



---

**Universidad de Valladolid**

**Facultad de Derecho**

**Trabajo Fin de Grado**

**Grado en Derecho**

**Las políticas de moderación  
de contenidos de los  
proveedores de servicios de  
internet**

Presentado por:

***Lucía Pérez de la Fuente***

Tutelado por:

***Juan M<sup>a</sup> Bilbao Ubillos***

*Valladolid, 8 de mayo de 2025*

## RESUMEN

Internet ha transformado el modelo tradicional de comunicación, permitiendo a los ciudadanos pasar de ser meros receptores de información a convertirse en creadores y difusores de contenido a través de redes sociales y foros. Este fenómeno ha supuesto grandes ventajas, pero también plantea desafíos jurídicos.

En este escenario, que posibilita el intercambio de información masivo, los intermediarios digitales han implementado sistemas de moderación de contenido que plantean interrogantes sobre su transparencia y compatibilidad con el ejercicio de la libertad de expresión. La regulación de estas plataformas varía según las jurisdicciones, con diferencias notables entre el enfoque de Estados Unidos y el de la Unión Europea respecto a la responsabilidad de los intermediarios.

Este Trabajo de Fin de Grado analiza estos desafíos, abordando el equilibrio en la protección de los derechos fundamentales.

**Palabras clave:** responsabilidad de los intermediarios digitales, moderación de contenido, derechos fundamentales, libertad de expresión, censura, autorregulación de plataformas.

## **ABSTRACT**

The Internet has transformed the traditional model of communication, allowing citizens to move from being mere receivers of information to becoming creators and disseminators of content through social networks and forums. This phenomenon has brought great advantages, but it also poses legal challenges.

In this scenario, which enables the massive exchange of information, digital intermediaries have implemented content moderation systems that raise questions about their transparency and compatibility with the exercise of freedom of expression. Regulation of these platforms varies across jurisdictions, with notable differences between the US and EU approaches to intermediary liability.

This Final Degree Project analyzes these challenges, addressing the balance in the protection of fundamental rights.

**Key words:** digital intermediary liability, content moderation, fundamental rights, freedom of expression, censorship, platform self-regulation.

## Índice de contenidos

|      |                                                                                                             |    |
|------|-------------------------------------------------------------------------------------------------------------|----|
| 1.   | INTRODUCCIÓN .....                                                                                          | 5  |
| 2.   | LA RESPONSABILIDAD DE LOS PROVEEDORES DE SERVICIOS DE INTERNET<br>6                                         |    |
| 2.1. | Concepto de “conocimiento efectivo” .....                                                                   | 9  |
| 2.2. | Mecanismos de defensa de los afectados.....                                                                 | 11 |
| 2.3. | Prohibición de la obligación de supervisión .....                                                           | 11 |
| 2.4. | Evolución de la doctrina .....                                                                              | 12 |
| 3.   | COMPARATIVA ENTRE LA REGULACIÓN DE LA UE Y EEUU .....                                                       | 15 |
| 3.1. | El caso NetChoice, LLC v. Paxton .....                                                                      | 18 |
| 4.   | LAS POLÍTICAS DE MODERACIÓN DE CONTENIDOS EN INTERNET .....                                                 | 23 |
| 4.1. | La legitimidad de las políticas de moderación de contenidos en su doble<br>dimensión pública y privada..... | 23 |
| 4.2. | Las obligaciones que impone la Digital Service Act a los intermediarios....                                 | 24 |
| 4.3. | El caso Twitter vs Vox .....                                                                                | 27 |
| 4.4. | Los riesgos de la co-regulación.....                                                                        | 29 |
| 5.   | IMPACTO EN LOS DERECHOS FUNDAMENTALES .....                                                                 | 31 |
| 5.1. | El derecho a la libertad de expresión .....                                                                 | 31 |
| 5.2. | El derecho al honor, a la intimidad y a la propia imagen .....                                              | 36 |
| 5.3. | Colisión de derechos fundamentales en internet .....                                                        | 39 |
| 5.4. | La eficacia horizontal de los derechos fundamentales .....                                                  | 44 |
| 6.   | CONCLUSIONES.....                                                                                           | 48 |
| 7.   | BIBLIOGRAFÍA .....                                                                                          | 51 |

