratadas como editoras de los innumerables *links*, *blogs*, *posts*, comentarios y actualizaciones que aparecen en sus plataformas» (2014).

Pero las redes sociales no están solamente protegidas, en cuanto a la moderación de contenidos, por la sección 230. También lo están por la Primera Enmienda de la Constitución, según la cual «el Congreso no promulgará ninguna ley que restrinja la libertad de expresión» y, en concreto, por la interpretación del Tribunal Supremo sobre lo que se puede considerar «expresión» a los efectos de la protección de la Primera Enmienda y sobre extensión de la protección que ofrece a los distintos tipos de contenidos. Así, Volokh y Falk dicen que las plataformas, como las redes sociales, expresan un juicio editorial con las decisiones que toman respecto a la retirada y la priorización de contenidos (2012) de manera que ni el Gobierno puede impedirles esa libertad ni sus usuarios pueden demandarles con éxito, ya que la Primera Enmienda se aplica solamente a los actores estatales<sup>25</sup>. Las

---

<sup>24</sup> *Stratton-Oakmont, Inc. v. Prodigy Services Co.* (1995) N.Y. Sup. Ct. Por el contrario, en *Cubby Inc. v. CompuServe Inc.*, (1991), el Tribunal del Distrito del Sur de Nueva York concluyó que CompuServe no era responsable del contenido difamatorio publicado por alguno de sus usuarios, ya que la empresa permitía la circulación de todo tipo de contenidos, sin ningún tipo de moderación ni implicación editorial, lo que le hacía más un «distribuidor» que un «editor». Sobre todo esto, ver Pinkus (2021).

<sup>25</sup> El texto de la Primera Enmienda únicamente impide al Congreso aprobar leyes que restrinjan la libertad de expresión. Esta protección fue posteriormente extendida a los Estados y a los Gobiernos locales por la Doctrina de la Acción del Estado (*State Action*

redes sociales pueden, pues, establecer sus propias políticas sobre contenidos en sus condiciones del servicio y, de hecho, han recurrido a las mismas de forma consistente cuando se han opuesto, con éxito, a demandas de usuarios por violar sus derechos de libertad de expresión, proceso debido, igual protección, discriminación, incumplimiento de contrato, competencia desleal o interferencia en beneficio económico prospectivo, entre otros motivos<sup>26</sup>.

2. *Unión Europea: la Directiva de comercio electrónico, las normas más estrictas sobre determinados contenidos, la autorregulación y las leyes nacionales*

El marco legal de la moderación en la Unión Europea es más complejo y variado que el norteamericano, ya que atiende a distintos tipos de contenidos, clases de plataformas y obligaciones impuestas a los Estados y a las empresas<sup>27</sup>, aunque su régimen básico, contenido en la Directiva de comercio electrónico (2000)<sup>28</sup>, presenta una sintonía fundamental con el de la sección 230 de la CDA, sobre todo al establecer en su artículo 14 que «los Estados miembros garantizaran que [...] el prestador del servicio no pueda ser considerado responsable de los datos almacenados a petición del destinatario» a condición de que «el prestador del servicio no tenga conocimiento efectivo de que la actividad o la información es ilícita [...]» o de que «en cuanto tenga

---

Doctrine) y la Cláusula del Proceso Debido (Due Process Clause) de la Decimocuarta Enmienda. Sin embargo, bajo la Doctrina de la Acción del Estado, las restricciones de la Primera Enmienda no se extienden por lo general a las personas físicas y jurídicas privadas (como pueda ser una empresa de red social). En otras palabras, una persona particular o una empresa privada no puede violar los derechos de libertad de expresión de un individuo, solo puede hacerlo un actor estatal. A menos que la parte privada que pretende restringir la expresión incurra en una de estas tres excepciones de la Doctrina de la Acción del Estado: (i) la acción de restricción de la expresión por parte de un privado implica una función tradicional y exclusivamente reservada al Estado, Doctrina de la Función Pública Exclusiva (Exclusive Public Function Doctrine), (ii) la acción es tal que el Estado ha estado significativamente implicado o «enredado» con el particular en la acción, y (iii) la acción del privado puede ser tratada como si fuera del Gobierno mismo, por superposición o solapamiento de identidades. Sobre esto, ver Pinkus, (2021).

<sup>26</sup> En relación a esto, ver Keller (2019).

<sup>27</sup> Sobre este punto, ver De Streel *et al.* (2020).

<sup>28</sup> Directiva 2000/31/CE del Parlamento Europeo y del Consejo, de 8 de junio de 2000, relativa a determinados aspectos jurídicos de los servicios de la sociedad de la información, en particular el comercio electrónico en el mercado interior.

conocimiento [...] actúe con prontitud para retirar los datos o hacer que el acceso a ellos sea imposible» y todo ello sin perjuicio de que «un tribunal o una autoridad administrativa, [...] exijan al prestador de servicios poner fin a una infracción o impedirla» o de que «los Estados miembros establezcan procedimientos por los que se rija la retirada de datos o se impida el acceso a los mismos». Por lo demás, esta Directiva, que ha sido objeto en los últimos veinte años de diversos intentos de reforma que nunca llegaron a buen puerto, prohíbe la imposición a las empresas de obligaciones de supervisión, búsqueda activa o monitorización de contenidos, para así proteger derechos fundamentales de los usuarios<sup>29</sup>, promueve la autorregulación, la corregulación y los mecanismos alternativos de resolución de disputas<sup>30</sup>, y afirma el principio de regulación por las normas del país de origen.

Además de la Directiva de comercio electrónico, cuyo régimen básico se complementa con normas sobre servicios de comunicación audiovisual y medidas para combatir contenidos ilícitos<sup>31</sup>, la Unión Europea cuenta con normas más estrictas para cuatro tipos de contenidos, a saber, terrorismo<sup>32</sup>, abuso y explotación sexual de la

---

<sup>29</sup> El artículo 15 de la Directiva establece que «los Estados miembros no impondrán a los prestadores de servicios una obligación general de supervisar los datos que transmitan o almacenen, ni una obligación general de realizar búsquedas activas de hechos o circunstancias que indiquen actividades ilícitas [...], sin perjuicio de que puedan establecer obligaciones tendentes a que los prestadores de servicios [...] comuniquen con prontitud a las autoridades públicas competentes los presuntos datos ilícitos o las actividades ilícitas».

<sup>30</sup> Ayres y Braithwaite ven la diferencia entre corregulación y autorregulación en que en la primera existe algún tipo de supervisión y/o de ratificación por parte del Gobierno, mientras que Julia Black considera que la autorregulación es una situación en la que un determinado colectivo o grupo impone regulación sobre sus miembros. Para estos conceptos ver Baldwin, Cave y Lodge (2012).

<sup>31</sup> El régimen básico de la Directiva de comercio electrónico se complementa con la Revisión de la Directiva de servicios de comunicación audiovisual (2018), que impone a una categoría concreta de empresas, las plataformas de compartición de videos, la obligación de tomar medidas apropiadas y proporcionadas para proteger al público, en general, de contenido ilegal y a los menores, en particular, de contenido dañino (*harmful*) y la Recomendación de la Comisión Europea (2018) sobre medidas para combatir eficazmente los contenidos ilícitos en línea, que se refiere tanto a terrorismo como a abuso sexual de la infancia.

<sup>32</sup> Directiva (UE) 2017/541 del Parlamento Europeo y del Consejo, de 15 de marzo de 2017, relativa a la lucha contra el terrorismo y por la que se sustituye la Decisión marco 2002/475/JAI del Consejo y se modifica la Decisión 2005/671/JAI del Consejo y Reglamento (UE) 2021/784 del Parlamento Europeo y del Consejo, de 29 de abril de 2021, sobre la lucha contra la difusión de contenidos terroristas en línea.

infancia<sup>33</sup>, racismo, xenofobia y delitos de odio<sup>34</sup> y propiedad intelectual<sup>35</sup> que, en los dos primeros casos, establecen la obligación para los Estados miembros de adoptar medidas para remover y bloquear sitios web que contengan o diseminen ese tipo de contenidos, en el tercero, obligan a los Estados a asegurar que el discurso del odio, racista y xenófobo sea punible, sin imponer obligaciones específicas a las plataformas, mientras que en el cuarto requieren que las empresas concluyan acuerdos con los titulares de derechos para la explotación de los contenidos protegidos, de manera que si no lo hacen se consideraran, en principio, responsables de violar derechos de propiedad intelectual. Estas normas de derecho comunitario «fuerte» se complementan con iniciativas de autorregulación, a menudo inspiradas o sugeridas por la Comisión Europea, como en el caso de la desinformación, que no siempre es ilegal, pero puede ser muy dañina para la sociedad y la democracia<sup>36</sup>, o el discurso de odio ilegal *online*<sup>37</sup>. En fin, aparte de las normas comunitarias, los distintos Derechos nacionales imponen, para algunos tipos de empresas (redes sociales u otras) obligaciones adicionales respecto a la moderación de contenidos ilegales o dañinos, y así Alemania ha aprobado la Ley para la mejora de la aplicación del derecho en las redes sociales (2017)<sup>38</sup>, etiquetada por algunos como «Ley Facebook», y Francia la Ley contra la manipulación de la información (2018)<sup>39</sup> y la Ley dirigida a luchar contra los contenidos de odio en Internet (2020)<sup>40</sup>.

---

<sup>33</sup> Directiva (UE) 2011/92 del Parlamento Europeo y del Consejo, de 13 de diciembre de 2011, relativa a la lucha contra los abusos sexuales y la explotación sexual de los menores y la pornografía infantil y por la que se sustituye la Decisión marco 2004/68/JAI del Consejo.

<sup>34</sup> Decisión marco 2008/913/JAI del Consejo, de 28 de noviembre de 2008, relativa a la lucha contra determinadas formas y manifestaciones de racismo y xenofobia mediante el derecho penal.

<sup>35</sup> Directiva (UE) 2019/790 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital y por la que se modifican las Directivas 96/9/CE y 2001/29/CE.

