

**LA LIBERTAD DE EXPRESIÓN Y LAS REDES SOCIALES:  
DE LA DOCTRINA DE LOS PUERTOS SEGUROS  
A LA MODERACIÓN DE CONTENIDOS**

*FREEDOM OF EXPRESSION AND THE SOCIAL NETWORKS:  
FROM THE SAFE HARBOUR LAWS TO CONTENT MODERATION*

FRANCISCO VALIENTE MARTÍNEZ  
Universidad Pontificia Comillas  
<https://orcid.org/0000-0002-4596-8037>

Fecha de recepción: 26-10-21

Fecha de aceptación: 22-3-22

**Resumen:** *El auge de las nuevas tecnologías ha propiciado una revolución en la comunicación humana, que ha pasado a ser global, masivamente multilateral y esencialmente digital. Regular la actividad online es indispensable para proteger a los internautas, pues las grandes empresas multinacionales propietarias de las aplicaciones que predominan en Internet diseñan unas políticas de uso que condicionan el ejercicio de nuestros derechos fundamentales, de un modo muy particular nuestra libertad de expresión. Las distintas alternativas para regular el funcionamiento de los proveedores de estos servicios y conseguir un modelo normativo equilibrado y viable a nivel internacional es, por lo tanto, un reto que no puede posponerse más.*

**Abstract:** *The rise of new technologies has led to a revolution in human communication, which has become global, massively multilateral and essentially digital. Regulating online activity is essential to protect Internet users, as the large multinational companies that own the main social networks design conditions and policies of use that limit the exercise of our fundamental rights, particularly our freedom of expression. The various alternatives for regulating the operation of Internet service providers and achieving a balanced and viable regulatory model at international level is a challenge that cannot be postponed any longer.*

**Palabras clave:** Internet, libertad de expresión, redes sociales, moderación de contenidos, puertos seguros

**Keywords:** Internet, freedom of speech, social networks, online content moderation, safe harbor laws

## 1. INTRODUCCIÓN: LA REVOLUCIÓN DE LAS NUEVAS TECNOLOGÍAS

El 21 de noviembre de 1969, las Universidades de Stanford y Los Ángeles culminaron con éxito un experimento que transformó para siempre el mundo de las comunicaciones: el primer enlace funcional que permitía interactuar a los usuarios de dos dispositivos informáticos remotos. Este proyecto, bautizado como ARPANET, había sido impulsado por la Defense Advanced Research Projects Agency (DARPA) y se acabaría convirtiendo en lo que hoy conocemos como Internet. El uso militar de estas aplicaciones sigue siendo crucial para nuestra defensa, pero ha sido su generalización como medio de comunicación a nivel mundial lo que ha supuesto una verdadera revolución, pues las páginas web (las llamadas world wide web o www), los hipertextos (http) y el sistema de códigos identificativos individuales (url) posibilitaron una estructura de creación de contenidos global y accesible, además de ser prácticamente gratuita<sup>1</sup>.

Pero el éxito mundial de Internet no se explica sin tener en cuenta las constantes innovaciones que desde las dos últimas décadas del siglo pasado han vivido los terminales de acceso. Los ordenadores personales y los teléfonos móviles, cuyas características originales eran su alto coste, su exclusividad y su empleo empresarial, se ha convertido en bienes de consumo asequibles para el gran público. Esta combinación de factores, englobados bajo la expresión “nuevas tecnologías”, han permitido definir Internet como “la imprenta del siglo XXI”<sup>2</sup>, hasta el punto que a quienes ya tenemos cierta edad nos cuesta recordar cómo vivíamos antes de su difusión y a las nuevas generaciones les resulta poco menos que inimaginable estar sin sus dispositivos móviles.

El siguiente paso en este imparable proceso de innovación tecnológica fue la aparición de plataformas que permitían la interacción multilateral en tiempo real (los primeros “chats”), que evolucionaron hasta las redes sociales actuales. Estos espacios son hoy un verdadero universo paralelo en el que

---

<sup>1</sup> El 6 de agosto de 1991, Tim Berners-Lee, investigador del CERN en Suiza, publicó la primera página web (w3.org) y el protocolo de creación de contenidos para futuros portales. Su diseño permitía que desde cada terminal conectado a la Red se pudiese acceder a otras webs e intercambiar audios, imágenes, vídeos y textos. A la inmensa contribución técnica de estos trabajos, Berners-Lee sumó un gran aporte de carácter ético: renunció a patentar sus diseños, de forma tal que su uso se convirtió en gratuito.

Vid T. BERNERS-LEE, *Tejiendo la Red*, Siglo XXI de España, Madrid, 2000.

<sup>2</sup> A. G. PISCITELLI, *Internet, la imprenta del siglo XXI*, Gedisa, Barcelona, 2005.

millones de personas publicamos a diario información que antes era privada o compartida sólo en núcleos íntimos, tales como fotografías, aficiones, opiniones de la más diversa índole o incluso estados de ánimo.

Los datos no admiten discusión. A principios de 2021, Facebook contaba con 2.740 millones de usuarios, YouTube con 2.291 millones, Whatsapp con 2.000, Instagram con 1.221, TikTok con 689, Telegram con 500, Snapchat con 498, Pinterest con 442 y Twitter con 353, todo ello sin citar redes profesionales como LinkedIn, de contactos como Tinder o Bumble, u otras de carácter regional, como la rusa VK. Cada usuario pasa conectado un promedio de dos horas y veinte minutos al día<sup>3</sup>. Además, un mismo individuo suele tener varios perfiles en distintas redes sociales simultáneamente. Como consecuencia, muchas personas se han creado una vida virtual que ocupa buena parte de su tiempo, una especie de realidad alternativa que se entremezcla de forma inextricable con la denominada, incluso a veces despectivamente, vida real<sup>4</sup>. Este uso masivo y constante no podía pasar desapercibido para disciplinas tan diversas como las ciencias sociales, la estadística, la educación, las ingenierías, la economía y, por supuesto, el Derecho.

El desafío que el uso global de las nuevas tecnologías supone es mayúsculo y debe afrontarse desde un ámbito plural, incluso dentro de las propias ciencias jurídicas. Pero, de todos los enfoques posibles, quizás sea su afectación a los derechos fundamentales y, en concreto, a la libertad de expresión, su aspecto más relevante. Así lo ha puesto de manifiesto el Tribunal Supremo de los Estados Unidos, que llegó a sentenciar que “un ciudadano con conexión a Internet puede convertirse en alguien cuya voz resuena más fuerte que la de cualquier orador desde una tribuna”<sup>5</sup>.

A tenor de lo expuesto, los objetivos del presente trabajo son tres: por un lado, analizar el origen y el contenido de las normas que regularon Internet desde sus inicios; por otro, estudiar algunos de los desafíos que se han generado con el uso masivo e internacional de esta tecnología y, por último, analizar algunas de las respuestas que se están dando a los mismos.

---

<sup>3</sup> H. TANKOVSKA, “Global social networks ranked by number of users 2021”. Disponible en: <https://datareportal.com/social-media-users>, consultado 01-09-2021.

<sup>4</sup> A. CONSTANTE, “¿Qué son las redes sociales?”, en VV.AA., *Las redes sociales. Una manera de pensar el mundo*, Ediciones Sin Nombre, México D. F., 2013, pp. 13-36.

<sup>5</sup> *Packingham v. North Carolina*, 582 U.S. 2017.

## 2. LOS PROVEEDORES DE SERVICIOS DE INTERNET Y LA DOCTRINA DE LOS PUERTOS SEGUROS

Los internautas no son meros receptores de información, como puede ocurrir con los consumidores de prensa, radio o televisión, sino también creadores y divulgadores de contenidos mediante procesos comunicativos que se materializan millones de veces al día a lo largo y ancho del mundo. Cuando accedemos a Internet activamos una cadena cuyos eslabones son los proveedores de servicios de Internet (Internet Service Providers, genéricamente denominados ISP), cada uno con un rol que permite la difusión digital de información. En primer lugar, encontramos a los proveedores de acceso, encargados de suministrar el hardware básico (los routers, el sistema de cableado, el wifi...) y de su instalación. En segundo lugar, están los servidores de tránsito, también llamados conductores, que permiten la interacción entre el terminal de cada usuario y el punto de alojamiento de la información consultada<sup>6</sup>. En tercer lugar, encontramos los servidores de alojamiento que, pese a ser desconocidos para el gran público, realizan la crucial labor de albergar todos los contenidos<sup>7</sup>. En cuarto lugar, intervienen los proveedores de servicios en línea, mucho más conocidos, pues son los que permiten que realicemos actividades cotidianas como disponer de correo electrónico, leer la prensa, conectarnos a las redes sociales, etcétera (es decir, Facebook, Outlook, YouTube...). Por último, los llamados proveedores de búsqueda facilitan los listados de enlaces url que nos llevan directamente a la información solicitada (nadie puede dudar que Google es el rey de estos proveedores, pero existen más, tales como Yandex, Yahoo o Bing).

Regular estos procesos es muy complejo. La legislación vigente para cada particular que accede a Internet es la del país en el que se encuentra, pero lo más probable es que el buscador que emplee opere a nivel mundial, que las páginas web que dicho buscador ofrezca estén alojadas en diversos países y, también, que su creador o administrador resida en un lugar diferente. ¿Cómo aplicar entonces el principio de *lex loci*, sabiendo las grandes

---

<sup>6</sup> En España, los proveedores de acceso y tránsito son por lo general los mismos, es decir, compañías operadoras como Movistar, MasMóvil, Vodafone, Orange...

<sup>7</sup> Cuando alguien desea crear su propia web necesita adquirir los servicios de alguna de estas empresas, denominadas proveedores de hosting, que mantienen el portal activo a cambio de una pequeña renta (algunos de los principales son DreamHost, HostGator, Hostpapa, SiteGround o GoDaddy). Adicionalmente suele ser necesario registrar un dominio, es decir, una url propia e identificativa de cada portal.

Vid M. MACDONALD, *Creación y diseño Web 2010*, Anaya, Madrid, 2009, pp. 99-115.

deficiencias que, en este y otros aspectos, presenta la colaboración internacional policial y judicial<sup>8</sup>?

Pero esta cuestión no es reciente. Tal dilema ya estaba en la mente del legislador cuando, a finales del siglo pasado, se vio ante la necesidad de elaborar una reglamentación que fomentase el crecimiento de Internet, pero, a la vez, no dejase nuestros derechos desprotegidos. Bien podría decirse que este doble objetivo era inasequible, una casi imposible cuadratura del círculo que obligaba a dar preferencia a una de las dos metas. El criterio que resolvió el debate fue el *momentum*, pues una normativa restrictiva de Internet en sus orígenes podría haber asfixiado su difusión. Ese temor fue determinante para el nacimiento de la “doctrina de los puertos seguros” (*safe harbor laws*), según la cual, los proveedores de servicios no son responsables legales de las actuaciones de los distintos usuarios de la Red, salvo en tres casos excepcionales: que sean ellos mismos los generadores del contenido, que lo hayan modificado o que seleccionen al receptor<sup>9</sup>. La plasmación legal de este principio la encontramos en la Sección 230 de la *Ley de Telecomunicaciones* aprobada en 1996 por el Congreso de los Estados Unidos; un artículo que ha sido definido como las veintiséis palabras que crearon Internet<sup>10</sup>: “*No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider*”.