## 1. INTRODUCCIÓN

El uso de internet ha facilitado la comunicación y difusión de ideas. Los ciudadanos han pasado de tener únicamente un papel pasivo de recepción de contenido informativo, emitido por los medios de comunicación tradicionales, como prensa, radio o televisión; a tener un papel activo como creadores de contenido en redes sociales y foros en internet, que posibilitan que su voz tenga un alcance universal. Esto es, sin duda, una ventaja para potenciar la difusión de ideas de todos los ciudadanos, pero también supone un desafío jurídico garantizar la libertad de expresión, al mismo tiempo que se salvaguardan otros derechos constitucionalmente protegidos.

Por otra parte, los intermediarios han pasado de ser medios de comunicación que elaboran y editan la información que publican a ser, en ocasiones, meros proveedores de acceso para los usuarios. Surge entonces la pregunta de qué responsabilidad se les debe atribuir y qué papel tienen en estas nuevas relaciones. Los proveedores de servicios de internet realizan funciones de intermediación que se reivindican “neutras” desde el punto de vista ideológico. Sin embargo, establecen unos sistemas de moderación de contenidos que deberían ser transparentes y respetuosos con el ejercicio de la libertad de expresión. De no ser así, existen una serie de riesgos a los que nos expondríamos, tanto como sociedad, como en el plano individual.

Otra cuestión que se expone en el trabajo es que la regulación del espacio digital no es homogénea en todo el mundo. Existen diferencias entre la postura adoptada por Estados Unidos respecto a la responsabilidad de los intermediarios y la adoptada por la Unión Europea.

El ejercicio del derecho fundamental a la libertad de expresión tiene unos límites en el escenario analógico. Estos límites están relacionados con la protección de los derechos fundamentales de otros sujetos, entre ellos, el derecho al honor, la intimidad y la propia imagen. El traspaso de estos límites al plano de internet supone ciertos retos

para el legislador. Algunos de estos retos pasan por determinar qué margen tienen las plataformas digitales para moderar los contenidos que circulan por ellas, qué valor tiene el contrato privado entre una plataforma y un usuario de esta, qué responsabilidades de transparencia deben exigírseles a las plataformas, y qué responsabilidad tienen sobre los contenidos ilícitos que circulan por ellas, entre otras cuestiones que se analizarán en este trabajo.

## **2. LA RESPONSABILIDAD DE LOS PROVEEDORES DE SERVICIOS DE INTERNET**

La primera regulación sobre la responsabilidad de los intermediarios surgió en Estados Unidos a través de la sección 230 de la Ley de Decencia de las Comunicaciones (Communications Decency Act), adoptada en 1996. Esta ley prevé una cláusula de exención de responsabilidad de las plataformas respecto de dos circunstancias: respecto de los contenidos proporcionados por terceros (de forma que no pueda considerárseles autores o editores) y respecto de sus decisiones de moderación de contenidos tomadas de buena fe y orientadas a impedir o restringir el acceso a información que se considere “obscena, lasciva, sucia, contenido excesivamente violento, acosador o de cualquier otro modo objetable, esté o no constitucionalmente protegido”.

De este modo, se trata de evitar el llamado «dilema del moderador»: ¿Para qué va una plataforma a asumir la carga de evitar cualquier contenido indeseable si ese ejercicio la lleva precisamente a afrontar responsabilidades legales? <sup>1</sup> De esta forma, se estableció la inmunidad para los proveedores de servicios, aunque ejercieran un cierto “control editorial” sobre discursos objetables<sup>2</sup>. La lógica de no considerar a las plataformas digitales como editoras, a diferencia de los medios de comunicación

---

<sup>1</sup> BARATA MIR, J.: “Libertad de expresión, regulación y moderación privada de contenidos”, Teoría y Derecho, n. 32, p.96.

<sup>2</sup> SÁENZ ROYO, E.: “Libertad de expresión, odio y redes sociales: las opciones del legislador democrático español.” p. 4

tradicionales, se basa en considerar que su función es neutral y no participan en el proceso de creación de contenido.