<sup>36</sup> Código de buenas prácticas contra la desinformación (2018).

<sup>37</sup> Código de conducta para luchar contra el discurso de odio ilegal online (2016).

<sup>38</sup> NetzDG, en su denominación abreviada, o Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken.

<sup>39</sup> Loi n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information.

<sup>40</sup> Loi n° 2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur Internet.

Mucho se ha escrito sobre la comparación, en cuanto al régimen de la libertad de expresión, entre la Unión Europea y los Estados Unidos, y sobre la opinión generalizada de que la protección de este derecho es más limitada en Europa, algo que se puede confirmar si uno lee detenidamente el artículo 10 del Convenio Europeo de Derechos Humanos, o incluso el artículo 11 de la Carta de los Derechos Fundamentales de la Unión Europea<sup>41</sup>. Que ello sea así depende, como es lógico, de factores históricos, políticos y culturales que hacen que, como apunta Oreste Pollicino, la Primera Enmienda se centre principalmente en la dimensión activa de poder expresar libremente ideas y opiniones, mientras que las normas europeas sobre derechos humanos antes citadas enfatizan la dimensión pasiva de recibir una información pluralista (2020). En cualquier caso, también hay un acuerdo extendido entre estudiosos y comentaristas de este asunto sobre la menor protección que las normas europeas ofrecen a las empresas de redes sociales, en cuanto a moderación de contenidos, comparada con la de la Primera Enmienda.

En fin, sobre la cuestión de si un usuario puede obligar a una red social a alojar contenido que el primero haya generado contra la voluntad de la red, Daphne Keller se refiere a algunas decisiones judiciales en Alemania, específicamente de los Tribunales Regionales de Munich, Colonia y Siegen (2018), donde se da la razón, frente a Facebook, a una serie de usuarios que pedían a la red social la restitución de sus *posts*, retirados por Facebook en aplicación de sus Normas comunitarias. En esas decisiones, la autora citada ve una

---

<sup>41</sup> El primero establece que «toda persona tiene derecho a la libertad de expresión» que incluye «la libertad de opinión y la libertad de recibir o de comunicar informaciones o ideas sin que pueda haber injerencia de autoridades públicas y sin consideración de fronteras», si bien el ejercicio de estas libertades, que entrañan deberes y responsabilidades, «podrá ser sometido a ciertas formalidades, condiciones, restricciones o sanciones, previstas por la ley, que constituyan medidas necesarias, en una sociedad democrática, para la seguridad nacional, la integridad territorial o la seguridad pública, la defensa del orden y la prevención del delito, la protección de la salud o de la moral, la protección de la reputación o de los derechos ajenos, para impedir la divulgación de informaciones confidenciales o para garantizar la autoridad y la imparcialidad del poder judicial», mientras que el segundo declara que «toda persona tiene derecho a la libertad de expresión», que «comprende la libertad de opinión y la libertad de recibir o comunicar informaciones o ideas sin que pueda haber injerencia de autoridades públicas y sin consideración de fronteras» y añade que «se respetan la libertad de los medios de comunicación y su pluralismo».

aplicación de la doctrina de la *Drittwirkung*, que utiliza a los derechos humanos como guía a la hora de interpretar las obligaciones entre privados (2019:12).

### 3. *Las iniciativas de reforma y sus razones*

En estos momentos hay iniciativas legislativas en curso, tanto en Norteamérica como en Europa, para modificar el régimen de responsabilidad, derechos y obligaciones en relación a los contenidos que se suben y circulan por Internet, que lógicamente afectan a la moderación de los mismos por las redes sociales. Todas ellas buscan cambiar el marco actual, el *statu quo*, constituido por la sección 230 de la CDA y la Directiva de comercio electrónico, que se entiende insuficiente y desbordado. Todas, además, han venido precedidas o acompañadas, según los casos, de otros proyectos normativos, medidas de regulación, comparecencias parlamentarias, investigaciones e informes y demandas judiciales, dirigidas a limitar y controlar el poder de las grandes empresas tecnológicas, no solo en cuanto a los contenidos, sino también a otras áreas como la privacidad, la defensa de la competencia, la fiscalidad o la propiedad intelectual. El impulso de reforma del régimen de los contenidos en Internet se sustenta, en definitiva, en una creciente toma de conciencia política y social sobre los problemas de todo tipo que plantean las redes sociales, esa «ventana con vistas al Hades» de la que habla Janan Ganesh<sup>42</sup>, en un mejor conocimiento de la manera en que las redes disponen, filtran y

---

<sup>42</sup> Janan Ganesh, «Don't judge the US by its politics», *Financial Times*, 1 de junio de 2021. Sobre este punto, el Tribunal Constitucional español, en su sentencia 93/2021, de 10 de mayo de 2021 (recurso de amparo 3223-2019), cuyo objeto era decidir si el contenido de un mensaje que la demandante de amparo publicó en su cuenta de Facebook, al día siguiente de la muerte del torero Víctor Barrio en la plaza de toros de Teruel, en el que le llamaba asesino y decía que no podía sentir su muerte, estaba amparado por el ejercicio de su derecho fundamental a la libertad de expresión (art. 20.1 CE), dijo que «[...] la generalización del uso de las redes sociales [...] ha supuesto una transformación sin parangón del modelo tradicional de comunicación, dando lugar a un modelo comunicativo que, entre otras notas, se caracteriza por la fragilidad de los factores moderadores del contenido de las opiniones. Dicha transformación ha supuesto un drástico cambio en el perfil del emisor y también de los receptores, cuya facilidad para interactuar entre sí y con el emisor les distancia del carácter pasivo del modelo tradicional. Además estos, en muchas ocasiones, actúan con precaria conciencia de la proyección de las opiniones emitidas [...]», FJ 2.

presentan los contenidos y de sus motivaciones económicas<sup>43</sup> y, por último, en el convencimiento generalizado de que la «espiral de privatización de la regulación» (Wagner, 2018: 219 y ss.) por la que los Gobiernos amenazan a las empresas con medidas draconianas, cuya adopción en realidad desean evitar (por costosas, impopulares o casi imposibles de aplicar), para forzar a aquellas a cambiar sus procesos de moderación y autorregularse no ha terminado de funcionar o, como mínimo, no lo ha hecho de forma medianamente satisfactoria. Por lo demás, y a diferencia de lo que ocurrió en la infancia del Internet comercial, cuando fue Estados Unidos quien sentó las bases del régimen legal de los contenidos en la red, ahora es Europa quien lidera la iniciativa de su reforma, como una muestra más del denominado «efecto Bruselas», que refleja, según Anu Bradford, que «el poder de la normativa de la Unión Europea para influir en los mercados globales ha crecido significativamente en las dos últimas décadas, particularmente en el área de la economía digital» (2020: 131-132).

Yendo, brevemente, al contenido concreto de las reformas propuestas, que comparten una orientación básica de exigencia de mayor responsabilidad a las redes sociales y control de sus actividades de moderación, hay que empezar por el proyecto de Ley de Servicios Digitales<sup>44</sup> presentado por la Comisión Europea, que no define lo que es «ilegal» en la red, ni afecta de modo directo a las disposiciones nacionales o comunitarias sobre la cuestión pero que establece, para las empresas (o plataformas, en sus propios términos) a las que va dirigido, obligaciones sobre la retirada del contenido *online*, medidas para luchar contra el contenido ilícito, mecanismos de denuncia y cooperación, para que usuarios y «alertadores fiables» (*trusted flag-*

---

<sup>43</sup> El modelo económico de las grandes plataformas online, entre ellas las redes sociales, ha dado lugar a un «capitalismo de vigilancia» que Shoshana Zuboff define como «un nuevo orden económico que considera la experiencia humana como materia prima para la realización de prácticas comerciales opacas de extracción, predicción y ventas» o a un «capitalismo de la información» (*informational capitalism*), que para Julie Cohen constituye «una alineación del capitalismo como modo de producción y el “informacionalismo” como modo de desarrollo, orientado hacia la acumulación del conocimiento y a mayores niveles de complejidad en el procesamiento de la información, todo ello para extraer y apropiarse de excedentes de los usuarios».

<sup>44</sup> Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a un mercado único de servicios digitales (Ley de Servicios Digitales) y por el que se modifica la Directiva 2000/31/CE, 15 de diciembre de 2020.

gers) notifiquen la existencia de tal contenido, salvaguardias para los usuarios cuyo contenido haya sido erróneamente eliminado, la posibilidad de impugnar decisiones de las plataformas sobre moderación de contenidos, la armonización de las exenciones de responsabilidad legal para las plataformas, medidas de transparencia de amplio alcance, que incluyen los algoritmos utilizados para las recomendaciones, poderes para investigar cómo trabajan las plataformas, entre ellos el acceso a datos clave y, en fin, la consideración como «sistémicas» para las plataformas usadas por más de un 10% de la población europea (45 millones de usuarios), lo que las sujetará, no solo a obligaciones específicas, sino también a una nueva estructura de supervisión, con capacidad de imponerles sanciones de manera directa<sup>45</sup>.

Por lo que respecta a Estados Unidos, el proyecto de Ley de Tecnología Segura<sup>46</sup> busca la reforma de la sección 230 de la CDA para alcanzar tres objetivos: primero, limitar el objeto de su inmunidad, que cubrirá cualquier tipo de «discurso» pero no de «información», y excluir a los anuncios y contenidos de pago; segundo, hacer que sea el demandado quien, para poder beneficiarse de la inmunidad, soporte la carga de la prueba de que es, en realidad, el prestador de un servicio informático interactivo; y, tercero, restringir el alcance de la cláusula del «buen samaritano», para permitir que el demandante pueda solicitar medidas cautelares cuando el prestador del servicio no actúe en caso de contenido que pueda causar, con probabilidad, un daño irreparable.