Desde su entrada en vigor, la Sección 230 ha sido la regulación fundamental del ciberespacio. Sus características son tres: en primer lugar, fue aprobada conjuntamente por demócratas y republicanos, lo que aseguró su pervivencia en los años siguientes; en segundo lugar, su razón de ser consistía en fomentar la innovación y el desarrollo de esta nueva tecnología, que no ha parado de crecer; en tercer lugar, había consenso en que debían ser los usuarios y proveedores de servicios, y no el Gobierno, quienes regulasen el funcionamiento de Internet<sup>11</sup>.

---

<sup>8</sup> J. CARRASCO MEDINA, “Derecho de autor e Internet: la muerte del principio de *lex loci protectionis*”, *UISRAEL Revista Científica*, núm. 19, 2019, pp. 25-38.

<sup>9</sup> R. XALABARDER, “La responsabilidad de los prestadores de servicios en Internet (ISP) por infracciones de propiedad intelectual cometidas por sus usuarios”, *Revista de Internet, Derecho y Política*, núm. 2, 2006, pp. 1-13.

<sup>10</sup> J. KOSSEFF, *The twenty-six words that created the Internet*, Cornell University Press, Londres, 2019.

<sup>11</sup> D. S. ARDIA, “Free Speech Savior or Shield for Scoundrels: An Empirical Study of Intermediary Immunity under Section 230 of the Communications Decency Act”, *Loyola of Los Angeles Law Review*, núm. 43, 2010, pp. 373-506.

Sin embargo, los legisladores estadounidenses eran muy conscientes de que Internet, incluso en esta etapa primigenia, se estaba empleando también de forma ilícita. Por un lado, se producían constantes vulneraciones de derechos de autor; por otro, había páginas web con contenidos peligrosos cuando no execrables. Además, el hecho de que los menores empezasen a acceder a la Red de forma habitual obligaba a tratar de poner coto a según qué contenidos. Por ello, la *Communications Decency Act* tipificó como delito la difusión de contenidos “indecentes” y “obscenos”, un evidente límite a la Sección 230. La American Civil Liberties Union entendió que esta redacción generaba inseguridad jurídica por fundamentarse en dos conceptos indeterminados y que, aunque persiguiese un fin legítimo, se exigía a los ISP un grado de control inviable<sup>12</sup>. Finalmente, el Tribunal Supremo, en un célebre fallo, resolvió que “no es exagerado concluir que el contenido en Internet es tan diverso como el pensamiento humano”, por lo que debe fomentarse el libre intercambio y acceso al mismo, de suerte que tanto el control previo como las restricciones imprecisas o subjetivas son contrarias a la Primera Enmienda. El Tribunal entendió la necesidad de proteger a los menores de contenidos inadecuados, pero a la vez consideró inconstitucional tanto responsabilizar a los proveedores de servicios por su difusión como obligarles a identificarlos entre los miles de portales existentes<sup>13</sup>.

Sin embargo, un año después se aprobó la Sección 512 de la *Digital Millennium Copyright Act*, que establecía el principio “notice and takedown”. A través de este mecanismo, cuando un proveedor de servicios recibía la notificación de que algún contenido vulneraba los derechos de propiedad intelectual, debía eliminarlo de inmediato o bloquear el acceso al mismo. De esta forma, no era necesario ya un monitoreo previo por parte de los ISP, pero sí se les exigía la capacidad de reaccionar una vez tenían conocimiento de alguna infracción. Si, por alguna razón, el proveedor mantenía accesible el contenido notificado, debía asumir la carga de la prueba para demostrar que no había vulneración normativa alguna<sup>14</sup>.

El resto del mundo siguió con atención cómo se resolvían estas cuestiones en los Estados Unidos. Al fin y al cabo, Internet debía ser global o no

---

<sup>12</sup> J. L. COX, “Defining a new medium of communication under the First Amendment: The Supreme Court tackles speech on the Internet in *Reno v. American Civil Liberties Union*”, *Nova Law Review*, vol. 22, núm. 3, 1998, pp. 761-782.

<sup>13</sup> *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997).

<sup>14</sup> J. WANG, *Regulation hosting ISP’s responsibilities for copyright infringement*, Springer, Nueva York, 2018, pp. 73-76.

sería, por lo que muchos otros países no dudaron en imitar la doctrina de los puertos seguros una vez se vio ratificada. Podemos encontrar dos ejemplos muy ilustrativos. Por un lado, Japón aprobó en 2001 una normativa que exime de toda responsabilidad al proveedor de servicios si no tenía conocimiento previo de que el contenido fuese ilícito<sup>15</sup>. Por otro, en 2008, la Federación Rusa estableció una normativa que excluye toda responsabilidad del prestador de servicios salvo que hubiese tenido conocimiento previo de la ilicitud, especificando además que, aun en este caso, afectaría únicamente a quienes cobren una tarifa por sus servicios pero no a quienes los prestasen gratuitamente (como por ejemplo hacen las empresas que dan soporte a blogs y también a las redes sociales)<sup>16</sup>.

El Consejo de Europa siguió unos pasos similares y publicó el 28 de mayo de 2003 la *Declaración sobre la libertad de comunicación en Internet*, que ha servido de pauta para las legislaciones nacionales posteriores y ha orientado la jurisprudencia del Tribunal Europeo de Derechos Humanos. Su Principio 1, sobre la regulación de contenidos, dice así: “Los Estados miembros no han de establecer restricciones a los contenidos en Internet que vayan más allá de las aplicadas a otros medios de difusión de contenidos”.

Por su parte, la Unión Europea reguló Internet mediante la *Directiva 2000/31/CE sobre Comercio Electrónico*. Sus artículos 13 y 14 dan una protección equivalente a la de los puertos seguros, en tanto en cuanto establecen que “cuando se preste un servicio de la sociedad de la información consistente en almacenar datos facilitados por el destinatario del servicio, el prestador de servicios no podrá ser considerado responsable de los datos almacenados”. A su vez, el artículo 15 prohíbe que los Estados miembros impongan una obligación general de supervisión previa, aunque sí existe el deber de frenar las actividades ilícitas si se tuviera conocimiento de ellas. Posteriormente, la *Directiva 2002/22/CE, relativa al servicio universal y de los derechos de los usuarios en relación con las redes y los servicios de comunicación electrónica* fijó el trato igualitario a todo contenido online, prohibiendo discriminaciones, restricciones o interferencias<sup>17</sup>. En lo referido al ejercicio específico de la libertad

---

<sup>15</sup> T. HAGIWARA, “Liability of Internet Service Providers”, *X Symposium International Softic*, 2001.

<sup>16</sup> V. NAUMOV y A. AMOSOVA, “Provider’s liability”, *Amcham News*, núm. 88, 2008, pp. 26-28.

<sup>17</sup> R. RODRÍGUEZ PRIETO, “De la ‘neutralidad’ a la ‘imparcialidad’ en la red. Un análisis crítico de la política de la UE sobre internet y algunas propuestas de mejora”, *Cuadernos europeos de Deusto*, núm. 57, 2017, pp. 217-246.

de expresión, la Unión Europea ha ido tomando medidas para contener los posibles abusos, de entre las que destaca la *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*, que en su artículo 1.2 señala que los Estados miembros “podrán” castigar únicamente las conductas que alteren el orden público, que sean amenazadoras, que sean insultantes o abusivas; y señala también en su artículo 5 que las personas jurídicas pueden ser responsables “cuando la falta de vigilancia o control” por su parte haga posible la materialización de una de estas conductas<sup>18</sup>.

En España, la *Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y de comercio electrónico*, transpuso la normativa europea y recogió lo ya expuesto sobre la responsabilidad de los proveedores de servicios digitales. Más adelante, *Ley 56/2007, de 28 de diciembre, de medidas de impulso de la sociedad de la información*, y el *Real Decreto Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias* desarrollaron esta protección al considerar que las plataformas vigentes por entonces, aunque mucho más sencillas y con menos posibilidades que las actuales, eran indudablemente prestadoras de un servicio<sup>19</sup>. Por último, la reciente *Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*, reconoce en su Título X la naturaleza fundamental de muchos de los derechos que se ejercen a diario en la Red, entre ellos, obviamente, la libertad de expresión.

En resumen, prácticamente toda la legislación sobre Internet ha sido claramente influenciada por el derecho estadounidense, lo cual ha favorecido una cierta homogeneización y un fácil diálogo institucional. La premisa central era que los prestadores de servicios no tenían responsabilidad sobre los contenidos digitales, pero se esperaba su colaboración en la persecución de posibles ilícitos. Como señalara el Consejo de Europa, las regulaciones no debían ser más restrictivas que las aplicadas en otros ámbitos. Pero nos es menos cierto que el uso generalizado de Internet ha provocado gran cantidad de supuestos en los que hay afectación de derechos y, si bien no son aceptables restricciones adicionales, tampoco parece razonable que haya menos.

---

<sup>18</sup> Además, en 2010, la Unión Europea creó el Organismo Europeo de Reguladores de Comunicaciones Electrónicas (BEREC), al que se dio autoridad para emitir dictámenes y recomendaciones, si bien carecen de carácter vinculante.

<sup>19</sup> A. ROIG, “E-privacidad y redes sociales”. *IDP. Revista de Internet, derecho y política*, núm. 9, 2009, pp. 42-52.



La homogeneidad internacional ha comenzado a resquebrajarse cuando los principales tribunales nacionales e internacionales han ido consolidando una dispar jurisprudencia en los conflictos surgidos en el entorno digital. Los dos casos Facebook Irlanda v. Schrems<sup>20</sup>, resueltos por el Tribunal de Justicia de la Unión Europea, son quizás el más clarificador ejemplo, pues pusieron de manifiesto que una norma comunitaria, incluso un tratado internacional, puede ser invalidada si se acreditan vulneraciones de derechos

---

<sup>20</sup> En el año 2000 la Comisión Europea aprobó el Acuerdo de Puerto Seguro, que regulaba el intercambio de datos personales entre EE.UU. y los países miembros de la Unión Europea (no debe confundirse, por tanto, con la doctrina de irresponsabilidad de los intermediarios antes comentada). El fundamento de esta libre circulación de datos se encontraba en el reconocimiento de un sistema de protección equivalente entre los distintos firmantes, por lo que la privacidad de los usuarios de Internet parecía quedar siempre garantizada. Sin embargo, Maximilian Schrems, un estudiante de Derecho austriaco, presentó una denuncia ante el Comisario Irlandés de Protección de Datos reclamando que, aunque él era usuario de Facebook Irlanda (país en el que se encuentra la sede europea de dicha empresa), sus datos estaban siendo tratados por la matriz, sita en Estados Unidos, donde el régimen de garantías es menor. El caso acabó ante el Tribunal de Justicia de la Unión Europea, que sentenció que, efectivamente, no existía una protección equivalente y que el Acuerdo de Puerto Seguro rebajaba la protección a la privacidad de los ciudadanos europeos de forma ilegítima (STJUE de 6 de octubre de 2015, Facebook Ireland v. Schrems, C-362/14). Como consecuencia, en 2016, la Comisión Europea aprobó un nuevo acuerdo marco entre la Unión Europea, Estados Unidos y Suiza, que se vino a llamar “Escudo de Privacidad”. Éste obligaba a las empresas estadounidenses a fijar nuevos mecanismos de protección para los datos de los ciudadanos europeos y creaba la figura del Mediador, al cual los posibles afectados podían acudir para defender sus intereses. La *Decisión 2016/1250 de la Comisión de 12 de julio de 2016 con arreglo a la Directiva 95/46/CE del Parlamento Europeo y del Consejo, sobre la adecuación de la protección conferida por el Escudo de la privacidad UE-EE. UU.* consideró que este nuevo modelo daba protección suficiente y afirmó que “la confianza y la protección de datos son imprescindibles, porque es lo que impulsará nuestro futuro digital”. Pero Schrems no estuvo de acuerdo: convertido ya en un reconocido activista en defensa de los derechos digitales y creador de la ONG “Europa versus Facebook”, denunció que seguían sin subsanarse las deficiencias en la protección de datos personales de los ciudadanos europeos. En julio de 2020, el TJUE acabó dándole nuevamente la razón y sentenció que la Unión Europea no podía asegurar que en Estados Unidos se aplicase “un nivel de protección sustancialmente equivalente al garantizado por la Carta”, por lo que el “Escudo de Privacidad” seguía siendo insuficiente (STJUE de 16 de julio de 2020, Facebook Ireland v. Schrems, C-311/18). Existía, pues, el riesgo de que Meta Platforms, Inc. y otros operadores estadounidenses dejaran de prestar sus servicios en Europa. Sin embargo, en marzo de 2022 la Casa Blanca y la Comisión Europea anunciaron un principio de acuerdo donde se esbozan las premisas que regularán en adelante las transferencias de datos.