La sección 230, cuyo objetivo era posibilitar el desarrollo del potencial de internet, es conocida por algunos autores como las 26 palabras que crearon Internet.

Actualmente, según recientes pronunciamientos jurisprudenciales, esa inmunidad se mantiene, aun cuando conlleve la proliferación de discursos de odio mediante los sistemas algorítmicos de recomendación. Las grandes plataformas justifican el uso de estos sistemas algorítmicos con la labor de filtrado automático que sería prácticamente imposible llevar a cabo con recursos humanos. No obstante, además de esa función también alteran la difusión de los contenidos favoreciendo algunos en perjuicio de otros.

Por otra parte, la Unión Europea adoptó la Directiva 2000/31/CE<sup>3</sup> (DCE) que fue incorporada al ordenamiento jurídico español mediante la Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y de comercio electrónico (en adelante LSSI-CE).

La DCE define a los proveedores de servicios de internet (o ISP ,Internet Service Provider) como cualquier persona física o jurídica que suministre un servicio de la sociedad de la información <sup>4</sup>. Esta Directiva establece un régimen de exención de la responsabilidad de los proveedores si se cumplen ciertas condiciones en función del servicio que presten:

---

<sup>3</sup> Directiva 2000/31/CE del Parlamento Europeo y del Consejo, de 8 de junio de 2000, relativa a determinados aspectos jurídicos de los servicios de la sociedad de la información, en particular el comercio electrónico en el mercado interior (Directiva sobre el comercio electrónico, DCE).

<sup>4</sup> Se entenderá por "servicios de la sociedad de la información": servicios en el sentido del apartado 2 del artículo 1 de la Directiva 98/34/CE, modificada por la Directiva 98/48/CE.

- 1- Mera transmisión (art. 12): servicios que consisten en transmitir, en una red de comunicaciones, datos facilitados por el destinatario del servicio, o en facilitar acceso a una red de comunicaciones.
- 2- Memoria tampón (Caching) (art. 13): servicios que consisten en la transmisión y almacenamiento automático, provisional y temporal de datos facilitados por el destinatario del servicio.
- 3- Alojamiento de datos (Hosting) (art. 14): servicios que consisten en almacenar datos facilitados por el destinatario del servicio.

Con posterioridad a la DCE del año 2000 han ido surgiendo otras actividades que no ha sido fácil encajar en alguna de estas tres categorías. Por ejemplo, los servicios que permiten cargar software o contenido como música o películas, o los chats, plataformas de vídeos, páginas de comercio electrónico, blogs o redes sociales han sido genéricamente englobados en la actividad de hosting recibiendo una protección diferente por los Estados Miembros. Sin embargo, estos servicios que procesan, gestionan e indexan la información no tienen un anclaje claro en la DCE.

La DCE establece una distinción entre prestadores activos y pasivos de forma que sólo se puedan acoger a la exención estos últimos. El requisito de pasividad radica en que la actividad que realicen sea meramente técnica y automática, es decir que tengan una función neutral. El Tribunal de Justicia de la Unión Europea desarrolló el criterio de la neutralidad como la ausencia de control del negocio, lo que equivale a no tener participación en la creación o distribución de contenidos <sup>5</sup>. Además, la exención sólo procede respecto de contenidos creados por terceros ajenos al prestador. <sup>6</sup>

---

<sup>5</sup> ARROYO AMAZUELAS, E.: La responsabilidad de los intermediarios en internet ¿puertos seguros a prueba de futuro?, Cuadernos de Derecho Transnacional, vol. 12, Nº 1, marzo 2020

<sup>6</sup> ARRESE IRIONDO, M. N., La responsabilidad de los diarios digitales por los comentarios de sus usuarios, RVAP, 99, 2014, p.391