Finalmente, en el Reino Unido, el proyecto de Ley de Seguridad *Online*<sup>47</sup> impone a redes sociales y plataformas que permitan conectar usuarios una obligación general de cuidado (*duty of care*), que incluye, entre otras, obligaciones específicas de evaluación de

---

<sup>45</sup> La nueva estructura de supervisión propuesta por el proyecto de Ley de Servicios Digitales incluye la creación de un consejo de Coordinadores Nacionales de Servicios Digitales y la atribución de poderes especiales para la Comisión en la supervisión de las plataformas de naturaleza sistémica.

<sup>46</sup> Safe Tech Act (Safeguarding Against Fraud, Exploitation, Threats, Extremism and Consumer Harms Act), propuesta el 8 de febrero de 2021 por el senador demócrata Mark Warner y copatrocinada por las senadoras Mazie Hirono y Amy Klobuchar, también demócratas.

<sup>47</sup> Draft Online Safety Bill, presentado al Parlamento por el Ministro de Estado de Cultura y Digital en mayo de 2021.

riesgos del contenido ilegal, de impactos sobre derechos de libertad de expresión y privacidad, de reporte y notificación y de ofrecimiento de acciones y soluciones, así como obligaciones para proteger contenido que revista «importancia democrática»; además, el proyecto establece cuantiosas multas (de hasta 18 millones de libras o equivalentes al 10% de los ingresos totales anuales) para las empresas que no retiren contenido ilegal o dañino (como propaganda terrorista o abuso sexual infantil), y confiere competencias de supervisión y aplicación de la ley a *Ofcom*, el regulador británico de las telecomunicaciones<sup>48</sup>.

#### IV. EL DEBATE SOBRE LA MODERACIÓN DE CONTENIDOS EN LAS REDES SOCIALES

Si bien una mayor conciencia social y política sobre los problemas que plantean los contenidos en las redes a individuos, grupos, comunidades y Estados democráticos, ha sido clave en la formulación de las iniciativas de reforma antes citadas (y lo seguirá siendo, seguramente, en la materialización de las mismas, cuando eso ocurra y con el contenido final con que se produzca), el debate de ideas y propuestas –jurídicas, filosóficas, regulatorias o económicas– sobre esa cuestión ha tenido, y sigue teniendo, aparte de valor intelectual y ético, una relevancia práctica, ya que ha contribuido a aumentar el interés y el conocimiento (esto es, la conciencia) sobre la moderación de contenidos en las redes sociales y ha inspirado principios y normas que han sido incorporados a los proyectos de nueva legislación.

El debate en cuestión no es solamente académico, aunque sí lo sea de forma predominante: participan en él profesores destacados de Derecho constitucional norteamericano, como Cass Sunstein, Tim Wu o Jack Balkin, de Derecho europeo como Alexandre de Streel y especialistas en Derecho y Tecnologías, como Frank Pasquale, Kate Klonick, Julie Cohen o Nicolas Suzor, pero también otro tipo de profesionales, como la periodista Rana Foroohar o el escritor e informático Jaron Lanier. No se encuentra, tampoco, estrictamente limitado a unas fechas concretas, aunque todas las aportaciones

---

<sup>48</sup> Según el comentarista Mark Scott, la iniciativa británica «intenta situarse a medio camino entre las normas del proyecto de Ley de Servicios Digitales de la Unión Europea y el enfoque más *laissez-faire* de los Estados Unidos» *Politico*, 12 de mayo de 2021.

principales son relativamente recientes, en torno a los últimos cinco años, ni circunscrito a un área geográfica muy específica, si bien hay una mayoría de autores norteamericanos, algunos europeos y unos pocos australianos y canadienses. No sigue el debate, por lo demás, un itinerario de argumentación, crítica y refutación, y sucesivas réplicas y dúplicas, pero sí participa de un mismo marco de análisis y de unas preocupaciones comunes y en él coinciden monografías extensas, como las de Sarah T. Roberts (2019) o Tarleton Gillespie (2018) con artículos, comentarios y notas de distinta extensión y empaque.

Los autores que participan en el debate sobre moderación de contenidos en redes sociales recurren con cierta frecuencia a analogías y metáforas para ilustrar, reforzar o justificar sus propuestas. A fin de cuentas, es razonable trazar una relación de semejanza entre cosas distintas (por ejemplo, la televisión por ondas terrestres y el Internet) para defender o criticar un régimen normativo determinado para la más nueva de ellas, o menos regulada. Igual ocurre cuando se pretende hacer una comparación tácita entre un concepto con un sentido recto ya establecido y otro distinto (por ejemplo, la «plaza del pueblo» o el «libre mercado» con el Internet) con la intención de atribuir a uno de los términos de comparación, habitualmente el menos explorado, notas o elementos de la caracterización o de la regulación del otro. En el caso de la analogía, se busca el apoyo para una determinada regulación en el pasado, mientras que en el de la metáfora, se proyecta un cierto presente hacia el futuro. Así, se han buscado analogías de las redes sociales, que son plataformas privadas que ofrecen de manera pública y gratuita servicios y productos a usuarios de todo tipo y a cambio de contraprestación a anunciantes y publicitarios, con entidades como la *company town*, comunidad comercial y residencial creada por una empresa, propietaria además de la mayor parte de sus casas y comercios, para uso público o el *shopping mall*, espacio privado comercial que puede ser ocasionalmente foro de expresión, y ello para argumentar que las redes sociales tienen la consideración de «agentes estatales» y están, en consecuencia, vinculados por la Primera Enmienda, de manera que no pueden afectar con su actividad de moderación la libertad de expresión de los usuarios<sup>49</sup>. Lo mismo ha

---

<sup>49</sup> Resulta interesante la caracterización que hace el Tribunal Federal de Apelación para el Segundo Circuito en *Knight Institute v. Trump* (2019) de la cuenta gubernamental

ocurrido con otras industrias o negocios, como el agua, la electricidad, los ferrocarriles, las navieras, la radio, la televisión o las telecomunicaciones, para mantener que, al igual que todas ellas, las redes sociales son «canales esenciales de comunicación», a las que se puede aplicar, respectivamente, el régimen del servicio público (*utilities*), la doctrina de las instalaciones esenciales (*essential facilities*), el derecho común del transporte (*common carriage*), la doctrina de la equidad (*fairness doctrine*) del audiovisual o el principio de neutralidad de la red (*net neutrality*)<sup>50</sup>.

En cuanto a la metáfora, aparecen en el debate con alguna frecuencia, al igual que ha ocurrido antes en la jurisprudencia del Tribunal Supremo de los Estados Unidos<sup>51</sup>, referencias a conceptos de prestigio, bien asentados en la historia de la filosofía política y jurídica, como «neutralidad», «contrato social», «constitucionalismo», «esfera pública» o «mercado de las ideas», que se usan como apoyo, e incluso fundamento, de propuestas de regulación. Esto no tiene nada de extraño ya que, como señala Oreste Pollicino, «el lenguaje metafórico encaja muy bien con el razonamiento jurídico, pero debe de manejarse con propiedad y cuidado» (2020:14). Así, si la atención se fija en el «mercado de las ideas», una expresión hoy generalizada para aludir a la situación descrita por John Stuart Mill donde los individuos intercambian ideas libremente como mejor forma de separar lo verdadero de lo falso, cuesta concluir que las redes sociales, por

---

del Presidente Trump en Twitter como «foro público limitado».

<sup>50</sup> Sobre este punto, ver Keller (2019:15-17).

<sup>51</sup> Sobre el concepto de «plaza pública», el Tribunal Supremo ha caracterizado en *Packingham v. North Carolina* (2017) a las redes sociales como la «moderna plaza pública» y ha reconocido que «el ciberespacio es el lugar más importante [...] para el intercambio de puntos de vista».

Sobre el mercado de ideas, es famoso el disenso del Juez Oliver Wendell Holmes Jr. en *Abrams v. United States* (1919): «...pero cuando los hombres se han dado cuenta de que el tiempo ha trastornado a muchas convicciones en lucha, pueden llegar a creer, incluso más de lo que creen en los fundamentos mismos de su propia conducta, que el deseado bien supremo se alcanza mejor mediante el libre comercio de ideas – que la mejor prueba de la verdad es el poder del pensamiento para lograr la aceptación en la competencia del mercado, y que esa verdad es la única base sobre la cual sus deseos pueden llevarse a cabo con seguridad». Ver también, al respecto, *Gertz v. Welch* (1974), donde se dice que «bajo la Primera Enmienda no existe algo que sea una idea falsa. Por perniciosa que pueda parecer una opinión, dependemos para su corrección, no de la conciencia de jueces y jurados, sino de la competencia de otras ideas».

su vinculación inescindible con la publicidad comercial, la continua recogida de datos personales de sus usuarios, la monitorización de sus interacciones y los complejos algoritmos que filtran, ordenan y priorizan el contenido que en ellas se vierte, sean un «mercado» en el que las ideas compitan libremente<sup>52</sup>. Ocurre que, como dice Vincent Blasi, «los mercados de “bienes y servicios” generan precios y niveles de producción», mientras que «el mercado de las “ideas” genera una serie de creencias individuales y, de alguna forma, [...] observaciones y argumentos»; son dos planos tan diferentes que la metáfora que trate de relacionarlos tenderá a ser engañosa (2004:8). Otros argumentos críticos de la capacidad explicativa del «mercado de las ideas» se refieren a la distancia que existe entre el sistema económico vigente cuando el concepto empezó a usarse en la filosofía política y jurídica (el capitalismo liberal de mediados del siglo XIX al primer tercio del siglo XX) y el sistema de las primeras décadas del siglo XXI (neoliberalismo, o capitalismo de vigilancia, o capitalismo de la información), que invalida cualquier propósito de comparación, o a la asimetría de poder y capacidad de influencia que preside la relación entre redes sociales y usuarios. Algo parecido ocurre con el concepto de «esfera pública», que se desarrolló en la posguerra de la Segunda Guerra Mundial, tras la publicación en 1962 del libro *Strukturwandel der Öffentlichkeit* de Jürgen Habermas<sup>53</sup>. Allí, Habermas sostenía que la Edad Moderna había contemplado una transformación estructural de la manera en que «lo público» y «la publicidad» se practicaban, y que la emergencia de la «esfera pública burguesa», coincidente y, en buena medida, constitutiva del proyecto filosófico de la Ilustración, suponía que los individuos privados podían y debían discutir materias de interés común, de manera racional y desinteresada, y así contribuir a conseguir mejores comunidades políticas. La «esfera pública» no solo era un «ideal normativo», una aspiración elevada que perseguir, sino un «hecho histórico», algo que realmente se había producido en una época concreta: la *coffeehouse* inglesa, el salón francés, el club de cena alemán. Pero, como señala Brian Cowan, lo que una vez se pensó

<sup>52</sup> Sobre este punto, ver Lombardi (2019: 209).