Vid. I. GARCÍA-PERROTE MARTÍNEZ y T. G. GARCÍA-MICÓ, “Identidad, cesión de datos personales y la decisión Privacy Shield tras la STJUE Schrems II”, *InDret: Revista para el Análisis del Derecho*, núm. 3.2020, 2020, pp. 551-559.

fundamentales derivados de ella. En otras palabras, el TJUE ha establecido que la Carta de Derechos Fundamentales de la Unión Europea es una norma de referencia que los proveedores de servicios deben respetar<sup>21</sup>.

De todo ello podemos concluir que, en su génesis, la regulación de Internet tenía como fin desarrollar y generalizar su uso, permitiendo todo tipo de innovaciones y fomentando la participación de todo tipo de usuarios. Una especie de “creced y multiplicaos” que ha permitido que en apenas un par de décadas el mundo haya saltado de lo analógico a lo digital, se haya interconectado a nivel global como jamás antes lo había hecho y se haya abierto un universo de posibilidades cuyo techo no es posible aún avistar. Pero, conseguido este primer objetivo, cabe preguntarse si los daños y vulneraciones de derechos que han ido apareciendo por el camino deben seguir tolerándose y si no sería necesario y conveniente desandar algunos pasos.

### 3. CONSIDERACIONES SOBRE EL EJERCICIO DE LOS DERECHOS FUNDAMENTALES EN EL ENTORNO DIGITAL

#### 3.1. La libertad de expresión online: un ejercicio condicionado a un medio

El artículo 19 de la Declaración Universal de los Derechos Humanos establece que “todo el mundo tiene derecho a la libertad de opinión y expresión”; la Primera Enmienda a la Constitución estadounidense supuso la prohibición de aprobar leyes que la restrinjan; el Convenio Europeo de los Derechos Humanos y la Carta de Derechos Fundamentales de la Unión Europea señalan que este derecho corresponde a “toda persona”. Podemos concluir, por tanto, que todos somos sus titulares y disponemos de la facultad para beneficiarnos de sus efectos tanto frente al poder como frente a otros particulares<sup>22</sup>. Pero ello no supone que exista identidad en la protección que se da a la libertad de expresión a ambos lados del Atlántico, pues mientras que en Estados Unidos encontramos una defensa a ultranza de la misma, en el marco normativo europeo el valor predominante es la dignidad, fuente

---

<sup>21</sup> M. I. PUERTO y P. SFERRAZZA-TAIBI, “La sentencia Schrems del Tribunal de Justicia de la Unión Europea: un paso firme en la defensa del derecho a la privacidad en el contexto de la vigilancia masiva transnacional”, *Revista de Derecho del Estado*, núm. 40, 2018, pp. 209-236.

<sup>22</sup> VV.AA., *Teoría general de los derechos fundamentales*, Tecnos, 1ª ed., Madrid, 1994, p. 84.

de todos los derechos fundamentales<sup>23</sup> y mínimo irrenunciable a la hora de impedir la cosificación de la persona<sup>24</sup>.

El artículo 20 de nuestra Carta Magna señala que la libertad de expresión se “reconoce y protege”, pero establece también sus primeros y más obvios límites. En lo sustancial, dichos límites no cambian porque cambie el canal de distribución de los contenidos, de resultas que lo que está prohibido fuera de la Red está prohibido en ella. Por consiguiente, del mismo modo que ya no se admite la censura previa, no parecería necesaria una actuación activa, ex ante, que ordenase el ejercicio de esta libertad, pues difícil justificación jurídica encontraría la aplicación de estándares distintos que dependiesen exclusivamente del medio en que nos expresamos<sup>25</sup>.

Para comprender bien esta cuestión debemos entender la propia estructura de Internet<sup>26</sup>. En sus orígenes, se concibió como un nuevo espacio donde cada cual podía configurar su página web y tratar cualquier temática, de manera que el libre mercado de las ideas se manifestaba en todo su esplendor<sup>27</sup>. Es más: el diseño de algunos portales fue reconocido como obra original con creatividades propias y, por consiguiente, quedaron protegidos mediante el

---

<sup>23</sup> H. NOGUEIRA ALCALÁ, *Derechos fundamentales, bloque constitucional de derechos, diálogo interjurisdiccional y control de convencionalidad*, Ubijus Editorial, México D. F., 2014, pp. 23 y ss.

<sup>24</sup> M. GONZÁLEZ PASCUAL, *El Tribunal Constitucional alemán en la construcción del espacio europeo de los derechos*, Civitas, Madrid, 2010, pp. 28 y ss.

<sup>25</sup> A. BOIX PALOP, “La construcción de los límites a la libertad de expresión en las redes sociales”, *Revista de estudios políticos*, núm. 173, 2016, pp. 55 a 112.

<sup>26</sup> Resulta necesario diferenciar entre tres conceptos. Por un lado, la denominada *clearnet* o *surface web* estaría compuesta por todos aquellos enlaces rastreables, fundamentalmente los indexados en buscadores generalistas como Google. Un segundo concepto sería la *deepweb*, donde se engloba toda aquella información online que no es pública, como por ejemplo los correos que se mantienen en las bandejas de Outlook, las conversaciones de Whatsapp o Skype y los archivos guardados en espacios como Dropbox. Un segmento especial de esta *darkweb* sería el contenido diseñado para que su acceso sólo sea posible a través de navegadores especiales, lo cual no significa necesariamente que sea material ilícito, como muchas veces se afirma.

*Vid* P. Biddle, P. England, M. Peinado y W. Bryan, “The Darknet and the Future of Content Distribution”, *Microsoft Corporation*, 2002, pp. 2-8.

<sup>27</sup> Brillante en su sencillez, el planteamiento del Juez Oliver Wendell Holmes Jr. sostiene que la mejor política en materia de libertad de expresión es la ausencia de política, pues en una sociedad sin restricciones a la misma, frente a cada opinión polémica surgiría una opinión crítica y serían los ciudadanos quienes, al contrastarlas, aprenderían a ser libres en el discernimiento entre la verdad y la mentira, el acierto y el error.

*Vid* *Abrams v. United States*, 250 U.S. 616 (1919), voto particular del Juez Oliver Wendell Holmes Jr.

derecho a la creación artística<sup>28</sup>. Esta dimensión primigenia sigue existiendo pese a la progresiva y creciente complejidad del ciberespacio.

Un segundo escenario nació con los primeros chats, que se convirtieron en foros donde un número relativamente reducido de ciudadanos conversaban sobre cualquier cosa, compartían experiencias o pasaban largas horas en la compañía digital de desconocidos. Gracias en parte a ese anonimato hablaban con plena libertad, se establecían vínculos de confianza, se forjaban amistades virtuales y los usuarios más consolidados se convertían de facto en protectores de la paz en los mismos, con la potestad de expulsar a quienes incumpliesen unos protocolos básicos de corrección que recibieron el nombre de “netiquetas”<sup>29</sup>.

Una tercera dimensión se alcanzó cuando los medios de comunicación tradicionales lanzaron sus propias páginas web para ofrecer a sus lectores información en tiempo real. En cada noticia incorporaban además un foro temático para fomentar el debate público y que los internautas compartieran sus opiniones. Cuando algunos usuarios empezaron a aprovecharlos para realizar comentarios ofensivos, violentos o amenazantes, se planteó si, al producirse en un espacio dependiente de un medio de comunicación regulado por la libertad de información, seguía imperando la irresponsabilidad de los intermediarios. El Tribunal Europeo de Derechos Humanos entendió que determinados portales de Internet no podían ser tenidos por meros focos de expresión ciudadana, sino que cabía exigirles un control editorial, hasta el punto que la dejación de esta función podía devenir en culpa *in vigilando*. Se argumentó para ello que, aunque es incuestionable que la libertad de información se centra en la transmisión veraz de hechos noticiables, la redacción y la presentación que haga cada medio puede desencadenar comentarios hostiles, lo que justifica la exigencia de disponer de mecanismos de control y la obligación de emplearlos adecuadamente<sup>30</sup>. Sin embargo, posteriormente

---

<sup>28</sup> F. PANTALEÓN PRIETO y A. SOLER PRESAS, “La protección jurídica de las páginas web”, *Anuario de Derecho Civil*, vol. 54, núm. 3, 2001, pp. 1051-1104.

<sup>29</sup> A. CHARALABIDIS, *The Book of IRC: The Ultimate Guide to Internet Relay Chat*, No Starch Press, San Francisco, 2000, pp. 11-33.

<sup>30</sup> El portal de noticias Delfi fue acusado por las autoridades estonias de permitir que sus lectores publicasen en los comentarios de sus artículos mensajes que podían ser tenidos por un discurso del odio. Se atribuyó a Delfi un deber como editor pues, por las características de su foro, los usuarios no podían modificar los comentarios una vez enviados y, como podían ser anónimos, se favorecían estas acciones.

Asunto Delfi AS v. Estonia, STEDH de 15 de junio de 2015.

matizó este criterio y señaló que con el empleo efectivo de los mecanismos de notificación y retirada de comentarios ilícitos es suficiente, pues no todas las expresiones son iguales y la responsabilidad en sentido estricto se ciñe a las que sean verdaderas incitadoras a la violencia o la discriminación<sup>31</sup>. Lo cierto es que la distinción entre expresión e información online es más que necesaria, pues contribuye a distinguir las noticias ciertas o al menos diligentemente contrastadas de las meras opiniones, sobre todo si se refieren a los asuntos públicos<sup>32</sup>. Lo cual se refuerza con el deber de transparencia de los Estados sobre el funcionamiento de las administraciones públicas, para lo cual el empleo de vías digitales es crucial<sup>33</sup>.