En ocasiones, no es fácil discernir si el proveedor juega un papel activo o pasivo. Por ejemplo, en el caso de *L'oréal vs eBay* <sup>7</sup>, en el ámbito del comercio electrónico, el TJUE consideró que *eBay* tenía un papel activo en el servicio de asistir a los vendedores y promocionar y fomentar las ventas de productos falsificados o que se comercializan sin el consentimiento del titular de la marca. Y que, por lo tanto, ese papel activo debía llevarle al conocimiento de que estaba almacenando ofertas de productos de forma ilícita.<sup>8</sup>

Las condiciones para que los proveedores se puedan acoger al régimen de exención de responsabilidad, además de tener un papel neutral en el servicio que presta, son las siguientes: que el prestador no tenga conocimiento efectivo de que la actividad o la información es ilícita, o que en cuanto lo tenga actúe con prontitud para retirar los datos o imposibilitar su acceso.

Dado que el concepto de “conocimiento efectivo” puede ser ambiguo se analiza a continuación.

## **2.1. Concepto de “conocimiento efectivo”**

Es conveniente aclarar el significado que debe darse al concepto de “conocimiento efectivo”, de forma que la LSSI aclara que el conocimiento será efectivo *“cuando un órgano competente haya declarado la ilicitud de los datos, ordenado su retirada o que se imposibilite el acceso a los mismos, o se hubiera declarado la existencia de la lesión, y el prestador conociera la correspondiente resolución, sin perjuicio de los procedimientos de detección y retirada de contenidos que los prestadores apliquen en virtud de acuerdos voluntarios y de otros medios de conocimiento efectivo que pudieran establecerse”* <sup>9</sup>.

---

<sup>7</sup> STJUE C-324/09, de 12 de julio de 2011, *L'Oréal v eBay* (ECLI:EU:C:2011:474)

<sup>8</sup> ARROYO AMAZUELAS, E. *Op. Cit.*, p.815

<sup>9</sup> Art. 16.1, apartado 2.º LSSI

En principio, se podría interpretar que sólo se considera que los intermediarios han tenido “conocimiento efectivo” de una ilicitud cuando una resolución así lo dicte. No obstante, el Tribunal Supremo en 2009 amplió esa interpretación de forma que se incluyen los casos<sup>10</sup> en los que existan “*hechos o circunstancias aptos para posibilitar, aunque medianamente o por inferencias lógicas al alcance de cualquiera, una efectiva aprehensión de la realidad de que se trate*”<sup>11</sup>. De esta forma, no en todos los casos en los que no existe una resolución que declare la ilicitud podrá acogerse el intermediario a la exención de responsabilidad. Es decir, el conocimiento indiciario también es considerado conocimiento efectivo. Los casos en los que el Tribunal Supremo se ha pronunciado en este sentido han sido casos en los cuales la ilicitud del contenido era patente y evidente por sí sola.<sup>12</sup>

El riesgo de la ampliación de la interpretación del “conocimiento efectivo” radica en que el intermediario ya no sólo tiene la obligación de retirar un contenido declarado ilícito, sino que, además, tiene que hacer el juicio de valorar si un contenido es ilícito. Esta valoración que antes era potestad de los jueces y tribunales ahora comienza a estar en manos de los intermediarios. Esto tiene sentido si consideramos los foros, redes sociales y chats de plataformas digitales como espacios privados autorregulados por sus administradores. Por el contrario, no resulta tan coherente si los consideramos espacios públicos de intercambio de ideas, donde ahora se producen las conversaciones que antes se producían en bares y otros espacios públicos. La consideración del espacio digital como espacio público o cuasi público ha justificado la progresiva regulación de la red y el establecimiento de límites para abordar la protección de derechos.

---

<sup>10</sup> Esto sólo aplica para los casos en los que se exija responsabilidad civil, es decir, que se pretenda reclamar por daños y perjuicios.

<sup>11</sup> SSTs de 9 de diciembre de 2009 (recurso 914/2006), de 10 de febrero de 2011 (recurso 1953/2008) y de 4 de diciembre de 2012 (recurso 1626/2011). ACEDO PENCO, Ángel, «La responsabilidad civil extracontractual por atentados contra la dignidad divulgada mediante los servicios de la sociedad de la información en los ordenamientos comunitario y español», en Anuario de la Facultad de Derecho, núm. XXIV, 2006, p. 114.