<sup>53</sup> La obra fue traducida al inglés como «The Structural Transformation of the Public Sphere» y al español como «Historia y crítica de la opinión pública. La transformación estructural de la vida pública».

como un fenómeno producido por una dialéctica histórica inexorable, se entiende ahora como algo que puede encontrarse en tiempos y espacios muy diferentes, cada uno con sus propias complejidades y variaciones locales. Así, los conceptos o formulaciones posteriores a Habermas de la esfera pública, como algunos de los que superficialmente aparecen en el debate sobre la moderación de contenidos, son más pluralistas y están menos enraizados en la narrativa del desarrollo histórico de Europa, y tienden a ver la relación entre «públicos» y «Estados» como mutuamente constitutiva, más que como la relación de dos entidades separadas o incluso antagónicas (Cowan, 2021). Se trata, también aquí, de trasplantar un concepto de prestigio, que ahora se refiere a un lugar donde se produce un debate crítico racional sobre asuntos de interés público, a las redes sociales. Pero si bien es innegable que las redes sirven, entre otras muchas cosas, para la difusión de ideas e información social, económica o política, tampoco se puede afirmar que sea, ni mucho menos, el contenido mayoritario que circula a su través. Las redes sociales no están pensadas ni diseñadas para hacer posible y facilitar la deliberación, y pese a que Facebook diga que la finalidad de sus normas comunitarias es «crear un espacio donde todo el mundo pueda expresarse y dar a conocer su opinión» y Twitter proclame que su objetivo es «servir a la conversación pública», la estrategia de ambas compañías se fija por objetivos comerciales y financieros, por número de usuarios, horas pasadas en la red e ingresos generados por la publicidad: esas son las métricas que determinan para sus directivos el éxito de las redes sociales y no su contribución a la consecución del bien común a través de la participación y la deliberación. Por si hiciera alguna falta, Frances Haugen la antigua empleada de Facebook a quien se ha hecho referencia al principio de este trabajo, lo ha vuelto a recordar recientemente.

Los autores que toman parte en el debate, en fin, no utilizan un mismo concepto de regulación sino que, por lo general, usan el término en un sentido más bien amplio y poco preciso. La regulación, en este contexto, se entiende como una intervención pública en la actividad privada de moderación que realizan las empresas de redes sociales, y que puede ir desde ser su complemento hasta operar como una sustitución plena de aquella. No se trata, pues, de una regulación entendida como «control sostenido y focalizado que ejerce

una agencia pública sobre actividades valoradas por la comunidad», según la definición clásica de Philip Selznick, sino que es vista desde «una perspectiva más amplia de aplicación de normas públicas y privadas dirigidas a la solución de problemas específicos» (Feyck y Werle, 2010: 525).

### 1. Una cuestión previa: la neutralidad

El primer embate teórico que recibe la moderación, entendida como capacidad de la red social para retirar el contenido que no sea apropiado según sus condiciones de servicio, y priorizar, ordenar, presentar o jerarquizar ese contenido según sus normas comunitarias y políticas de uso, proviene de la neutralidad. En Internet, la idea de neutralidad puede aplicarse en diferentes niveles: así, en el de la infraestructura que despliegan y mantienen los operadores de telecomunicaciones y aseguran la conectividad y el transporte de las señales<sup>54</sup>, en el de los terminales que diseñan y producen los fabricantes de *hardware* y permiten disfrutar de servicios y productos en las redes, y, por último, en el de esos mismos servicios y productos que ofrecen las plataformas *online* y todo tipo de diferentes empresas a los usuarios finales. Pues bien, en el ámbito de los contenidos, tercero de los niveles reseñados, la neutralidad supondría permitir la circulación, sin restricciones, preferencias, interferencias ni cualificaciones de ningún tipo, de todo contenido legal por las redes sociales.

En este sentido, Frank Pasquale, crítico de las grandes tecnológicas y autor del libro *The Black Box Society* (2015), uno de los tempranos alegatos escritos contra el control que ejercen Gobiernos y empresas sobre los ciudadanos a través del uso de algoritmos secretos y *big data*, afirma que el objetivo central de la idea de neutralidad es impedir que intermediarios gigantescos distorsionen el comercio privado o la esfera pública, simplemente en virtud de su tamaño, de los efectos de red o de sus capacidades de vigilancia (2016), y se muestra de acuerdo con la apreciación del Consejo Nacional Digital francés (Conseil National du Numérique) de que «el principio de

---

<sup>54</sup> El concepto de «neutralidad de la red» implica que todas las comunicaciones transmitidas a través de una red han de ser tratadas de igual forma, con independencia de su contenido, aplicación, servicio, terminal, emisor y receptor (Wu, 2003: 141).

neutralidad debe de insertarse en el desarrollo de las plataformas digitales, para que su crecimiento no dificulte el uso de Internet como foro de creación, libre expresión e intercambio de ideas»<sup>55</sup>.

Dos especificaciones del principio de neutralidad son la algorítmica y la de las interfaces de programación de las aplicaciones. La neutralidad algorítmica, según Tarleton Gillespie, se justifica porque a medida que nuestro discurso público *online* tiene lugar, cada vez más, en redes sociales, y que las empresas propietarias recurren a algoritmos complejos para gestionar y organizar cantidades masivas de contenido generado por los usuarios, que codifican opciones políticas determinadas, se hace preciso, si queremos que tales algoritmos sean dignos de confianza, que sean neutrales. O, lo que es lo mismo, y dicho así parece que sea tarea sencilla, que realicen un cálculo simple, neutral, sin filtros y sin intervención humana ni necesidad de manipulación que los haga «correctos» según cualquier criterio predeterminado o conformes con los intereses de las empresas que prestan el servicio (Gillespie, 2018). A su vez, la neutralidad en las interfaces de programación de las aplicaciones (*application programming interface* o *API*), que son descripciones técnicas sobre cómo usar o interactuar con determinadas aplicaciones, y que permiten, en última instancia, que aplicaciones sucesivas sean construidas de manera que puedan dialogar, entenderse y ser compatibles entre sí, es defendida por Jonathan Zittrain (2008). Para este autor, la falta de neutralidad en las *API* pone en peligro la creatividad de Internet, limita su desarrollo y conduce a la consolidación de redes cada vez más cerradas y sometidas a tecnologías patentadas, por lo que una manera muy directa y efectiva de evitarlo sería que todas las aplicaciones estuvieran, en el plano lógico, al alcance de todos por igual.

Por el contrario, Jack Balkin considera que la neutralidad solamente debe aplicarse a la provisión de los servicios básicos de Internet, a la conectividad y a la interoperabilidad, mientras que las redes sociales deben comprometerse en la moderación de los

---

<sup>55</sup> Otra versión de la neutralidad de las plataformas consiste en que estas no deberían poder discriminar entre quienes las utilizan para acceder a los mercados o vender en ellos sus productos o servicios, o que no deberían favorecer los propios productos y servicios de las plataformas frente a terceros que usen estas para acceder y vender en los mercados. Este sería el caso, por ejemplo, de Amazon en el comercio electrónico.

contenidos, algo que pueden hacer bien o mal, pero no eludir si actúan responsablemente. Balkin ve a las empresas de redes sociales como «curadores», organizaciones que tienen que cuidar «cosas valiosas» (como los conservadores de un museo, por ejemplo), y también como editores del discurso público. Para este autor, pues, el objetivo no es tratar de alcanzar una neutralidad ilusoria en cuanto a los contenidos, sino modular la organización y los incentivos de las empresas de redes sociales para orientarlas hacia la consecución de bienes públicos (2020). También se muestra en contra de la idea de neutralidad en los contenidos Daphne Keller, para quien una norma de difusión obligatoria (*must-carry*) aplicable a todo aquello que no fuera estricta e indudablemente ilegal equivaldría a una nacionalización de las redes sociales (2019).

## 2. *Las posiciones en el debate sobre moderación de contenidos*

### A) El pesimismo radical

Hay autores enormemente críticos con las redes sociales, que las consideran tóxicas para los individuos y absolutamente irrecuperables para las sociedades democráticas. Jaron Lanier, informático, persona muy conocida entre los activistas de Internet y autor de un libro rotundo y provocador (*Ten Arguments for Deleting Your Social Media Accounts*, 2018), es uno de ellos. Lanier sostiene que las redes sociales son «imperios de modificación de las conductas» orientados a la adquisición de la atención de sus usuarios, a la indagación sobre sus vidas, a colmar de contenidos sus mentes, a dirigir interesadamente su comportamiento para ganar dinero con disimulo y engaño, a crear falsas multitudes y, en última instancia, una sociedad falsaria. Por ello, propone abandonar las redes sociales, porque para el individuo suponen, entre otras cosas, la pérdida del libre albedrío, la destrucción de la capacidad de empatizar, la infelicidad, la pérdida de la dignidad económica o la imposibilidad de participar en la política<sup>56</sup>. Ningún sentido tiene, pues, desde esa perspectiva, entrar en los detalles de cómo regular los contenidos en las redes sociales y mejorar la actividad de moderación.

---

<sup>56</sup> En esta línea hipercrítica, conjugando una diversidad de planos en el análisis, ver Peirano (2019).