Estas dimensiones no hacían sino replicar en el ciberespacio las manifestaciones convencionales de la libre opinión. Sin embargo, Internet supuso también la aparición de un nuevo escenario con el uso masivo de las redes sociales, auténticas transformadoras de la comunicación interpersonal. Las redes sociales son aplicaciones online que permiten crear un perfil individual identificativo en el que se puede volcar información personal a la par que se conocen las actividades que realizan otras personas, así como participar en distintos grupos de opinión o de intereses. Hay una gran diversidad de redes sociales, pero todas presentan rasgos comunes: en primer lugar, tienen vocación expansiva, esto es, aspiran a sumar cada vez más usuarios; en segundo lugar, su fundamento es la interconexión, tanto cualitativa (aumentando las funciones disponibles en cada perfil) como cuantitativa (fomentando que cada usuario contacte con más personas); en tercer lugar, superan las barreras físicas y de distancia tradicionales; en cuarto lugar, viralizan la

---

<sup>31</sup> Dos empresas inmobiliarias húngaras denunciaron a un medio de comunicación y a un buscador online después de que varios usuarios anónimos publicasen en sus foros comentarios muy críticos sobre su proceder comercial. Entendieron que se les había difamado impunemente y, en efecto, los tribunales nacionales resolvieron que los intermediarios eran también responsables del perjuicio causado. Sin embargo, el Tribunal Europeo de Derechos Humanos entendió que, aunque había una innegable afectación a la imagen de las empresas denunciadas, no había un mensaje socialmente dañino; además, valoró que los denunciados hubieran retirado los comentarios una vez se interpuso la demanda civil como un proceder suficientemente diligente.

Asunto Magyar Tartalomszolgáltatók Egyesülete e Index.hu ZRT v. Hungría, STEDH de 2 de febrero de 2016.

<sup>32</sup> SSTC 107/88, de 25 de junio, y 223/92, de 14 de diciembre.

<sup>33</sup> L. COTINO HUESO, "El reconocimiento y contenido internacional del acceso a la información pública como derecho fundamental", *Teoría y realidad constitucional*, núm. 40, 2017, pp. 279-316.

comunicación, pues un usuario puede acceder a los contactos de aquéllos con quienes ya está conectado, generando una auténtica telaraña que afecta incluso a quienes no participan de la misma pero aparecen en fotografías o publicaciones; en quinto lugar, exigen que cada usuario facilite una serie de datos personales, algunos como condición necesaria para crear el perfil y otros generados por el uso cotidiano de la aplicación; finalmente, suele ocurrir que, al solicitar la baja, no todos los datos se borran de forma automática y permanente<sup>34</sup>. Estas características han hecho de las redes sociales un espacio diferente a cuanto antes se había conocido, un punto de encuentro para personas de todo el mundo que buscan manifestar distintos aspectos de su personalidad, sus deseos y aspiraciones, a la par que se marcan tendencias, se identifican héroes y villanos, se ensalza y se difama, se ama y se odia. Un mundo donde las mejores virtudes de la libertad de expresión han alcanzado cotas sin precedentes, pero también sus riesgos<sup>35</sup>.

Es innegable que las redes sociales se han convertido en un medio muy común en la comisión de abusos. Ello se debe a varios factores: por un lado, la posibilidad de actuar anónimamente mediante perfiles falsos transmite una sensación de relativa impunidad que desinhibe los comportamientos hostiles; por otra parte, el anonimato en el acceso anima a los usuarios a buscar información que no buscarían, por ejemplo, en una librería tradicional;<sup>36</sup> en tercer lugar, desde que es posible la traducción inmediata de contenidos, la barrera idiomática tampoco es ya un obstáculo; además, las redes suelen contar con un eficaz buscador que ayuda a cada usuario a hallar la información que le interesa, del tipo que sea.

Ahora bien, es necesario recordar que los límites propios de la libertad de expresión siguen vigentes en el mundo virtual. No existe el derecho al insulto ni a realizar manifestaciones vejatorias o humillantes, ni tampoco se permite el discurso del odio, es decir, la creación y difusión los mensajes denigratorios dirigidos contra colectivos vulnerables. Existe también otro límite, que sería el mantenimiento del orden público<sup>37</sup>, es decir, “aquella

---

<sup>34</sup> J. M. BELTRÁN CASTELLANOS, “Aproximación al régimen jurídico de las redes sociales”, *Cuaderno electrónico de estudios jurídicos*, núm. 2, 2014, pp. 61-90.

<sup>35</sup> J. LANIER, *Diez razones para borrar tus redes sociales de inmediato*, Debate, Barcelona, 2020, pp. 77-88.

<sup>36</sup> M. A. MORETÓN TOQUERO, “El ciberodio, la nueva cara del mensaje del odio: entre la cibercriminalidad y la libertad de expresión”, *Revista jurídica de Castilla y León*, núm. 27, 2012, pp. 1-18.

<sup>37</sup> Asunto Barford v. Dinamarca, STEDH de 22 de febrero de 1989.

situación y estado de legalidad normal en que las autoridades ejercen las atribuciones que les son dadas y los ciudadanos las respetan y obedecen sin oponer resistencia”<sup>38</sup>. La vinculación del uso de las redes sociales con posibles alteraciones del orden público, éticamente justificables o no, ha quedado patente tras comprobarse cómo los principales movimientos sociales de los últimos años fueron posibles gracias a ellas. Fue a través de estas aplicaciones que ciudadanos de todo el mundo tuvieron conocimiento inmediato de las (mal) llamadas “primaveras árabes”, del movimiento 15-M o de los más recientes #blacklivesmatter, #metoo o #prayforAmazonia antes de que los medios de comunicación se hiciesen eco de estos hechos. Y, aunque no es menos cierto que estas reacciones ya ocurrían con anterioridad –a modo de ejemplo, podemos recordar las manifestaciones en España tras el secuestro y asesinato de Miguel Ángel Blanco a manos de la banda terrorista ETA en 1997–, es incontestable que las nuevas tecnologías han aumentado su número, intensificado su frecuencia y variado su temática.

Para combatir estos excesos, los países europeos han recurrido con frecuencia a la vía penal, ya sea mediante el agravamiento de las sanciones o incluyendo nuevos tipos. La concreción de esta corriente en España ha sido el actual artículo 510 CP, artículo extenso cuya redacción ha sido tildada de confusa, compleja, arbitraria y carente de proporcionalidad<sup>39</sup>. Buena parte de la doctrina se ha mostrado muy crítica con este proceder, pues, por su propia naturaleza, el Derecho Penal debe ser excepcional. Por ello, comparto la opinión de quienes piensan que la generalización de la vía penal para perseguir opiniones y mensajes, online u offline, acaba a la postre por afectar a la dimensión institucional de la libertad de expresión pues provoca, en última instancia, un “efecto desaliento” para quienes mantienen opiniones e ideas quizá polémicas pero sin duda legítimas<sup>40</sup>.

La necesidad de encontrar un procedimiento generalista, rápido y efectivo ha tenido como consecuencia que las redes sociales hayan tomado una decisión controvertida: establecer sus propias políticas de uso y de moderación de contenidos, de manera que el posible conflicto ya no enfrenta exclu-

---

<sup>38</sup> J. C. MONTALVO ALBIOL, “Concepto de orden público en las democracias contemporáneas”, *Revista jurídica Universidad Autónoma de Madrid*, núm. 22, 2010, pp. 197-222.

<sup>39</sup> F. VALIENTE MARTÍNEZ, *La democracia y el discurso del odio: límites constitucionales a la libertad de expresión*, Dykinson, Madrid, 2020, p. 305.

<sup>40</sup> R. ALCÁCER GUIRAO, *La libertad del odio. Discurso intolerante y protección penal de minorías*, Marcial Pons, Madrid, 2020, pp. 206-211.

sivamente a los internautas entre sí, sino que se produce también entre éstos y las empresas que gestionan las redes sociales.

### 3.2. La moderación de contenidos: un cuestionable mecanismo de regulación

La expresión “moderación de contenidos” engloba las prácticas dirigidas a la evaluación, categorización y posterior ratificación o eliminación de materiales tenidos por inadecuados de acuerdo con los criterios establecidos por los administradores de un portal de Internet. Esta acción puede clasificarse siguiendo varios criterios. Si atendemos al momento de la intervención puede ser previa a la publicación o posterior a la misma; si atendemos al origen de la revisión puede ser proactiva (realizada de oficio por los administradores) o reactiva (tras la denuncia de otro usuario); dependiendo del carácter del revisor puede ser manual (si la realiza un equipo humano), automatizada (si se hace en función de algoritmos) o híbrida<sup>41</sup>. Como se ve, se trata de un escenario complejo, con múltiples aristas y opciones donde, además, el servicio de moderación puede quedar subcontratado a otras empresas, abriendo un nuevo modelo de negocio derivado del ejercicio de la libertad de expresión<sup>42</sup>.

Ahora bien, ¿con qué derecho se arrojan las empresas propietarias de las redes sociales la legitimidad para establecer mecanismos de control de las acciones de sus usuarios? En otras palabras, ¿no deberían ser éstos los responsables únicos de sus publicaciones? Dos elementos nos ayudan a responder a esta pregunta: por un lado, las características de los contratos entre las redes sociales y sus usuarios; por otro, la concepción de las personas jurídicas como titulares de derechos fundamentales.

---

<sup>41</sup> J. ALBERO VALDÉS, “La libertad de expresión ante la sociedad algorítmica: una mirada al problema de la moderación de contenidos online”, *Cátedra José María Cervelló*, Law IE School, 2020.

<sup>42</sup> En 2020, la Cátedra José María Cervelló, que homenajea a este insigne gaditano, jurista y protector de las artes, dedicó, con gran acierto, la XIV edición de su certamen sobre derecho de los negocios a la libertad de expresión. Aunque a priori pudiera parecer una decisión contradictoria, estamos comprobando que, efectivamente, la difusión de los contenidos online no puede considerarse ya como el mero ejercicio de una libertad individual pues, por suerte o por desgracia, toda una industria dedicada al análisis, categorización, control y eliminación de los mismos está creciendo en su seno.

Vid XIV Premios José María Cervelló de Derecho de los Negocios. Disponible en: <https://es.ontier.net/ia/bases-xiv-premio-jose-maria-cervello.pdf> consultado: 10-09-2020.



Ya hemos comentado que cada acción de cada usuario es creadora de contenidos a priori protegidos por el ejercicio de la libertad de expresión<sup>43</sup>. La mayor parte de las redes sociales, al igual que muchos foros temáticos especializados y, cada vez más, la prensa online, imponen el registro como condición de acceso. A tal efecto, el formato universalmente aplicado es el de los contratos de adhesión. Este tipo de contratos ofrece las ventajas de la inmediatez en el servicio y la uniformidad contractual por más que haya multitud de usuarios, pero también saca a la luz los inconvenientes de la restricción de libertades, pues sólo deja al consumidor/usuario dos opciones: aceptar las reglas de cada plataforma o no formar parte de las mismas<sup>44</sup>. Es decir, quien crea un perfil formaliza un contrato y, por ende, asume un compromiso de aceptación de sus políticas y condiciones uso, las haya leído o no. Un consentimiento que además es pleno, pues abarca distintos ámbitos y ninguna red social admite un consentimiento parcial<sup>45</sup>.

La consecuencia ha sido una considerable autorregulación de los distintos proveedores a través de un régimen jurídico privado –una especie de *lex digitalis*– bajo cierta –pero tenue– supervisión estatal<sup>46</sup>. Las empresas propietarias de estas plataformas actúan con plena autonomía para decidir qué mensajes quedan vedados y las exigencias de comportamiento a sus miembros<sup>47</sup>. En otras palabras, los ISP pueden limitar –o autolimitar– el ejercicio de las libertades de sus usuarios por cualesquiera motivos, por ejemplo por decisiones tomadas en el ámbito de la libertad de empresa. El verdadero debate radica en si al hacerlo restringen en demasía los derechos de sus usuarios o si, como ya defendía parte de la doctrina cuando comenzaba el uso generalizado de Internet, no puede entenderse que ello signifique autocensura, ni mucho menos, censura<sup>48</sup>.