<sup>12</sup> ARRESE IRIONDO, M. N. *Op. Cit.*, p. 394

## **2.2. Mecanismos de defensa de los afectados**

Los usuarios pueden comunicar la existencia de una ilicitud a la plataforma con el fin de que esta elimine un determinado contenido. Y además, si no lo hiciera, no podría esgrimir que no tenía conocimiento sobre la existencia del ilícito si en el futuro se le exigieran responsabilidades.

En algunos Estados Miembros, la doctrina exige que las personas al notificar a la plataforma de la existencia de contenido ilícito, incluyan en la notificación información como la descripción de los hechos controvertidos y su localización, las razones por las que debe suprimirse el contenido incluyendo las disposiciones legales a las que se haga referencia y la prueba de los hechos, así como una copia de la comunicación dirigida al autor de la información si su identidad es conocida. En España se han dado propuestas similares, como la exigencia de adjuntar copia de la resolución que acredite que la información a la que remitía el motor de búsqueda es falsa.<sup>13</sup>

## **2.3. Prohibición de la obligación de supervisión**

La DCE en su art. 15 señala que no se puede imponer a los prestadores de servicios una obligación general de supervisar los datos que transmitan o almacenen, ni de realizar búsquedas activas de indicios de actividades ilícitas. La LSSI, en España, contempla la posibilidad de que los intermediarios incorporen procedimientos de detección y retirada de contenidos como consecuencia de acuerdos voluntarios alcanzados por los propios prestadores<sup>14</sup>. Además, el TS interpretó que el establecimiento de tales mecanismos podría ser útil para probar la ausencia de conocimiento efectivo de una ilicitud llegado el caso<sup>15</sup>.

Progresivamente desde la UE se ha ido fomentando cada vez más la autorregulación. Algunos ejemplos de esto son el Código de conducta para la lucha

---

<sup>13</sup> Sentencia TJUE. Asunto C-131/12 (Google Spain, S.L., Google Inc. / Agencia Española de Protección de Datos, Mario Costeja González) de 13 de mayo de 2014

<sup>14</sup> Art. 16.1, segundo apartado LSSI.

<sup>15</sup> STS de 7 de enero de 2014

contra las declaraciones de incitación al odio en línea<sup>16</sup>, o el Memorandum de Acuerdo sobre la falsificación de mercancías<sup>17</sup>. Esto ha permitido a los Estados Miembros eludir la prohibición de control general y, de alguna forma, delegar la vigilancia de internet en operadores privados<sup>18</sup>.

## 2.4. Evolución de la doctrina

La doctrina de la Unión Europea ha pasado de la neutralidad a la imparcialidad de los operadores. 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>19</sup>.

En el ámbito del ejercicio de la libertad de expresión en internet, la UE ha ido tomando medidas para contener los posibles abusos, como 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*, en el cual se establece que las personas jurídicas pueden ser responsables “cuando la falta de vigilancia o control” por su parte haga posible la materialización de conductas que alteren el orden público, sean amenazadoras, insultantes o abusivas.<sup>20</sup> De forma que, como resumen, los prestadores de servicios no tienen responsabilidad

---

<sup>16</sup> Code of conduct for countering illegal hate speech online, disponible en: [https://ec.europa.eu/newsroom/just/item-de-tail.cfm?item\\_id=54300](https://ec.europa.eu/newsroom/just/item-de-tail.cfm?item_id=54300)

<sup>17</sup> MoU on Counterfeit Goods, disponible en: [https://ec.europa.eu/growth/industry/intellectual-property/enforcement/memorandum-understanding-sale-counterfeit-goods-internet\\_es](https://ec.europa.eu/growth/industry/intellectual-property/enforcement/memorandum-understanding-sale-counterfeit-goods-internet_es)

<sup>18</sup> ARROYO AMAZUELAS, E.*Op. Cit.*, p.831

<sup>19</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.

<sup>20</sup> VALIENTE MARTÍNEZ, F. (2023). La libertad de expresión y las redes sociales: de la doctrina de los puertos seguros a la moderación de contenidos. Revista de Filosofía del Derecho y derechos humanos (48) p.174.

sobre los contenidos digitales, pero se espera su colaboración en la persecución de posibles ilícitos.