## B) La intervención pública en las redes sociales

En el siguiente escalón de autores críticos con el funcionamiento de las redes sociales, pero ya dentro del marco del debate, están quienes defienden una intervención pública en la moderación de los contenidos, con distintos grados de intensidad. En sus formulaciones más extremas, hay autores que proponen que el Estado se haga cargo y preste el servicio de redes sociales o que estas sean tratadas como servicios públicos que ofrecen acceso y moderación de contenidos a cambio del pago de una suscripción anual.

Rana Foroohar, por ejemplo, cree que es hora de que las empresas de redes sociales admitan que no son la moderna plaza pública, pese a la caracterización en contrario hecha por el Tribunal Supremo en *Packingham v. North Carolina* (2017), sino negocios publicitarios que monetizan contenidos, como cualquier otro negocio de medios, ya sea radio, prensa o televisión. Para Foroohar, las redes sociales tienen capacidad más que suficiente para evitar la difusión de contenidos ilegales y dañinos, y si no invierten lo bastante en ello es porque tal objetivo no está entre sus prioridades. Por consiguiente, esta autora cree que ha llegado el momento de adoptar medidas de largo alcance, como puedan ser la separación estructural de contenidos y comercio, que a su juicio servirá para lograr un panorama cultural más sano y competitivo, la regulación de las redes sociales como servicios públicos, un régimen de tributación riguroso y exigente y la conclusión de un *new deal* digital y una carta de derechos digitales, sin entrar en estos dos últimos puntos en mayores detalles (2019). Por su parte, Frank Pasquale considera que Facebook y Google han alcanzado el papel de reguladores globales de la comunicación, imponiendo de una manera prácticamente irrestricta normas a usuarios y anunciantes, razón por la cual resulta, a su juicio, de crucial importancia una intervención estatal en la esfera pública digital, que no puede quedar en manos de «plutócratas más interesados en los retornos económicos de sus acciones y los avances de la inteligencia artificial que en las instituciones democráticas básicas y en la sociedad civil» (2016). Para Tim Wu, uno de los académicos con trayectoria más extensa y sólida en el estudio de la materia y actual consejero del presidente Biden en política de tecnología y competencia, Facebook, Google o incluso Twitter son monopolios de información, no empresas normales, ya

que «comercian» en expresión individual y resultan claves para el funcionamiento de la democracia, por lo que merecen un tratamiento especial. El éxito de estas empresas y el nivel de concentración que han producido en el mercado ha tenido como efecto la aparición de nuevos cuellos de botella y de un problema de escasez de canales de comunicación, siendo así, según Wu, que si bien ahora resulta más fácil hablar, es más difícil ser escuchado si no se tiene acceso (o se impide el acceso) a los sitios donde está la mayoría de la gente. Wu, al igual que Forohar, es partidario de medidas de separación estructural que establecerían una saludable distancia entre quienes controlan la red y la infraestructura, por una parte, los terminales y los puntos de acceso, por otra, y, en fin, la creación de los contenidos. Aunque Wu tampoco descende a detallar cómo se podría implementar este principio de separación, parece bastante claro que el mismo conllevaría una labor de desintegración y reordenación de las redes sociales (2018). Dos autores menos conocidos, Blayne Haggart y Natasha Tusikov, que estudian sobre todo los regímenes de Canadá y Australia, creen también que Facebook o YouTube son monopolios, se manifiestan a favor de su regulación estatal y dicen, entre otras cosas, que la nacionalización de las redes sociales como bienes públicos es una opción a considerar (ponen como ejemplo el caso de la radiotelevisión canadiense), que los Gobiernos deberían implicarse más en la regulación de los algoritmos –para que estos respondan a prioridades determinadas democráticamente, en vez de reflejar los motivos económicamente egoístas de las empresas–, que la intervención gubernamental en las comunicaciones debe ir más allá de maximizar la interconexión y la interoperabilidad, que no toda regulación es un ataque a la libertad de expresión y que, en definitiva, los Gobiernos y las sociedades tienen legítimas razones para regular tanto Internet, en cuanto infraestructura y conjunto de redes, como los monopolios que han crecido en Internet, hasta dominar sectores enteros de actividad, ya se trate de búsquedas, redes sociales o publicidad *online*<sup>57</sup>. Para terminar con este grupo, Cass Sunstein, uno de

<sup>57</sup> B. Haggart y N. Tusikov (2021): «Regulating the Digital Economy in Three Parts: The Conclusion», [www.cigionline.org](http://www.cigionline.org)

Ver también B. Haggart (2018): «Why not nationalize Facebook», *The Conversation*, 28 de marzo de 2018.

los constitucionalistas más prestigiosos, citados y prolíficos de los Estados Unidos, enfatiza la necesidad, en su reflexión sobre las redes sociales, de distinguir «ciudadano» de «consumidor» y «libertad» de «elección». Para Sunstein, si se acepta que la libertad de expresión está entre los fundamentos de la ciudadanía democrática y que no puede ceñirse o identificarse con la soberanía en las elecciones del consumidor, entonces la regulación gubernamental de Internet no tiene por qué resultar inaceptable (o al menos, no una regulación que sea neutral desde el punto de vista del contenido). A Sunstein le preocupa, desde el punto de vista democrático, que los ciudadanos no tengan suficiente exposición a contenidos, materiales y opiniones diferentes de las que, por pura preferencia personal, elijan, que como sociedad se llegue a carecer de una base común de experiencias que impida compartir, y que desaparezca la capacidad de atender a cuestiones sustantivas sobre políticas públicas o sobre principios de moralidad que deban inspirar el gobierno de la sociedad. Sunstein cree que se debe aspirar a un «debate público más rico» y defiende la «libertad de los antiguos», entendida como participación activa y comprometida en la «república» y, en consecuencia, propone diversas medidas como crear dominios en Internet dedicados a la deliberación, publicar las conductas relevantes de las empresas que gestionan las redes sociales, fomentar la autorregulación voluntaria, subvencionar públicamente determinadas páginas web, imponer normas de disposición obligatoria (*must-carry*) en temas educativos y de interés público, y usar de forma creativa *links* y otros recursos tecnológicos, para llamar la atención de los ciudadanos sobre puntos de vista distintos a los que, comúnmente y de forma casi exclusiva, tienen acceso (2017).

### C) La reforma público-privada de la moderación

Los autores incluidos bajo esta rúbrica mantienen puntos de vista distintos sobre el estado actual de la moderación de contenidos, siendo el de Klonick probablemente el más favorable, y realizan aproximaciones diferentes al objeto de estudio, destacando por su objetividad la del equipo dirigido por De Streel, en su informe para el Parlamento Europeo. Pero un rasgo fundamental que les une a todos es su apoyo a una mezcla de medidas, unas adoptadas en el ámbito de las empresas y otras aprobadas como legislación o regulación de

las mismas, como la opción preferida y más viable para mejorar ese estado de situación.

La contribución de Ben Wagner a la obra colectiva *Digital Dominance. The Power of Google, Amazon, Facebook and Apple* (2018), que trata precisamente de las implicaciones económicas, sociales y políticas del control que esas grandes empresas ejercen sobre la esfera digital, consiste en un análisis del impacto de los «intermediarios dominantes» en que se han convertido las redes sociales, en la expresión individual y en el discurso social y colectivo. Wagner pide transparencia y mecanismos procesales de salvaguardia de los derechos de los usuarios, similares a los que existen en los ordenamientos jurídicos, ya que el papel dominante de las plataformas hace que estas no sean ya simplemente empresas privadas, sino proveedoras de bienes públicos importantes. Consecuentemente, según este autor, hay que regular para reflejar esa dualidad<sup>58</sup>.

Para Nicolas Suzor, profesor e investigador australiano, autor de un libro importante sobre Internet y derecho –*Lawless. The Secret Rules that Govern Our Digital Lives* (2019)– y miembro, además, del Consejo de Supervisión de Facebook, las redes sociales son intermediarios de Internet que, a través del diseño de arquitectura, la capacidad de influir en normas y prácticas sociales y el poder de aplicar sus propias condiciones de servicio, normas comunitarias y políticas de todo tipo, tienen el poder de configurar y conformar la conducta y la comunicación en Internet. Suzor sigue, aquí, el planteamiento precursor de Lessig, para quien «derecho», «mercado», «normas sociales» y «arquitectura» son las cuatro fuerzas que regulan Internet, y el «código» (de programación) equivale a «derecho», ya que determina las posibilidades de uso, identificación, monitorización, o prestación de servicios o acceso a la información en Internet, entre otros extremos (1999). Bajando un nivel mayor de concreción, Suzor considera que el proceso de moderación de contenidos es falible, como lo son las propias leyes, pero, a diferencia de lo que ocurre con estas últimas, no está rodeado de una infraestructura

---

<sup>58</sup> En el mismo sentido Barack Obama en 2017, al urgir una reflexión de las redes sociales sobre sí mismas, que reconozca que son tanto un bien público como una empresa comercial, citado en V. Dhar (2021) «Nationalize Facebook and Twitter as public goods», *The Hill*, 15 de enero de 2021.

diseñada para disminuir la influencia de las decisiones individuales, la subjetividad de los decisores y la posibilidad de que aparezca el sesgo ideológico. Las normas que rigen el funcionamiento de las redes sociales no se determinan democráticamente ni están sujetas a revisión judicial, a lo que hay que añadir, según este autor, que el proceso de moderación es singularmente opaco. Ante una situación en la que los intermediarios *online* juegan un papel tan fundamental en la forma en la que los usuarios utilizan sus servicios, Suzor cree que hay que encontrar una manera para que las redes sociales tomen sus decisiones legítimamente. Para ello, y puesto que considera el «constitucionalismo», antes que nada, un modelo de limitación de poderes, propone uno de tipo digital que, sobre la base de los derechos humanos, sea capaz de limitar el poder de las plataformas *online*, mejorar y legitimar sus procesos de gobernanza, guiar a los Estados en la aprobación de las leyes necesarias, mejorar la transparencia, ampliar la participación en el proceso de toma de decisiones sobre moderación de contenidos y forzar a las empresas, en fin, a crear sus propios «límites constitucionales». Además de estas medidas, y de forma un tanto genérica y voluntarista, Suzor propone movilizar a las partes interesadas en este propósito reformista, resistir a los esfuerzos gubernamentales dirigidos a introducir mayores restricciones en Internet, crear nuevas colaboraciones entre distintas comunidades, grupos y organizaciones, así como alumbrar nuevas instituciones que sirvan para controlar a las grandes empresas tecnológicas y obligarlas a que rindan cuentas de su gestión, todo orientado al fin de alcanzar un Internet vibrante, diverso y responsable.