---

<sup>43</sup> A. K. YERGALIYEVA y T. K. ERDZHANOV, “Legal regulation in social network”, *World Science*, núm. 15, vol. 3, 2016, pp. 50-52.

<sup>44</sup> A. PLATERO ALCÓN y C. JIMÉNEZ ASENSIO, “El consumidor online en el ordenamiento jurídico español”, *Revista Nuevo Derecho*, núm. 21, 2017, pp. 1-23.

<sup>45</sup> B. RODRÍGUEZ ARROYO, “Las redes sociales: usuarios, términos y condiciones”, *EOB, Fashion, Luxury and Retail*, 2020. Disponible en: <https://enriqueortegaburgos.com/las-redes-sociales-usuarios-terminos-y-condiciones/> consultado 09-09-2021.

<sup>46</sup> J. A. CASTAÑEDA MÉNDEZ, “El efecto horizontal de los derechos fundamentales en el contexto de constitucionalización global del régimen jurídico privado digital”, *Revista Jurídica Mario Alario D Filippo*, núm. 15, pp. 29-47.

<sup>47</sup> J. BARATA I MIR, “Freedom of expression and digital platforms: challenges for content regulation in a global environment”, *Revista catalana de dret public*, núm. 61, 2020, pp. 1-13.

<sup>48</sup> X. O’CALLAGHAN MUÑOZ, “La autolimitación de la libertad de expresión”, *La Ley: Revista jurídica española de doctrina, jurisprudencia y bibliografía*, núm. 1, 1996, pp. 1559-1560.

Las personas jurídicas suelen emitir tres tipos de mensajes: un discurso comercial para dar a conocer sus productos y servicios; un discurso corporativo si se posicionan sobre asuntos públicos; y un discurso político si contribuyen a financiar los gastos de algún candidato durante un proceso electoral<sup>49</sup>. Para la jurisprudencia norteamericana clásica, nada de esto las hacía titulares de derechos fundamentales, pues las personas jurídicas se concebían como “un ser artificial, invisible, intangible y que existe sólo en virtud de la ley”<sup>50</sup>. Sin embargo, este status quo fue alterado en el año 2010: el controvertido fallo del Tribunal Supremo en el caso *Citizens United*<sup>51</sup> equiparó a personas físicas y corporaciones en su carácter de titulares del derecho a la libertad de expresión,<sup>52</sup> bajo el fundamento de que su ejercicio por cuantos

---

<sup>49</sup> FERGUSON III, C. Corporate Speech. “The First Amendment Encyclopedia”, *Middle Tennessee State University*, 2009. Disponible en: <https://www.mtsu.edu/first-amendment/article/939/corporate-speech>, consultado 09-09-2021.

<sup>50</sup> *Trustees of Dartmouth College v. Woodward*, 17 U.S. (4 Wheat.) 518 (1819).

<sup>51</sup> *Citizens United* es una organización sin ánimo de lucro que, durante las primarias presidenciales del Partido Demócrata celebradas en 2008, emitió un documental sumamente crítico con la entonces senadora y candidata Hillary Clinton. Tras su proyección en algunos cines, comenzó su distribución en DVD y a través de diversas plataformas de televisión online. El acuerdo con dichas plataformas consistía en la puesta a disposición del documental de manera gratuita para los espectadores a cambio de un canon que abonaría *Citizens United*. La Comisión Electoral Federal prohibió dicha distribución alegando que este acuerdo incumplía algunas normas relativas a la financiación de campañas electorales, pero *Citizens United* reaccionó argumentando que la Comisión Electoral Federal vulneraba la Primera Enmienda por impedirles contribuir a la forja de la opinión pública. Con un ajustado margen (cinco votos a favor y cuatro en contra) el Tribunal Supremo dio la razón a *Citizens United*. Los argumentos en contra de esta decisión se encuentran en el voto particular redactado por el Juez Stevens, al que se sumaron los jueces Ginsburg, Breyer y Sotomayor, defensores de que la innegable distinta naturaleza de las personas físicas y jurídicas permite afirmar que ninguna corporación tiene una idea propia, sino que reflejan las de las personas que las administran, y que la decisión tomada por la mayoría puede provocar que las grandes corporaciones acaben minando la libre opinión pública en lugar de mejorarla.

Vid A. VALERO HEREDIA, “*Citizens United* y la financiación de las campañas electorales en el Derecho norteamericano”, *Teoría y Realidad Constitucional*, núm. 30, 2012, pp. 433-453.

<sup>52</sup> En el caso *Burwell*, el Tribunal Supremo resolvió que el derecho a la libertad religiosa permitía a las empresas negarse a facilitar medios y tratamientos anticonceptivos, aunque la ley obligase a ello. El fundamento era que se debe proteger la actividad de iglesias y otras organizaciones religiosas sin ánimo de lucro en tanto en cuanto “existen para fomentar los intereses de las personas que se profesan una misma fe”, pero no ocurre así con las entidades con ánimo de lucro, cuyos trabajadores están unidos esencialmente por un vínculo económico. Resulta de igual interés el voto particular contra la decisión de la mayoría emitido por la juez Ruth Bader Ginsburg, donde se plantea que este reconocimiento abre la puerta a que las personas jurídicas puedan negarse al cumplimiento de cualquier

actores sociales haya, resulta “indispensable para la toma de decisiones en una democracia, lo cual no es menos cierto porque el discurso provenga de una corporación”<sup>53</sup>. En Europa, por su parte, esta titularidad ya está consolidada tanto a nivel supranacional<sup>54</sup> como nacional<sup>55</sup>.

Los motivos para tal reconocimiento radican sobre todo en que algunos de los valores inherentes a la dignidad están presentes también en las personas jurídicas. Si a ello sumamos que los derechos fundamentales cuentan con una dimensión axiológica, debemos concluir que empresas, fundaciones, asociaciones, partidos políticos o sindicatos son frecuentemente empleados por los ciudadanos para desarrollar actividades vinculadas al ejercicio de sus derechos<sup>56</sup>.

#### 4. LAS ALTERNATIVAS EN LA MODERACIÓN DE CONTENIDOS ONLINE

##### 4.1. Las deficiencias del modelo de puertos seguros

La Sección 230 y las normativas influenciadas por la misma han garantizado que los ISP y, en particular, las redes sociales, hayan sido durante mucho tiempo ajenas a los posibles abusos que cometían sus usuarios. Tal conducta podría resultar criticable, pero era legítima, pues la legislación las eximía de toda responsabilidad. Sin embargo, esta inacción terminó por provocar consecuencias de gran relevancia política y social, e inició una cadena de reacciones que aún están sucediéndose.

Los detonantes de este nuevo paradigma fueron esencialmente tres: por un lado, el mundo entero comprobó con horror cómo estas plataformas eran empleadas para retransmitir en directo auténticas tragedias, cuando no críme-

---

ley, salvo quizás leyes fiscales, siempre que puedan acreditar algún tipo de afectación a sus creencias religiosas o morales.

*Vid* Burwell Secretary of Health and Human Services v. Hobby Lobby Stores Inc., 573 682 (2014)

<sup>53</sup> Citizens United v. Federal Election Commission, 558 U.S. 310 (2010).

<sup>54</sup> Artículo 34 del Convenio Europeo de Derechos Humanos: El Tribunal podrá conocer de una demanda presentada por cualquier persona física, organización no gubernamental o grupo de particulares [...].

<sup>55</sup> En España, por todas, véase la STC 139/1995, de 26 de septiembre.

<sup>56</sup> L. M. Díez-PICAZO GIMÉNEZ, *Sistema de derechos fundamentales*, Tirant lo Blanch, 5ª ed., Valencia, 2021, pp. 129-132.

nes abyectos;<sup>57</sup> por otro lado, los avances tecnológicos han facilitado la creación de los llamados *socialbots* o *socbots*,<sup>58</sup> softwares que causaron un escándalo de dimensiones políticas a priori inimaginables; finalmente, hay lobbies y movimientos sociales de diversa índole que ejercen una presión cada vez más intensa sobre las redes sociales para que actúen sobre según qué contenidos<sup>59</sup>.

La conclusión a la que llegaron empresas como Facebook o Google fue que, si no transmitían a la sociedad el mensaje de que estaban haciendo algo

---

<sup>57</sup> Cuando las redes sociales incorporaron la opción de publicar vídeos en directo (los llamados “lives”), algunos usuarios vieron la vía para adquirir notoriedad a través de la retransmisión de hechos abominables: adolescentes suicidándose, homicidios, palizas, incluso actos terroristas, como el cometido en Christchurch (Nueva Zelanda) en 2019, en el que el autor del atentado fue retransmitiendo y comentando el tiroteo a medida que lo cometía. Es cierto que estos acontecimientos son cuantitativamente escasos, pero su impacto ha sido terrible.

Vid A. RAVEN, “Public Crime Reporting on Social Media: A Progressive or Regressive Phenomenon?”, en VV.AA., *Social Media Strategy in Policing*, Springer Nature, Londres, 2019, pp. 79-100.

<sup>58</sup> Se trata de aplicaciones que tienen la apariencia de usuarios corrientes y se registran como tales, pero están configurados para generar, apoyar o rechazar determinados términos o expresiones de conformidad con parámetros previamente establecidos.

Vid T. JACKSON y C. GUITIAN, *Fake news (What’s the issue?)*, Quarto Publishing, Londres, 2020, p. 80.

<sup>59</sup> En el año 2017, un reportaje de The Times puso de manifiesto que en YouTube existían una gran cantidad de vídeos de carácter filonazi, terrorista o violento y que, durante la difusión de los mismos, aparecían mensajes publicitarios de distintos anunciantes, como McDonalds, Mercedes-Benz, HSBC, The Guardian, la Fundación Marie Curie e incluso el Gobierno Británico. Su reacción no se hizo esperar: exigieron el control de dichos contenidos y no volver a verse relacionados en vídeos potencialmente similares o retirarían sus anuncios. Para afrontar la crisis, Google, empresa propietaria de esta red, anunció el uso de moderadores de inteligencia artificial para detectar y eliminar estos mensajes. Como resultado, entre julio y septiembre de 2018, YouTube eliminó 7,8 millones de vídeos, 1,6 millones de canales y 224 millones de comentarios considerados contrarios a sus políticas de uso.

Tres años después, en julio de 2020, las organizaciones Free Press y Common Sense Media, con el inestimable apoyo de la Liga Antidifamación (ADL) y la Asociación de Defensa de los Derechos de los Afroestadounidenses (NAACP), pidieron a las empresas que se anunciaban en Facebook que realizasen un boicot dirigido a forzar una mayor vigilancia y represión contra los grupos y mensajes que inciten al odio o la violencia. Alarmado tras la experiencia de YouTube y después de haber sufrido descensos en la cotización de sus acciones, Mark Zuckerberg anunció en su cuenta personal de Facebook nuevos mecanismos para perseguir estos mensajes en todas sus plataformas y aprobó medidas adicionales, algunas de carácter simbólico, como el bloqueo de algunas publicaciones realizadas por el entonces Presidente, Donald Trump.

Vid A. MOSTROUS, “Big brands fund terror through online adverts”, *The Times*, 2017. Disponible en: <https://www.thetimes.co.uk/article/big-brands-fund-terror-knnxfgb98> consultado: 10-06-2021.