Es paradójico que las condiciones para que los prestadores de servicios puedan acogerse al régimen de exención de responsabilidad sean tener un rol pasivo o no tener conocimiento efectivo, es decir, “no tocar nada” o “no saber nada”, al mismo tiempo que se les culpa de que ciertos ilícitos se hayan podido materializar por su falta de vigilancia. Es decir, existe el peligro de que la adopción voluntaria de medidas para prevenir el ilícito se identifique con un conocimiento de la ilegalidad que acabe terminando por calificar al proveedor de "activo".<sup>21</sup>

Los Estados han ido aportando normas de Derecho nacional imponiendo requisitos de diligencia a los intermediarios. No obstante, considerando el carácter transfronterizo de internet, la regulación nacional puede ocasionar divergencias que afecten negativamente al mercado interior en el que la libre circulación de mercancías y servicios está garantizada<sup>22</sup>.

El desarrollo de nuevos modelos de negocio y el uso cotidiano de sus servicios por la mayoría de los ciudadanos ocasiona nuevos riesgos como la difusión de contenidos ilegales o la utilización de sistemas algorítmicos que alteran artificialmente la propagación de unos contenidos o el silenciamiento de otros.

En este contexto se hizo necesaria la creación de una normativa moderna que se materializó con el Reglamento (UE) 2022/2065 de Servicios Digitales, o DSA por sus siglas en inglés Digital Services Act. El objeto de esta ley es regular la actividad de los servicios intermediarios ofrecidos a destinatarios que tengan su lugar de establecimiento o estén

---

<sup>21</sup> ARROYO AMAZUELAS, E. *Op. Cit.*, p.823

<sup>22</sup> MORETÓN TOQUERO, M. ARANZAZU: *Entre Scilla y Caribdis: Los intermediarios digitales y la moderación de contenidos*, Rivista N°: 4/2023 Associazione Italiana Costituzionalisti. p.335.

situados en la Unión, con independencia de dónde tengan su lugar de establecimiento dichos intermediarios (art. 1.1 DSA). Por lo tanto, se aplica a empresas estadounidenses cuando presten servicios en la UE.

La novedad respecto a la DCE es la incorporación de obligaciones previas de diligencia, publicidad y transparencia informativa, así como procedimientos de notificación y actuación, medidas de suspensión, garantías de los usuarios, gestión de riesgos, etc. Con estos instrumentos se persigue crear un entorno más seguro y predecible para empresas y usuarios.

Frente a los contenidos ilícitos y nocivos, la Ley ofrece mecanismos rápidos y eficaces de denuncia y retirada, y un mayor respeto a los derechos de los usuarios<sup>23</sup>. Se ha evolucionado al modelo de inmunidad o exención de responsabilidad condicionada. Así la DSA establece diferentes obligaciones de diligencia para los intermediarios en función de la naturaleza de sus servicios (establece la misma clasificación de la DCE de 2000 a la que se ha hecho referencia anteriormente), de su tamaño y de su impacto. De esta forma, las “plataformas en línea” tienen obligaciones específicas, en atención a su función no sólo de alojamiento de contenido sino también de difusión del mismo. Las “plataformas en línea de muy gran tamaño” (y motores de búsqueda en línea de muy gran tamaño), definidas en el art.33 DSA (más de 45 millones de usuarios mensuales en la UE), estarían sujetas a obligaciones específicas de mayor intensidad. Esto está en coherencia con su mayor relevancia y por lo tanto potencial riesgo, y al hecho de que su volumen les permite implementar tecnologías más efectivas.