Por su parte, Kate Klonick, en un artículo muy influyente publicado en 2018 califica a las redes sociales como los «nuevos gobernadores» de la libertad de expresión, propone abandonar las tradicionales doctrinas y analogías regulatorias al uso para justificar su regulación y, en cambio, cree que hay que ver a las plataformas que prestan servicios de contenidos como «sistemas de gobernanza», responsables de dar forma y permitir la participación en una nueva cultura democrática, pero con una responsabilidad muy limitada hacia sus usuarios. Para Klonick, las empresas de redes sociales son entidades privadas que se autorregulan y que están económica y normativamente motivadas para reflejar la cultura democrática y las

expectativas de sus usuarios sobre la libertad de expresión. La autora opina que los mayores problemas que plantea este sistema privado de gobernanza a la cultura democrática radican en una doble ausencia de igualdad de oportunidades efectiva en cuanto al acceso y a la participación en las redes y de responsabilidad frente a sus usuarios, así como en la inexistencia de rendición de cuentas. La solución de estos problemas no va a llegar, según Klonick, por la vía de modificaciones parciales o de una redacción alternativa de la sección 230 CDA, o a través nuevas interpretaciones judiciales de la Primera Enmienda, sino mediante simples cambios en la arquitectura y en los sistemas de gobernanza de las plataformas. Klonick es optimista respecto a la capacidad de las redes sociales para evolucionar de una manera positiva, y cita como ejemplo el tránsito desde unas normas de decisión y conducta basadas sobre «criterios vagos, caprichosos y formulados *ad hoc*» a normas precisas y en constante proceso de mejora. En fin, si las empresas fallan en implementar esos cambios de diseño y gobierno propio que Klonick considera necesarios, habrá que abrir la puerta a la regulación, que tendrá que ser equilibrada, pragmática y realista, y partir, por ende, de lo que las plataformas hacen en sus operaciones cotidianas.

Gillespie, en su monografía sobre las redes sociales (2018: 197 y ss.), afirma que hay muchas cosas que las redes pueden hacer para mejorar la moderación de contenidos, como contratar más moderadores humanos, ofrecer más transparencia sobre sus actuaciones, contar con software de detección de contenidos inapropiados más preciso... Pero, en última instancia, tales medidas serían arreglos menores, una especie de parches, porque lo fundamental del mecanismo de las redes sociales seguiría averiado. Gillespie piensa en medidas más ambiciosas, de un mayor alcance, aunque reconoce, con una sinceridad que desarma, que las mismas pueden resultar «políticamente insostenibles, económicamente extravagantes y que, con toda seguridad, no van a ocurrir nunca», como el diseño de una transparencia «deliberada y accionable», la distribución colectiva de la moderación (para que sea una responsabilidad compartida por los usuarios de las redes), la protección de los usuarios cuando cambian de plataforma (algo que, por cierto, es norma habitual

en las empresas de telecomunicaciones desde hace años)<sup>59</sup>, el rechazo del fundamento económico de la popularidad o la incorporación de diversidad real a las redes sociales. Gillespie propone, además, que las redes realicen una labor de recolección de las contribuciones de valor cívico que publican, así como que identifiquen y den visibilidad a los temas que preocupan a los usuarios, utilizando a tal efecto datos e inteligencia artificial. Este autor cree, a fin de cuentas, que las redes sociales deberían ser «custodios de Internet», esto es, organizaciones capaces de identificar asuntos de interés, apoyar un debate sobre los mismos enriquecido con datos y constituirse, así, como una especie de facilitador de la discusión pública.

Más claro y concreto como ejemplo de propuesta orientada a la reforma público-privada de la moderación de contenidos en las redes sociales es el informe preparado por Alexandre de Streel y una serie de colaboradores, de la Universidad de Namur y otras instituciones, a solicitud de la Comisión sobre Mercado Interior y Protección de los Consumidores del Parlamento Europeo (2020). En particular, en el informe se recomienda fortalecer la responsabilidad procesal de las plataformas *online*, incrementar el papel que juegan en cuanto a los contenidos los usuarios y los denominados «alertadores fiables» (*trusted flaggers*) para identificar los que resulten cuestionables, alinear responsabilidades con riesgos, mejorar la efectividad de la monitorización y de la aplicación de las normas en el cumplimiento de la ley y adoptar medidas complementarias de transparencia, de apoyo a periodistas y prensa, y de colaboración entre redes sociales, sociedad civil, academia y organizaciones no gubernamentales y demás partes interesadas. También en Europa, en una línea parecida, pero con un acento algo más crítico, Teresa Quintel y Carsten Ulrich (2020) ven a las redes sociales como importantes habilitadores de la libertad de expresión, con poderes casi-públicos, como «guardianes de acceso» (*gatekeepers*) a Internet y destacan que sus decisiones sobre contenidos están profundamente ensambladas en sus arquitecturas y afectadas, en la misma medida, por intereses comerciales. Por ello, creen que la autorregulación ya no resulta suficiente, siendo

---

<sup>59</sup> MacCarthy propone que el Congreso apruebe leyes procesales, de defensa de los consumidores, para regular prácticas de moderación de contenidos por las plataformas (2020).

precisa una mayor supervisión regulatoria. Para estos autores, hay que transitar hacia un régimen de corregulación, para ganar, no solamente en eficacia y mejora de resultados, sino también en términos de legitimidad democrática. Propugnan Quintel y Ulrich, en definitiva, la imposición a las redes sociales de una obligación general de cuidado (*duty of care*) definida conforme a principios de interés público y garantía de derechos fundamentales, y con inclusión de medidas de transparencia, rendición de cuentas y análisis de riesgos.

#### D) Los incentivos públicos para la reforma de las empresas

¿Pero, por qué hay que moderar el contenido?

Para responder a esta pregunta hay que entender bien, dice Jack Balkin, qué son las redes sociales y qué funciones cumplen (2020). Según este autor, las redes sociales son instituciones clave en la esfera pública digital del siglo XXI y, desde esta perspectiva, tienen las funciones de facilitar la participación, organizar la conversación y editar y asegurar la calidad de su contenido. Para Balkin, la esfera pública es el «hogar de la libertad de expresión», el lugar donde se realizan los valores de participación, cultura democrática y creación y difusión del conocimiento, que para él equivale al libre «mercado de las ideas»<sup>60</sup>. Pero una esfera pública no funciona sin instituciones dignas de confianza, dirigidas por profesionales y guiadas por normas de interés público, siendo así que, para Balkin, en la esfera pública del siglo XXI, ni existe el tejido conectivo de las instituciones necesarias para salvaguardar los valores de la libertad de expresión, ni las redes sociales cumplen sus funciones públicas, porque están guiadas por imperativos de mercado. Los modelos competitivos, sin intervención pública, no producirán por sí solos la clase de cultura y el conocimiento necesario para la realización de los valores de la libertad de expresión; o, por decirlo en otras palabras, producirán demasiados daños y muy pocos beneficios en términos de libertad de expresión, con independencia de que se financien a través de la publicidad o del

<sup>60</sup> Balkin parece seguir aquí el planteamiento de Thomas I. Emerson en su clásico *The System of Freedom of Expression* (1970), donde considera la libertad de expresión como un medio para conseguir la autorrealización individual, un proceso esencial para avanzar en el conocimiento y descubrir la verdad y un elemento esencial para que todos los miembros de una sociedad participen en el proceso de toma de decisiones que les afecten.

pago de suscriptores. Para Balkin, la conclusión es que hay que regular las redes sociales, porque nos interesa y preocupa la esfera pública digital, y el objetivo de tal regulación no puede ser otro que el de crear incentivos para que las redes sean instituciones responsables y dignas de confianza, capaces de promover una «sana y vibrante» esfera pública digital. ¿Cómo se consigue, pues, que empresas cuyo modelo de negocio está basado sobre la recolección masiva de datos y una publicidad fundamentada en comportamientos y pautas de conducta previamente observados, lleguen a ser ese tipo de instituciones? Pues Balkin responde que ofreciéndoles incentivos para que adopten normas profesionales y orientadas al bien público, y haciéndoles internalizar los costes que imponen a la sociedad, a través de medidas de derecho de la competencia (como incrementar el número de actores en el mercado, proceder a la separación estructural de las redes sociales de mayor tamaño...), privacidad (imposición de deberes de confidencialidad, cuidado de los datos y lealtad hacia los usuarios...) y de protección de consumidores (como una mayor inversión en moderación).

En una línea que guarda semejanzas con el planteamiento de Balkin, Dypayan Ghosh (2020) sostiene que lo fundamental para regular correctamente las redes sociales es centrarse en medidas que tengan impacto económico real sobre su modelo de negocio, en los ámbitos de la privacidad, la transparencia y la competencia. Este autor, que como otros tantos (e igualmente sin mayor detalle) propone también un nuevo contrato social digital, sostiene que Facebook está encantada con el debate sobre los contenidos, porque distrae la atención de otro debate más sustancial para la empresa, que es el relativo a la regulación económica. Ghosh cree que el tema de los contenidos es de carácter administrativo y no intelectual, (aunque añade, de forma que parece incongruente con lo anterior, que ha de abordarse para «salvar vidas, evitar la discriminación y proteger las elecciones»), y ha de plantearse atendiendo a políticas regionales y culturas diversas.