Vid Informe trimestral de actividad de YouTube (julio-septiembre de 2018).

al respecto, sufrirían consecuencias de diversa índole: por un lado, económicas, por el riesgo de perder anunciantes o sufrir boicots; por otro, reputacionales, dado el desprestigio que estos mensajes suponen para la imagen de una marca; verían también reducirse su número de usuarios, pues muchas personas se darían de baja o no se registrarían con el fin de reducir su exposición potencial (y la de los menores a su cargo) a estos materiales; por último, pero no menos importante, existía el riesgo de una escalada de este tipo de acciones. Podemos observar que la respuesta de las redes sociales ha sido prácticamente la misma: potenciar los sistemas de moderación de contenidos, establecer sanciones a los difusores de los materiales polémicos y realizar purgas periódicas de la información que alojan, asumiendo como un coste aceptable que, en ocasiones, se eliminan mensajes legítimos<sup>60</sup>.

El primer problema, referido a la difusión en directo de vídeos de gran violencia, es muy difícil de atajar. Es un hecho que los crímenes retransmitidos en directo por redes sociales fueron creciendo en intensidad y crueldad, a la par que también creció el número de usuarios que los veían tan pronto como tenían noticia de que estaban ocurriendo. Esta versión online del “efecto Lucifer”<sup>61</sup>, aplicada sobre una masa crítica de miles de millones de usuarios, resultaba dantesca. Pero la inmediatez de la retransmisión en directo de estos terribles hechos, el carácter previo habitualmente pacífico de sus autores y su rápida viralización imposibilitan su erradicación. De ahí que la respuesta de las redes sociales se centre en conseguir la máxima diligencia en la detección, suspensión de la emisión y pronto aviso a las autoridades policiales pertinentes, medidas que se toman a través de una revisión híbrida apoyada en la inteligencia artificial y controladores humanos<sup>62</sup>.

El segundo problema, relativo al uso descontrolado de los *sobots*, radica en que estas aplicaciones simulan el comportamiento de una persona física en un entorno digital, de manera que expresan opiniones o activan respuestas auto-

---

<sup>60</sup> T. GILLESPIE, *Custodians of the Internet: Platforms, content moderation and the hidden decisions that shape social media*, Yale University Press, New Haven (Connecticut), 2018.

<sup>61</sup> “Efecto Lucifer” fue el nombre que se dio a los resultados de un experimento realizado en 1971 en la Universidad de Stanford. Se demostró que, cuando no existen reglas claras que limiten roles y conductas, quienes se sienten con el poder de causar daño entran en una espiral de violencia cada vez más compleja y cruel.

Vid P. Zimbardo, *El efecto Lucifer: El porqué de la maldad*, Paidós, Barcelona, 2007.

<sup>62</sup> S. AGUDO, “Cómo puede resolver Facebook sus problemas con los vídeos violentos”, *Genbeta*, 2017. Disponible en: <https://www.genbeta.com/a-fondo/como-puede-resolver-facebook-sus-problemas-con-los-videos-violentos> consultado: 12-09-2021

matizadas partiendo de una configuración previamente definida. Su uso masivo acaba alterando el propio fundamento de las redes sociales, pues las opiniones más respaldadas no serán ya las que cuenten con más apoyo de sujetos reales, sino aquéllas con más *bots* configurados a tal efecto. La trascendencia de esta cuestión radica en dos factores: en primer lugar, no se puede saber a ciencia cierta cuántos están actualmente activos, aunque se estima que serán cientos de millones (sólo en Twitter, el número de usuarios que en realidad son programas oscila entre el nueve y el quince por ciento de sus perfiles); en segundo lugar, no podemos saber, en términos generales, qué fines persiguen sus impulsores<sup>63</sup>.

Esta “comunicación algorítmica” es sumamente sencilla de crear, pero muy difícil de controlar. Por un lado puede ser legítima, toda vez que la usan empresas, corporaciones y partidos políticos para difundir publicidad o información sobre sus distintas actividades; pero en otros casos su legitimidad es más que cuestionable, particularmente cuando su propósito es configurar la opinión pública, modificar conductas o alterar agendas políticas<sup>64</sup>. Un buen ejemplo de esta controversia tuvo lugar en las elecciones presidenciales estadounidenses de 2016, que fueron de gran polarización y quedaron bajo la sospecha de distorsiones provocadas por *bots* en los foros online (en concreto, se ha estimado que entre el 16 de septiembre y el 21 de octubre estuvieron activos más de cuatrocientos mil programas que generaron casi cuatro millones de tuits, es decir, la quinta parte de los mensajes políticos de esos días)<sup>65</sup>.

El Comité Judicial del Senado abrió una investigación sobre estos hechos y uno de los temas tratados fue los límites de la libertad de expresión cuando ésta se ejerce mediante aplicaciones online<sup>66</sup>. Una de las consecuencias de esta investigación ha sido la recientemente aprobada Ley BOT del Estado de

---

<sup>63</sup> O. VAROL, E. FERRARA et alios, “Online human-bot interactions: Detection, estimation and Characterization”, *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*, 2017.

<sup>64</sup> E. CAMPOS-DOMÍNGUEZ y B. GARCÍA-OROSA, “Comunicación algorítmica en los partidos políticos: automatización de producción y circulación de mensajes”, *El profesional de la información*, vol. 27, núm. 4, 2018, pp. 769-777.

<sup>65</sup> A. BESSI y E. FERRARA, “Social bots distort the 2016 U.S. Presidential election online discussion”, *First Monday*, vol. 21, núm. 11, 2016, pp. 1-14.

<sup>66</sup> Mark Zuckerberg, CEO de Facebook, Jack Dorsey, CEO de Twitter, y Sandar Pichai, CEO de Google, tuvieron que comparecer al respecto ante una comisión del Senado. Los tres reconocieron que sus servidores y aplicaciones no estaban preparados para afrontar una campaña de desinformación cuidadosamente planificada y aseguraron que tomarían medidas humanas y tecnológicas para evitar nuevas adulteraciones. “No estoy orgulloso de cómo el libre mercado de las ideas fue manipulado para falsear la realidad y dividir nuestra nación”, dijo

California, que exige su identificación como tales y, en caso de no haberse identificado, confirmar que lo son o no en un plazo máximo de 72 horas tras ser requeridos para ello, so pena de ser eliminados de la Red. Sin embargo, esta normativa es cuestionable desde un punto de vista constitucional. El debate se centra en si el libre mercado de las ideas se distorsiona por este uso de la inteligencia artificial y si su restricción vulnera los derechos que la Primera Enmienda reconoce a las personas físicas que los hayan configurado, pues se les podría equiparar a los editores de periódicos que durante décadas orientaron y configuraron la opinión pública<sup>67</sup>. En todo caso, no es descartable que en algunos años encontremos la resolución definitiva por parte del Tribunal Supremo, así como nuevas legislaciones a tal efecto<sup>68</sup>.

El tercer problema quizás sea el más complejo a la hora de analizar las dificultades que tendría regular la libertad de expresión en redes sociales: ¿qué hacer con el discurso del odio y los mensajes que alientan a la violencia y cuál es su frontera con el pensamiento polémico, hiriente o discrepante, que es legítimo por más que sea incómodo? Las redes sociales han elaborado políticas y normas de obligado cumplimiento para sus usuarios y, bajo la premisa “en Facebook manda Facebook”, ejercen su derecho a elegir qué contenidos son conformes con los valores que –al menos públicamente– inspiran su labor. Como consecuencia, se arrojan poderes moderadores prácticamente absolutos<sup>69</sup>.

---

Dorsey. Zuckerberg dio un paso más allá y afirmó: “No teníamos una visión suficientemente amplia de cuál es nuestra responsabilidad y eso fue un gran error. Mi error. Lo siento”.

Vid “Facebook, Social Media privacy and the use and abuse of data”, *Joint Hearing before the Committee on Commerce, Science and Transportation and the Committee on the Judiciary United States Senate*, April 10, 2018.

Vid “Testimony of Jack Dorsey, Chief Executive Officer Twitter, Inc.”, *United State Senate Select Committee on Intelligence*, September 5, 2018.

<sup>67</sup> L. BARDAJÍ GÁLVEZ, “¿Should robots have a right to free speech?” *Do Better by Esade*, 2019. Disponible en: <https://dobetter.esade.edu/en/robots-right-free-speech> consultado: 14-09-2021.

<sup>68</sup> No en vano, las influencias externas y la distorsión de la opinión pública online han sido frecuentemente mencionadas en acontecimientos políticos de lo más variopinto. Se ha hablado de ellos durante el proceso independentista catalán, durante el referéndum del Brexit, en las elecciones francesas o en las elecciones alemanas. Determinar cuánto hay de realidad, de tentativa o de ficción es una labor compleja que excede en mucho el objeto de este trabajo.

Vid D. A. MARTÍN, J. N. SHAPIRO y M. NEDASHKOVSKAYA, “Recent Trends in Online Foreign Influence Efforts”, *Journal of Information Warfare*, vol. 18, núm. 3, special edition Winter 2019, pp. 15-48.

<sup>69</sup> Así, en las políticas de uso de Facebook e Instagram encontramos que “si nuestros productos se usan de forma inapropiada o si alguien muestra una conducta perjudicial para los demás [...] aplicaremos las medidas correspondientes, tales como ofrecer ayuda, eliminar

Precisamente han sido Facebook y Twitter quienes han tomado los caminos más complejos y quizá arriesgados. En octubre de 2020 comenzó a funcionar el *Facebook Oversight Board*, que pronto fue conocido como “Tribunal Supremo de Facebook”, una especie de consejo de sabios integrado por veinte miembros elegidos entre juristas, periodistas, políticos, activistas y académicos. Se trata de una entidad independiente responsable de velar por el legítimo ejercicio de la libertad de expresión, de suerte que Facebook se compromete a ejecutar sus resoluciones. Hasta tal punto es así que el Consejo depende de un fideicomiso, cuenta con unos estatutos propios y establece una operativa específica que aspira a la imparcialidad en la toma de decisiones, con un actuar ajeno a presiones externas de cualquier índole<sup>70</sup>. Twitter ha optado por un modelo incluso más polémico: la calificación de sus contenidos. Sus moderadores se reservan el derecho a colocar una etiqueta que, a modo de advertencia, avisa a los lectores de que están ante un comentario considerado “falso o dañino”, sea quien sea el autor del

---

el contenido, bloquear el acceso a ciertas funcionalidades, inhabilitar una cuenta o ponernos en contacto con las autoridades pertinentes”. Por su parte, el blog corporativo de YouTube recoge que “damos otro paso en nuestra política contra la incitación al odio al prohibir específicamente los videos que alegan que un colectivo es superior para justificar la discriminación, la segregación o la exclusión en función de cualidades como la edad, el género, la raza, la casta, la religión, la orientación sexual o la condición de veterano. Esto incluiría, por ejemplo, videos que promueven o glorifican la ideología nazi, que es intrínsecamente discriminatoria. Finalmente, eliminaremos el contenido que niega que ocurrieron eventos violentos bien documentados, como el Holocausto o el tiroteo en la escuela Sandy Hook”. TikTok, la más reciente de las redes sociales usadas masivamente a nivel internacional, indica que se reservan el derecho “a suspender temporal o permanentemente o a rescindir su cuenta de usuario o a restringir o imponer límites a su acceso a alguna parte o a la totalidad de los servicios, con o sin notificación previa, en cualquier momento, por cualquier motivo o sin motivo”. Twitter y Periscope prevén dos tipos de acciones contra quienes contravengan sus condiciones de uso: por un lado, eliminar contenidos y bloquear cuentas; en su defecto, etiquetarlos como información potencialmente falsa: “Podremos suspender o cesar a su cuenta, o bien dejar de suministrarle todos o parte de los Servicios, en cualquier momento y por cualquier motivo, o sin motivo. [...] Está prohibido promocionar, con la intención de engañar, contenido multimedia falso o alterado que pueda dar lugar a daños graves. Asimismo, es posible que etiquetemos los Tweets que incluyen contenido multimedia falso y alterado para ayudar a los usuarios a comprender su autenticidad y para ofrecer más contexto”.