Cabe mencionar también, que la DSA establece un deber de actuación de los intermediarios *“de manera diligente, objetiva y proporcionada”* y *“con la debida consideración de los derechos e intereses legítimos de todas las partes implicadas, incluidos los derechos fundamentales de los destinatarios del servicio, como la libertad*

---

<sup>23</sup> MORETÓN TOQUERO, M. ARANZAZU. *Op. Cit.* p.336

*de expresión, la libertad y el pluralismo de los medios de comunicación y otros derechos y libertades fundamentales amparados por la Carta”.*

### **3. COMPARATIVA ENTRE LA REGULACIÓN DE LA UE Y EEUU**

Como ya se ha descrito en el apartado anterior, las primeras regulaciones europeas en materia de responsabilidad de intermediarios en internet, materializadas en la DCE, se inspiraron en la americana Sección 230 de la Communications Decency Act. En un principio, seguían una lógica liberal basada en la irresponsabilidad de los intermediarios por los contenidos que circulen por sus redes. Sin embargo, en Europa, ya desde la DCE y más claramente en la DSA, se apreció una diferencia sustancial en comparación con Estados Unidos, la imposición del deber de diligencia. Este deber comporta la obligación de retirar el contenido sobre el cual se tenga el conocimiento efectivo de su carácter ilícito. Otra diferencia fundamental fue el requisito de que los proveedores de servicio no tengan un papel activo en la creación o edición del contenido para poder beneficiarse de la exención de responsabilidad.

Por otra parte, para la jurisprudencia norteamericana clásica, las personas jurídicas no eran titulares de derechos fundamentales ya que se concebían como “un ser artificial, invisible, intangible y que existe sólo en virtud de una ley”<sup>24</sup>. Este status quo fue alterado en el año 2010 con la sentencia de la Corte Suprema en el caso *Citizens United*.

*Citizens United* es una entidad sin ánimo de lucro que, en el marco de las primarias presidenciales del Partido Demócrata de 2008, lanzó un documental altamente crítico con la entonces senadora y candidata Hillary Clinton. Tras su exhibición en algunas salas de cine, el documental comenzó a distribuirse en DVD y a través de varias plataformas de televisión en línea. El acuerdo con estas plataformas establecía que el documental estaría disponible de forma gratuita para los espectadores, mientras

---

<sup>24</sup> Trustees of Dartmouth College v. Woodward, 17 U.S. (4 Wheat.) 518 (1819)

que *Citizens United* asumiría el pago de un canon. Sin embargo, la Comisión Electoral Federal prohibió dicha distribución, argumentando que vulneraba ciertas normas sobre la financiación de campañas electorales. En respuesta, *Citizens United* alegó que esta restricción violaba la Primera Enmienda al impedir su contribución a la formación de la opinión pública.

El caso llegó a la Corte Suprema, que falló a favor de *Citizens United* <sup>25</sup>. En la sentencia se equiparó el derecho a la libertad de expresión de personas físicas y jurídicas, bajo el fundamento de que su ejercicio por cuantos 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>26</sup>.

Las razones para dicho reconocimiento se basan principalmente en el hecho de que ciertos valores asociados a la dignidad también se encuentran presentes en las personas jurídicas. Si a esto añadimos que los derechos fundamentales poseen una dimensión axiológica, podemos concluir que empresas, fundaciones, asociaciones, partidos políticos y sindicatos son comúnmente utilizados por los ciudadanos como medios para llevar a cabo actividades relacionadas con el ejercicio de sus derechos<sup>27</sup>.

---

<sup>25</sup> El TS falló con un estrecho margen de cinco votos a favor y cuatro en contra. Resulta igualmente relevante la opinión discrepante, redactada por el juez Stevens y respaldada por los jueces Ginsburg, Breyer y Sotomayor, argumentaba que la diferencia entre personas físicas y jurídicas es evidente y que ninguna corporación tiene una opinión propia, sino que refleja las de sus administradores. Además, sostenían que la decisión de la mayoría podría llevar a que las grandes corporaciones influyan desproporcionadamente en la opinión pública, en lugar de fortalecerla

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>26</sup> *Citizens United v. Federal Election Commission*, 558 U.S. 310 (2010).

<sup>27</sup> L. M. DÍEZ-PICAZO GIMÉNEZ, *Sistema de derechos fundamentales*, Tirant lo Blanch, 5ª ed., Valencia, 2021, pp. 129-132.