Siva Vaidhyathan, quien en la parte positiva apunta que Facebook se ha convertido en el medio a través del cual más de dos mil millones de personas conducen sus vidas políticas, se conectan con quienes comparten intereses, pasiones, identidades y afiliaciones, se entretienen e informan, añade, en la negativa, que ha amplificado alguno de los peores hábitos de las personas y, sobre todo, ha perju-

dicado la capacidad para pensar colectivamente sobre los problemas sociales y políticos, y ha empobrecido instituciones de confianza para filtrar la calidad de la información, como la prensa (2018). Vaidhyathan es escéptico ante las medidas regulatorias inspiradas por el derecho de la competencia, que califica como una «molestia menor» para las empresas, o relativas a los contenidos, que en su opinión suponen una vía aún más complicada e inefectiva, y apoya, sin embargo, una regulación de la privacidad como la de la Unión Europea. Su propuesta, en definitiva, consiste en limitar la capacidad de Facebook de rastrear todo lo que los individuos hacen y dicen en la red, así como las maneras en que puede usar los datos personales para influenciar conexiones sociales y actividades políticas. O dicho de una manera más sintética, hacer que Facebook importe menos<sup>61</sup>.

#### E) Las propuestas participativas y comunitarias

En este epígrafe se consideran autores que enfatizan la participación y el compromiso de usuarios —en su condición de tales y como ciudadanos—, comunidades, asociaciones y organizaciones no gubernamentales como elemento capital para la reforma y mejora de la moderación de contenidos en las redes sociales.

Shannon Vallor, filósofa de la tecnología, constata que empresas como Facebook, YouTube o Twitter han transformado prácticas sociales, políticas e informativas, a nivel individual e institucional, por lo que cabe preguntarse si fomentan el ejercicio de la razón deliberativa, a la manera de Habermas, si fragmentan la esfera pública mediante la formación de «cámaras de resonancia» (*echo chambers*) y «filtros burbuja» (*filter bubbles*) y si facilitan actitudes políticas como el activismo o la desobediencia civil, entre otras cosas. Vallor estima, en cualquier caso que, si las redes sociales van a promover el fortalecimiento de la esfera pública «habermasiana» y las virtudes de un discurso público razonado (y ese es, en su propio análisis, un gran «si» de carácter condicional), los usuarios tendrán que movilizarse para explotar esa oportunidad y cultivar nuevas normas, valores y virtudes de cultura

<sup>61</sup> S. Vaidhyathan (2021), «What if regulating Facebook fails», *Wired*, 2 de julio de 2021.

cívica, o por el contrario preparase para la llegada de la autodestrucción o de la tiranía impuesta por una masa ignorante (2015).

Daphne Keller (2019), por su parte, propone modelos híbridos de reforma, como atemperar la intervención de las plataformas de redes sociales, de manera que solo actúen contra contenidos que puedan considerarse muy ofensivos, así como aplicar leyes de moderación solo en el caso de que sean justas, transparentes y no discriminatorias, dejar siempre espacio en las redes para discursos minoritarios y de colectivos desfavorecidos, o abrir los servicios de la redes a terceros, para que estos puedan, a su vez, desagregar, introducir novedades y más competencia en el mercado, al igual que se hizo a finales del siglo pasado en las telecomunicaciones, y en línea con la propuesta de neutralidad en las interfaces de programación de las aplicaciones (Zittrain), antes mencionada. También autores como Kyle Langvardt (2019), que defiende la intervención gubernamental en las redes sociales y Harold Feld (2019), igualmente partidario de una profunda reforma del actual marco legal, hacen propuestas que implican una mayor participación social o comunitaria, como la entrega del control de la moderación de los contenidos a los usuarios, el primero, o la obtención de una diversidad de opiniones y contenidos representativos de amplios sectores de la sociedad, como resultado del uso de algoritmos más flexibles e «imperfectos» (*wobbly*), el segundo.

Por último, Sarah T. Roberts, en su obra *Behind the Screen. Content Moderation in the Shadows of Social Media* (2019: 201-202, 2017), probablemente el estudio más completo, detallado e incisivo sobre la cuestión, afirma que entender la moderación y a las personas que realizan ese trabajo es crucial para hacerse una idea cabal de Internet tal cual es. Y añade que si hoy esa actividad ha ganado prominencia en la conciencia pública y a los ojos de legisladores y reguladores en todo el mundo ha sido, en buena parte, porque una serie de investigadores interesados en el impacto social de Internet han hecho de ello un área de estudio propia, a menudo con poco apoyo y, en ocasiones, hasta con la hostilidad activa de la academia y de las empresas tecnológicas. Roberts se pregunta si la moderación de contenidos, realizada en nombre de las redes sociales más grandes del mundo, puede constituir un baluarte efectivo contra la manipulación, la explotación y la propaganda (que a veces tiene efectos mortíferos) o si todo lo negativo asociado a las

redes viene incluido en ellas por defecto, como algo inescindible de la operación de una industria que se basa en contenidos generados por usuarios y que, en consecuencia, puede instrumentalizarse fácilmente para producir resultados nefastos. Roberts, tras descartar la posibilidad de limitar la cantidad de contenido publicable en las redes sociales, por no considerarla realista y factible, ya que los contenidos son extremadamente valiosos para aquellas, propone implicar a las bibliotecas y a los bibliotecarios en la solución de los problemas de moderación, ya que son profesionales con buena formación y naturalmente orientados hacia el servicio público, que pueden asistir a los usuarios en sus necesidades de información, sin incurrir en los conflictos de interés que afectan a los empleados o contratados por las redes sociales.

#### F) Los defensores de la autorregulación

Aunque nadie llegue a defender, por ideal, inmejorable o simplemente bueno, el estado actual de la regulación de contenidos en las redes sociales, hay autores que consideran que mantener la autorregulación de las empresas, con las menores cortapisas, junto al interés que estas tienen en rebajar críticas de usuarios y partes interesadas y frenar amenazas de intervención de los poderes públicos, es la mejor de las alternativas posibles.

Así, Jeff Kosseff (2019) afirma que ni los jueces, que no responden lógicamente ante los usuarios de redes sociales por las sentencias que dicten imponiendo obligaciones o responsabilidades a las redes, ni los legisladores, que no tienen la capacidad ni están situados en la mejor posición para hacer juicios editoriales, ni tampoco el Gobierno, que carece de incentivos claros y precisos al respecto, deben ser guardianes de acceso (*gatekeepers*) del discurso *online*. Kosseff estima que esa función la tienen que cumplir las empresas de redes sociales, porque tienen la relación directa con los usuarios, quienes de alguna forma son sus «clientes», de manera que si no cumplen con sus expectativas podrán llegar a verse fuera del mercado, como les ocurrió a precursores como Friendster o MySpace, de los que ya pocos se acuerdan.

En la misma línea, John Samples (2019), que publica en el Instituto Cato, una entidad de simpatías «libertarias» (término equivalente, en la cultura política norteamericana, a ultra liberalismo o liberalismo conservador) sostiene que, pese a que lo pidan demócratas y republi-

canos al mismo tiempo, el Gobierno no debería regular la moderación de contenidos en las redes sociales, ni siquiera para alcanzar objetivos de equidad, equilibrio o de otro tipo. Primero, y sobre todo, porque la Primera Enmienda y la sección 230 de la CDA limitan enormemente la capacidad gubernamental para regular el discurso en Internet. Y luego, porque a este autor no le parecen convincentes ninguno de los argumentos que se dan a favor de la intervención: ni el del poder de mercado –ya que las redes sociales hoy populares y dominantes podrían dejar de serlo mañana y existen, además, alternativas para quienes no tengan acceso o hayan sido excluidos de una determinada red–, ni el de la creación de «filtros burbuja» (*filter bubbles*) que impiden la deliberación democrática –ya que su evidencia empírica no es sólida, y si lo fuera habría pocos remedios para ese problema que fuesen compatibles con la Constitución–, ni el de la lucha contra el discurso violento –que puede ser efectivamente regulado por el Gobierno, pero limitado a terrorismo y casos severos–, ni el de la preocupación por la interferencia extranjera en la elecciones –que a juicio de Samples se ha exagerado indebidamente, más allá de que los extranjeros pueden expresarse en términos constitucionalmente admisibles con ciertas condiciones–, ni el de la prevención de los daños causados por la desinformación o el discurso del odio –que excede la jurisdicción del Gobierno–. Para Samples, las redes sociales tienen una presunción de legitimidad para retirar de sus espacios contenidos que resulten indeseables o dañinos y les corresponde, dice, impedir que intervengan funcionarios públicos deseosos de suprimir o restringir la libertad de expresión en nombre del «bien público». Samples afirma que a los líderes de esas empresas les puede resultar incómodo o penoso tener que tomar ese tipo de decisiones, que son complejas y peligrosas; pero, a fin de cuentas, es su trabajo y nadie más puede hacerlo por ellos.

## V. CONCLUSIÓN

En el debate sobre la moderación de contenidos en las redes sociales es frecuente el recurso a analogías con regímenes como los del transporte, la prensa, la televisión o las telecomunicaciones así como a categorías como las de «servicio público», «instalaciones esenciales» o «neutralidad de la red». También lo es la invocación de conceptos

con tradición en la historia de la filosofía política y jurídica, como «contrato social», «constitucionalismo», «esfera pública» o «mercado de las ideas», de forma metafórica y, en general, eufemística. Unas y otras, analogías y metáforas, sirven de apoyo, e incluso basamento, a muchas de las propuestas examinadas. Pero sin negarles por completo cualquier tipo de utilidad o valor explicativo, su uso, en particular en el caso de las metáforas, que tienen un mayor componente de sugestión o de proyección hacia el futuro, entorpece más que aclara la cuestión debatida y, en definitiva, no ayuda especialmente a encontrar vías plausibles para su solución. El recurso a analogías y metáforas en este debate hace que, en ocasiones, se puedan dar por supuestas ciertas afirmaciones sin estar suficientemente justificadas (por ejemplo, ¿quién ha atribuido funciones públicas a las redes sociales?), y se idealicen rasgos de los conceptos que se comparan o proponen como respuesta para un problema determinado, sin suficiente concreción (así, ¿qué supone hablar de un «contrato social digital», o un «constitucionalismo digital», aparte de que suene bien y garantice, de entrada, una aceptación casi unánime?).