Vid <https://www.facebook.com/legal/terms> consultado 01-09-2021

Vid <https://blog.YouTube/news-and-events/our-ongoing-work-to-tackle-hate> consultado 01-09-2021

Vid <https://www.tiktok.com/legal/terms-of-use?lang=es> consultado 01-09-2021

Vid <https://help.twitter.com/es/rules-and-policies/manipulated-media> consultado 01-09-2021

<sup>70</sup> K. KLONICK, “The Facebook Oversight Board: Creating an independent institution to adjudicate online free expression”, *The Yale Law Journal*, núm. 129, 2020, pp. 2418-2499.



mismo. Es más, su nuevo propietario, Elon Musk, está buscando fórmulas para desarrollar mecanismos de calificación cada vez más eficientes y ha llegado a plantear la posibilidad de que sean los usuarios quienes valoren las publicaciones.

Sea como fuere, la mayor parte de las redes sociales ejercen esta facultad de forma opaca y automatizada. De ahí que no falten voces que reclaman, más que órganos de control y revisión, una especie de “constitución” que consagre un compromiso más firme de cada empresa con los usuarios, la máxima transparencia en el funcionamiento de los procesos de revisión y su adecuación a los estándares internacionales en materia de actuación contra mensajes discriminatorios<sup>71</sup>. Porque, por mucho cuidado que se pueda poner en esta labor, es innegable el riesgo de destrucción de opiniones y expresiones de interés para el debate público<sup>72</sup>.

#### 4.2. Una propuesta innovadora: vigilar al vigilante

Una de las características de la Administración Trump fue la frecuencia con la que este Presidente estadounidense opinaba sobre asuntos de lo más diverso a través de su cuenta personal de Twitter. Sus declaraciones, muchas veces polémicas, se convertían inmediatamente en tendencia y generaban

---

<sup>71</sup> K. KLONICK y T. E. KADRI, “How to make Facebook’s ‘Supreme Court’ Work”, *Digitalcommons*, School of Law, University of Georgia, 2018.

<sup>72</sup> Uno de los mejores exponentes de esta situación se vivió el 5 de junio de 2019, fecha en que YouTube anunció la eliminación de la película “El Triunfo de la Voluntad” (1934), de la directora alemana Leni Riefenstahl, por ser de contenido “intrínsecamente discriminatorio”. No cabe duda de que esta obra tenía un fin propagandístico del nazismo, pero no es menos cierto que supuso un hito en la historia del cine y que tiene un innegable valor histórico. Desde cualquier punto de vista lógico, cuesta aceptar que eliminar estas obras ayude a combatir las nuevas formas de discriminación.

Una polémica similar se produjo cuando la cadena HBO anunció la retirada “Lo que el viento se llevó” (1939) por idealizar un contexto histórico en el que la esclavitud resultaba aceptable y por representar a los esclavos negros o como leales sirvientes devotos de sus amos blancos o como ineptos. Ante el aluvión de críticas, HBO recuperó la película, pero añadió una contextualización de la misma para concienciar sobre los horrores de la esclavitud.

Una decisión parecida ha tomado Disney+, que ha agregado una etiqueta de advertencia por contenido inadecuado, discriminatorio o xenófobo a algunos de sus clásicos más conocidos, como *Dumbo* (1941), *Peter Pan* (1953), *La dama y el vagabundo* (1955), *Los robinsones de los mares del sur* (1960), *El libro de la selva* (1967) o *Los aristogatos* (1970). Una decisión legítima, pues *The Walt Disney Company* es propietaria de estas películas, pero que, a mi modo de ver, resulta incomprensible y, además, supone una banalización del discurso del odio.

todo tipo de comentarios, tanto favorables como desfavorables. Debido a ello, Trump bloqueó a algunos usuarios particularmente críticos, lo que provocó una demanda judicial por parte de quienes alegaban que el Presidente había hecho de su cuenta personal un espacio público de discusión y que, al bloquearles, vulneraba la Primera Enmienda. Así lo entendieron los tribunales,<sup>73</sup> que defendieron que el carácter privado de las redes sociales impedía considerar la moderación de contenidos o el bloqueo de usuarios como censura pública... salvo en las cuentas de dimensión institucional. Es decir, en este caso cabía aplicar la doctrina de la State Action, pues aunque fuese la cuenta personal de Trump, éste la usaba para anunciar y discutir asuntos públicos, por lo que no tenía derecho a negar a los ciudadanos su participación como seguidores y comentaristas<sup>74</sup>.

Pero el choque entre la Administración Trump y las redes sociales alcanzó su cénit con la aprobación de la *Orden Ejecutiva de 28 de mayo de 2020, de prevención de la censura online*<sup>75</sup>. En ella, el Presidente expresaba su preocupación ante el hecho de que Twitter, Facebook, Instagram y YouTube dispongan de “un poder inmenso, sin precedentes, que usan para censurar, restringir, editar, dar forma, ocultar o alterar” información clave para la opinión pública, e instaba a la *Federal Communications Commission* a revisar la aplicación de la Sección 230 y a proponer una posible regulación que castigue a los actores privados en caso de vulnerarla. De forma sibilina, pero no por ello carente de lógica, el Presidente planteaba que, si las redes sociales son la nueva plaza pública, resultaría incongruente que un vigilante privado dé o niegue el acceso a la misma en función de sus propios criterios.

Aunque la decisión de Donald Trump pudiera haber obedecido a una reacción colérica por las restricciones a su cuenta de Twitter, las sospechas sobre estas compañías no han parado de crecer, sobre todo por la opacidad

---

<sup>73</sup> Knight First Amendment Institute v. Trump, Southern District of New York, May 23<sup>th</sup> 2018

<sup>74</sup> V. J. VÁZQUEZ ALONSO, “Twitter no es un foro público pero el perfil de Trump sí lo es. Sobre la censura privada de y en las plataformas digitales en los EE.UU.”, *Estudios de Deusto*, vol. 68/1, pp. 475-508.

<sup>75</sup> Cuando Twitter calificó como “falso o dañino” un tuit del Presidente Trump relativo a un uso fraudulento del voto por correo, fue como abrir la caja de Pandora. En la mencionada orden ejecutiva, el Presidente denunciaba la hipocresía y falsa neutralidad de los administradores de las redes sociales y, en particular, acusaba a Twitter de realizar activismo político, tomar decisiones editoriales y perjudicar abiertamente las ideologías conservadoras.

Vid “Executive Order 13925 of May 28, 2020 Preventing online censorship”, *Presidential Documents, Federal Register*, vol. 85, núm. 106, Tuesday, June 2, 2020.

de sus procesos de toma de decisiones<sup>76</sup>. El denominado “Facebookgate” ha sacado a la luz prácticas más que cuestionables, en particular relacionadas con la gestión del *big data*, para manipular, polarizar y dividir a la opinión pública, para obtener rendimientos económicos gracias al tratamiento de datos personales e incluso para influir en los hábitos de consumo y de salud de los usuarios. Un comportamiento empresarial denominado “capitalismo de vigilancia”, pues es ya un hecho que grandes corporaciones como Google, Amazon, Apple, Microsoft y, desde luego, Meta, analizan sistemáticamente nuestras conductas online y las convierten en datos, de suerte que puede pensarse que hemos dejado de ser sus clientes y nos hemos convertido en objetos de extracción de una materia prima única: la información<sup>77</sup>. Un proceder preocupante que se ha calificado como el exponente de la “bancarrota moral” de estas empresas y está llevando a que cada vez más voces soliciten una intervención del legislador para poner coto a su omnímodo poder<sup>78</sup>.

A todo ello hay que sumar el proceso judicial que Trump ha emprendido contra Twitter después del cierre de su cuenta, tras ser acusado de causar el inconcebible asalto al Capitolio del seis de enero de 2021<sup>79</sup>. Resulta difícil prever cómo acabará este proceso, pues si bien la Primera Enmienda protege de posibles abusos cometidos por los poderes públicos (interpretación que ya es de por sí amplia, pues la literalidad de la norma dice “del Congreso”), no cabría solicitar dicha protección cuando no ha habido intervención pública alguna en la restricción ni existen leyes que limiten la moderación privada de contenidos<sup>80</sup>. Sin embargo, no podemos olvidar que sobre esta misma cuenta ya se ha sentenciado su carácter de foro público. ¿No sería entonces incongruente que

---

<sup>76</sup> A. JUDGE, “30 Questions that Facebook has yet to Answer -- Gaps in the testimony of Mark Zuckerberg at a US Senate hearing?” *Laetus in praesens*, 2018. Disponible en: <https://www.laetusinpraesens.org/docs10s/facebook.php> consultado: 10-10-2021.

<sup>77</sup> S. ZUBOFF, *La era del capitalismo de vigilancia*, Paidós, Barcelona, 2020, pp. 24 y 25.

<sup>78</sup> G. LEVY, “La bancarrota moral de Facebook”, *AndinaLink*, 2021. Disponible en: <https://andinalink.com/la-bancarrota-moral-de-facebook/> consultado: 19-10-2021.

<sup>79</sup> Ante la Corte Federal de Florida, Trump ha denunciado que el debate democrático está en peligro y que, si realmente su cuenta era un foro público, Twitter viola la Primera Enmienda al cerrar su cuenta pues afecta a sus más de ochenta millones de seguidores. Argumenta, además, que la decisión tomada por la compañía que dirige Jack Dorsey es arbitraria, toda vez que mantiene abiertas cuentas claramente violentas y antidemocráticas, como por ejemplo las de líderes talibanes, con la misma impunidad con que cerró la suya.

<sup>80</sup> J. MORRIS, citado por: A. PHILLIPS. “No, Twitter is not violating Trump’s freedom of speech”, *The Washington Post*, 2020. Disponible en: <https://www.washingtonpost.com/politics/2020/05/29/no-twitter-did-not-violate-trumps-freedom-speech/> consultado 03-09-2021.

su titular –es decir, Trump– no pudiese bloquear usuarios, pero que la red que le da soporte –o sea, Twitter– sí pudiese cerrar su perfil? Habrá que estar atentos a la evolución de este proceso, pues, aunque el 20 de noviembre de 2022 Twitter devolvió a Trump el acceso a su cuenta, éste no parece dispuesto a retirar su demanda, sino todo lo contrario: la ha ampliado contra Meta y Google, exigiendo además una indemnización por daños y perjuicios.

### 4.3. La propuesta europea: las exigencias legales a los ISP

En el seno de la Unión Europea, las obligaciones de las redes sociales respecto a su posible contenido ilícito o difamatorio han acabado ante el Tribunal de Luxemburgo<sup>81</sup>. El debate se centraba en determinar hasta qué punto eran responsables de detener la difusión de comentarios conflictivos, habida cuenta de que los mensajes pueden compartirse de forma inmediata y su viralización sigue una progresión geométrica. El TJUE dio un primer paso al entender legítimo que un tribunal competente exija a los administradores de las redes el bloqueo de dicho contenido en su origen, así como su rastreo y la posterior eliminación de todo el que sea idéntico. Por otro lado, aunque en el caso Delfi el Tribunal Europeo de Derechos Humanos sentenció que los deberes de un editor no pueden extrapolarse de forma general a las redes sociales, validó el uso de los distintos filtros que permiten la moderación de contenidos<sup>82</sup>.