Una buena respuesta del Estado democrático al problema de la moderación de los contenidos por las redes sociales no debe fundamentarse en propuestas que nieguen a las empresas realizar esta actividad, como puedan ser las de nacionalización, imposición de obligaciones de servicio público, o exigencia de neutralidad respecto a todos los contenidos que no sean abiertamente ilegales. En un plano normativo, el Estado no debe interferir en la libertad de las empresas que ofrecen servicios de redes sociales y ha de mantenerse neutral tanto respecto al estilo, orientación, normas de conducta y valores que una red social proponga como en cuanto a las preferencias, opiniones y gustos que sus usuarios expresen en un marco de libertad y respeto a la ley. No parece que el establecimiento y gestión de redes sociales públicas (alternativas o sustitutivas de las redes privadas actuales) deba ser considerada una función esencial del Estado democrático, o al menos es algo que no ha quedado convincentemente demostrado en el debate, y, como cuestión de hecho, no existe seguridad alguna de que el Estado sea capaz de gestionar eficazmente tales servicios. Tampoco son plausibles, en términos prácticos, las propuestas que pretenden, creativa y novedosamente, encomendar la moderación a una hipotética comunidad de actores públicos y privados, vagamente

descrita en el mejor de los casos, o socializar, dispersar o «distribuir» esta actividad, para intentar la superación, de esta forma, del conflicto de intereses ínsito en el modelo de negocio de las redes (negocio de publicidad versus «bien público»). La evanescencia o la falta de concreción de muchas de ellas contrastan frontalmente con la realidad y las magnitudes de las redes sociales, que nos dan cuenta de las dificultades y limitaciones con que se topan las iniciativas individuales y la indiferencia o pasividad generalizada de los usuarios frente a las demandas de mayor participación y compromiso en la gestión de las redes. Pero la solución no puede ser, desde luego, seguir confiando, sin más y frente a casi todas las evidencias, en las bondades, al menos comparativas, de la auto-regulación, así como en la capacidad de regeneración de las grandes empresas tecnológicas, o en el cumplimiento de sus compromisos, frecuentemente ignorados. El Estado democrático debe, por el contrario, exigir mediante ley, e incluso normas constitucionales, mayores responsabilidades a las empresas de redes sociales. Esa es la línea que siguen, en lo principal, las iniciativas de reforma del marco legal de la moderación de contenidos que aquí se han considerado, y también un buen número de las aportaciones al debate sobre la cuestión. El Estado tiene que permitir a las empresas como Facebook, YouTube o Twitter tomar decisiones legales, con todas las consecuencias, sobre el tipo de servicios y productos que quieran ofrecer a sus usuarios. Pero también ha de sancionar sus incumplimientos, tratarlas sin la deferencia con que lo ha hecho hasta bien recientemente, y recuperar un espacio público para la participación y la deliberación social, no interviniendo foros ya existentes, como puedan ser las propias redes, ni planteando la creación de otros nuevos, sino fortaleciendo las instituciones y los procesos de la democracia representativa.

## BIBLIOGRAFÍA

- BALKIN, J. (2014). Old School/New School Speech Regulation. *Harvard Law Review*, 127, pp. 2296-2342.
- (2020). How to Regulate (and Not Regulate) Social Media. *Knight First Amendment Institute at Columbia University*. <https://knightcolumbia.org/content/how-to-regulate-and-not-regulate-social-media>
- BALDWIN, R., CAVE, M. y LODGE, M. (2012). *Understanding Regulation. Theory, Strategy and Practice*. Oxford University Press.
- BLASI, V. (2004). Holmes and the Marketplace of Ideas. *Supreme Court Review*, 2004, pp. 1-46.
- BRADFORD, A. (2020). *The Brussels Effect. How the European Union Rules the World*. Oxford University Press.
- COHEN, J.E. (2019). *Between Truth and Power. The Legal Constructions of Informational Capitalism*. Oxford University Press.
- COWAN, B. (2021). Public Sphere. En A. BLAIR, P. DUGUID, A.J. GOEING y A. GRAFTON (eds.), *Information: A Historical Companion*. Princeton University Press.
- DASKAL, J. (2019). Speech Across Borders. *Virginia Law Review*, 105 (8), pp. 1605-1666.
- DE STREEL, A. et al. (2020). *Online Platforms' Moderation of Illegal Content Online*. Study for the committee on Internal Market and Consumer Protection, Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament. doi:10.2861/831734
- FEICK, J. y WERLE, R. (2010). Regulation of Cyberspace. En R. BALDWIN, M. CAVE y M. LODGE (eds.), *The Oxford Handbook of Regulation* (pp. 523-547). Oxford University Press.
- FELD, H. (2019). *The Case for the Digital Platform Act: Market Structure and Regulation of Digital Platforms*. Roosevelt Institute.
- FOROOHAR, R. (2019). *Don't be evil. How Big Tech Betrayed its Founding Principles and All of Us*. Currency.
- GHOSH, D. (2020). *Terms of Disservice. A New Digital Social Contract for Our Technological Future*. Brookings Institution Press.
- GILLESPIE, T. (2018). *Custodians of the Internet. Platforms, Content Moderation and the Hidden Decisions that Shape Social Media*. Yale University Press.
- KELLER, D. (2019). Who Do You Sue? State and Platform Hybrid Power over Online Speech. *Aegis Series Paper*, 1902. [https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech\\_0.pdf](https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech_0.pdf)

- KLONICK, K. (2018). The New Governors: The People, Rules and Processes Governing Online Speech. *Harvard Law Review*, 131, pp. 1598-1670.
- KOSSEFF, J. (2019). *The twenty-six words that created the Internet*. Cornell University Press.
- LANGVARDT, K. (2019). A New Deal for the Online Public Sphere. *George Mason Law Review*, 26, pp. 341-394.
- LANIER, J. (2018). *Ten Arguments for Deleting Your Social Media Accounts Right Now*. Macmillan.
- LESSIG, L. (1999). *Code and other Values of Cyberspace*. Basic Books.
- LOMBARDI, C. (2019). The Illusion of a “Marketplace of Ideas” and the Right to Truth, *American Affairs*, 3 (1), pp. 198-209.
- MACCARTHY, M. (2020). A consumer protection approach to platform content moderation in the United States. En B. PETKOVA y T. OJANEN (eds.), *Fundamental Rights Protection Online .The Future Regulation of Intermediaries* (pp. 115-139). Elgar.
- PASQUALE, F. (2016). Platform Neutrality: Enhancing Freedom of Expression in Spheres of Private Power. *DigitalCommons@UM Carey Law*, 1603.
- PEIRANO, M. (2019). *El enemigo conoce el sistema*. Debate.
- PINKUS, B. M. (2021). The Limits of Free Speech in Social Media. *Accessible Law*, Spring 2021 Issue. <https://accessiblelaw.untdallas.edu/limits-free-speech-social-media>
- POLLICINO, O. (2020). Metaphors and judicial frame: why legal imagination (also) matters in the protection of fundamental rights in the digital age. En B. Petkova y T. Ojanen (eds.), *Fundamental Rights Protection Online. The Future Regulation of Intermediaries* (pp. 2-15). Elgar.
- QUINTEL, T. y ULRICH, C. (2020). Self-regulation of fundamental rights? The EU Code of Conduct on Hate Speech, related initiatives and beyond. En B. Petkova y T. Ojanen (eds.), *Fundamental Rights Protection Online. The Future Regulation of Intermediaries* (pp. 197-229). Elgar.
- ROBERTS, S. T. (2017). Content moderation. En L. SCHINTLER y C. MCNEELY (eds.), *Encyclopedia of Big Data*, Springer.
- (2019). *Behind the Screen. Content Moderation in the Shadows of Social Media*. Yale University Press.
- SAMPLES, J. (2019). Why the Government Should Not Regulate Content Moderation of Social Media. Cato Institute.
- SUNSTEIN, C. (2017). *#Republic: divided democracy in the age of social media*. Princeton University Press.
- SUZOR, N. (2019). *Lawless. The Secret Rules that Govern Our Digital Lives*. Cambridge University Press.

- VAIDHYANATHAN, S. (2018). *Anti-social media. How Facebook Disconnects Us and Undermines Democracy*. Oxford University Press.
- VALLOR, S. (2015). Social Networking and Ethics. *Stanford Encyclopedia of Philosophy*. [https://www.researchgate.net/publication/335970815\\_The\\_impact\\_of\\_social\\_networking\\_on\\_business\\_and\\_business\\_ethics](https://www.researchgate.net/publication/335970815_The_impact_of_social_networking_on_business_and_business_ethics)
- VOLOKH, E. y FALK, D. (2012). Google: First Amendment Protection for Search Engine Search Results. *Journal of Law, Economics and Policy*, 8 (4), pp. 883-900.
- WAGNER, B. (2018). Free Expression? Dominant Information Intermediaries as Arbiters of Internet Speech. En M. MOORE y D. TAMBINI (eds.), *Digital Dominance: The Power of Google, Amazon, Facebook and Apple* (pp. 219-240). Oxford University Press.
- WU, T. (2003). Network Neutrality, Broadband Discrimination. *Journal of Telecommunications and High Technology Law*, p. 2, p. 141.
- (2018). Is the First Amendment Obsolete? *Michigan Law Review*, 117 (3), pp. 547-581.
- ZITTRAIN, J. (2008). *The future of the Internet— and How to Stop It*. Yale University Press.
- ZUBOFF, S. (2019). *The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power*. Public Affairs.