La decisión sobre qué barrera aplicar no es menor, pues aunque las multinacionales que operan las redes sociales puedan gestionar libremente lo que no es sino su propiedad, sus acciones podrían acabar convirtiéndose en un séptimo sello<sup>83</sup> que silencie las opiniones tenidas por indeseables. A nivel institucional ya comprobamos que la Unión Europea, el Consejo de Europa y los Estados miembros han adoptado estrictas medidas contra el discurso del odio y la difamación, hasta el punto que en los últimos años han cobrado forma los primeros proyectos normativos que imponen obligaciones específicas a las redes sociales y otros portales análogos. El primer paso en esta senda fue la publicación de una serie de “medidas operativas” a cumplir por los

---

<sup>81</sup> STJUE de 3 de octubre de 2019, *Eva Glawischnig-Piesczek v. Facebook Ireland Limited*, C-18/18.

<sup>82</sup> E. ARROYO AMAYUELAS, “La responsabilidad de los intermediarios en Internet. ¿Puertos seguros a prueba de futuro?” *Cuadernos de derecho transaccional*, vol. 12, núm. 1, 2020, pp. 808-837.

<sup>83</sup> “Cuando se abrió el séptimo sello, se hizo el silencio en el cielo” (Apocalipsis, 8:1).

ISP. Aunque no tenían fuerza vinculante, se les solicitaba la incorporación de sistemas de detección, notificación y borrado de contenidos ilícitos u ofensivos, así como acciones concretas para la detección del terrorismo online<sup>84</sup>. El siguiente paso fue la aprobación del *Código de Conducta sobre el discurso del odio online*, cuya primera evaluación insiste en la necesidad de que los ISP cuenten normas que prohíban esta conducta y habiliten mecanismos para revisar los contenidos que puedan tenerse como tales<sup>85</sup>.

A nivel nacional, Alemania ha sido pionera en este campo. El 1 de octubre de 2017 el Bundestag aprobó la *Ley de Aplicación en la Red (Netzwerkdurchsetzungsgesetz – NetzDG)*, expresamente diseñada para combatir los delitos de odio, la desinformación y los contenidos delictivos de las redes sociales. Las obligaciones impuestas a las mismas son claras: i) contar con procedimientos de denuncia sencillos, de acceso inmediato y constantemente disponibles para los usuarios; ii) presentar cada seis meses un informe sobre el volumen de denuncias y los recursos técnicos y humanos habilitados para su tratamiento; iii) en el caso de los delitos de odio, se exige a la propia red social que presente la correspondiente denuncia ante la Oficina Federal de Policía Criminal; iv) establecer un mecanismo de apelación cuando un contenido sea borrado o una denuncia desestimada; v) resolver las denuncias en un plazo de veinticuatro horas. Incumplir estas normas puede acarrear sanciones de hasta cincuenta millones de euros<sup>86</sup>.

Por su parte, el Parlamento francés aprobó la *Loi du 24 juin 2020 visant à lutter contre les contenus haineux sur internet*, análoga a la norma alemana. Su nacimiento fue complejo, pues el proyecto original, sumamente ambicioso, exigía la eliminación de todo discurso del odio y de carácter pornográfico de las redes sociales en un plazo máximo de veinticuatro horas tras notificarse a los administradores, so pena de multas que alcanzarían hasta el cuatro por ciento de su facturación mundial. Sin embargo, el Consejo Constitucional francés entendió que doce de sus diecinueve disposiciones resultaban inconstitucionales por vulnerar el principio de proporcionalidad, restringir en

---

<sup>84</sup> Comisión Europea, “A Europe that protects: Commission reinforces EU response to illegal content online”, 2018, Disponible en: [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_18\\_1169](https://ec.europa.eu/commission/presscorner/detail/en/IP_18_1169) consultado: 14-09-2021.

<sup>85</sup> Comisión Europea, “Assessment of the Code of Conduct on Hate Speech on line. State of Play”, 2019.

<sup>86</sup> Ministerio Alemán de Justicia y Protección al Consumidor. Disponible en: [https://www.bmjjv.de/DE/Themen/FokusThemen/NetzDG/NetzDG\\_node.html](https://www.bmjjv.de/DE/Themen/FokusThemen/NetzDG/NetzDG_node.html) consultado: 14-09-2021.

demasía la libertad de expresión e imponer sanciones abusivas<sup>87</sup>, de manera que el texto finalmente aprobado ha quedado mermado y apenas contiene mandatos expresos.

En ciernes está el *Reglamento Europeo relativo a un mercado único de servicios digitales*, norma que aspira a regular todos los servicios de intermediación en línea. En su texto se recoge que las denominadas “plataformas online de gran tamaño”, es decir, aquellas que prestan servicio al menos al diez por ciento de los ciudadanos europeos, representan un riesgo especial cuando difunden contenidos ilícitos. Por consiguiente, tanto si están establecidas en el seno de la Unión Europea como si no, deben contar con un mecanismo de denuncia por parte de los usuarios y con alertas fiables que permitan su detección de oficio<sup>88</sup>, so pena de ver suspendido el acceso a su url desde el territorio comunitario.

Como puede comprobarse, se está exigiendo a las redes sociales que lleguen a su máximo alcance el procedimiento de “notice and takedown”, pero de una forma mucho más generalista y no sólo para proteger los derechos de autor. Esta política, tan alejada de la doctrina de los puertos seguros, no carece de defensores; al fin y al cabo, el Derecho debe brindar protección jurídico-constitucional a quienes emplean las redes sociales. Además, si otros medios de difusión afrontan distintos escenarios de responsabilidad, ya sea administrativa, civil o aun penal, no parece sostenible la exención de los ISP’s<sup>89</sup>.

Es necesario añadir que, si bien la regulación estadounidense y la europea parecen distanciarse, no se están produciendo problemas irresolubles. De facto, las medidas que se han tomado por la presión mediática y social que se vive en Estados Unidos superan los estándares mínimos de actuación que exige la Unión Europea en lo referido a libertad de expresión, aunque no ocurra lo mismo en la protección y tratamiento de datos personales. Puede generar desconfianza que sean corporaciones privadas las que gestionen el ejercicio de esta libertad, pero tal suele ser el modo americano de hacer las cosas, ya que la

---

<sup>87</sup> Consejo Constitucional Francés. Décision n° 2020-801 DC du 18 juin 2020. Loi visant à lutter contre les contenus haineux sur internet.

<sup>88</sup> Comisión Europea, “Ley de Servicios Digitales: para un entorno online seguro y responsable”. Disponible en: [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_es](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_es) consultado: 21-09-2021.

<sup>89</sup> L. COTINO HUESO, “Libertades informativas y responsabilidad de los prestadores de servicios de Internet” en VV.AA., *La protección jurisdiccional de los derechos. Actas del XI Congreso de Constitucionalistas de España*, 2013, Tirant lo Blanch, Valencia, pp. 171-223.

libertad se concibe muchas veces como la menor intervención estatal posible. Con todo, la Administración Biden ha mostrado su preocupación por la falta de regulación expresa de las redes sociales. En su agenda está una batería de medidas que convergerían hacia el modelo europeo, pero el riesgo de que acaben siendo declaradas contrarias a la Primera Enmienda complica su tramitación. Sin embargo, no conviene olvidar que el libre mercado de las ideas no fue concebido en un escenario donde la información se viralizase a nivel mundial de forma instantánea, ni donde aplicaciones virtuales pudieran alterar el sano debate público, por lo que exigir legalmente estándares mínimos de calidad y respeto no parece, al menos a priori, una idea tan descabellada<sup>90</sup>.

## 5. CONCLUSIONES

La legislación que durante las últimas décadas ha regulado Internet se diseñó con un propósito concreto: dinamizar al máximo esta magnífica forma de comunicación y atraer al mayor número posible de usuarios. En esta nueva década debemos preguntarnos, no obstante, si las renunciadas que se hicieron en pos de ese objetivo siguen siendo necesarias, toda vez que resulta inimaginable un mundo en el que las nuevas tecnologías no se empleen de forma masiva.

Lo cierto es que, a día de hoy, la principal amenaza a la autonomía de la voluntad y al libre ejercicio de buena parte de nuestros derechos civiles ya no procede de los poderes públicos, sino de grandes actores privados. Estos gigantes, tanto por su acción como por su inacción, generan consecuencias que acaban teniendo un gran impacto sobre la sociedad: con su acción, cuando restringen en demasía la libertad de expresión y emplean mecanismos de big data para condicionar a la opinión pública; por su inacción, cuando permiten campar sin freno contenidos ilícitos y preocupantes. La reacción de los poderes públicos ante esta disyuntiva es compleja: si no se interviene, el ciberespacio quedará regulado por grandes corporaciones sometidas a la presión de lobbies y anunciantes; si se interviene, se limita un espacio nacido para ser libre.

Una conclusión parece cierta: confiar en el buen hacer de los administradores de las redes sociales es arriesgado, pues hasta el momento parecen actuar con opacidad, arbitrariedad y de forma desproporcionada, emplean-

---

<sup>90</sup> M. MINOW, *Saving the News: Why the Constitution Calls for Government Action to Preserve Freedom of Speech*, Oxford University Press, Nueva York, 2021, pp. 58-101.

do técnicas de moderación de contenidos que bien pueden calificarse como censoras. De ahí que podamos afirmar que, si bien la legislación primigenia fue una especie de “creced y multiplicaos”, las políticas de uso y las restricciones al mensaje discrepante comienzan a convertirse en un séptimo sello. Uno particularmente peligroso, pues se presentan como necesarias en pos de un fin legítimo, como es combatir el odio, la violencia y la discriminación.

A tal efecto, la normativa europea parece estar dando los primeros pasos, que se centran en exigir a las redes sociales que habiliten mecanismos de notificación de actividades ilícitas, así como el deber de retirar estos contenidos tras la recepción de una denuncia. Pero ése es sólo uno de los polos de la cuestión: queda pendiente la más compleja de las tareas, que es poner coto a las acciones de moderación de contenidos para que no se conviertan en un mecanismo de anulación del pensamiento discrepante. No hay hasta el momento regulación o mecanismo de reclamación que permita a los usuarios reivindicar la legitimidad de sus mensajes más que los que habilite cada red social, lo que las convierte, de facto, en auténticos árbitros de la libertad de expresión online.

No es un problema menor: si siempre ha habido polémica en torno a los actores que configuraban la opinión pública, mucho más habrá de haberla ahora que ésta se centra en operadores que actúan a nivel internacional y que cuentan con unas capacidades de intervención sin precedentes. Afrontamos, a mi modo de ver, uno de los grandes debates de nuestro tiempo, pues la libre opinión es la esencia de todo sistema democrático y se encuentra en una encrucijada: jamás fue tan fácil hacer llegar nuestras expresiones a tantas personas, pero, a la vez, jamás el acceso a la plaza pública estuvo en manos de tan pocos actores, ni nunca tuvieron éstos tal poder y envergadura.

FRANCISCO VALIENTE MARTÍNEZ  
*Derecho Constitucional*  
Universidad Pontificia Comillas  
c/Alberto Aguilera, 23  
Madrid 28015  
e-mail: [fvaliente@comillas.edu](mailto:fvaliente@comillas.edu)