La regulación norteamericana considera que internet es el entorno perfecto para la comunicación y formación de la opinión pública y, por lo tanto, deben proyectarse sobre él las garantías constitucionales de la Primera Enmienda que protege la libertad de expresión<sup>28</sup>. No obstante, cae en la siguiente contradicción: la Primera Enmienda protege la libertad de expresión de los usuarios, pero también la de las plataformas para cancelar a los usuarios. Esta contradicción se puede apreciar más claramente en el caso de Trump contra Twitter que se expone a continuación.

La Corte Suprema caracterizó como foros públicos espacios como el perfil de representantes públicos y por lo tanto no pueden bloquear a los usuarios por los comentarios que hagan en su perfil. Pero, de forma irónica, el propio Donald Trump fue censurado y cancelado por la misma red social, que actuaba al amparo de la sección 230 y de la Constitución. Es decir, Trump no puede bloquear a ningún usuario en su perfil porque es un foro público, pero Twitter sí que puede bloquearlo a él<sup>29</sup>. La Corte Suprema en EEUU ha reiterado que la Primera Enmienda sólo es eficaz frente a restricciones de la libertad de expresión procedentes del Gobierno y no para las procedentes de compañías privadas<sup>30</sup>. De hecho, sólo resulta una vulneración de la libertad de expresión por parte de las plataformas si se demuestra que están operando bajo el mando del Gobierno, según la *state-action doctrine*. Esto es lo que ocurrió cuando se desestimó la pretensión de Trump de que ese considerase vulnerada su libertad de expresión cuando Twitter le bloqueó su cuenta. Esto sólo podría haber sido estimado si se demostrara que Twitter actuaba siguiendo instrucciones del poder público.

De forma que en EEUU se niega la eficacia directa del derecho a la libertad de expresión frente a empresas privadas. En definitiva, la facultad de las plataformas de moderar los contenidos queda amparada no solamente por la sección 230 de la

---

<sup>28</sup> Reno vs. American Civil Liberties Union 521 U.S. 844, de 1997

<sup>29</sup> Vázquez Alonso, V. 2022, *La censura “privada” de las grandes corporaciones digitales y el nuevo sistema de la libertad de expresión*. Teoría y Derecho, 32, p.113

<sup>30</sup> SÁENZ ROYO, E. *Op. Cit.* p. 5

Communications Decency Act sino también por la Primera Enmienda. Por otra parte, la imposición legal o judicial de la obligación de mantener contenidos en contra de su criterio o de sus propias políticas sería *compelled speech* y, por tanto, una violación de la Primera Enmienda<sup>31</sup>.

La reacción de Trump ante estos sucesos fue en la línea de intentar limitar la discrecionalidad de las plataformas a la hora de moderar contenidos. Así estableció la *Executive Order on Preventing Online Censorship*, de 28 de mayo de 2020. Esta orden proponía que los reguladores federales pudieran sancionar a las compañías tecnológicas si usaban criterios discriminatorios en la moderación de contenidos. Esta orden fue revocada por Biden en 2021.

### **3.1. El caso NetChoice, LLC v. Paxton**

En septiembre de 2021 el Estado de Texas promulgó la ley HB 20 que prohíbe que las redes sociales más grandes (más de cincuenta millones de usuarios activos) como Facebook, X y YouTube bloqueen, eliminen o desmoneticen contenido de expresión según las opiniones de los usuarios. Esta ley también exige transparencia en las políticas de moderación de contenido y establece procedimientos de quejas para los usuarios.

*NetChoice y la Computer & Communications Industry Association* presentaron una demanda impugnando dos disposiciones de la ley por ser inconstitucionales:

- La Sección 7, que prohíbe la censura basada en el punto de vista de las publicaciones de los usuarios, excepto para contenidos que inciten a actividades delictivas o sean ilegales.
- La Sección 2, que obliga a las plataformas a revelar cómo moderan y promueven los contenidos, publicar una «política de uso aceptable» y mantener un sistema de quejas y recursos para sus usuarios.

---

<sup>31</sup> BARATA MIR, J.: *Op. Cit.* p. 104.

En la demanda argumentaban que estas disposiciones vulneran la libertad editorial amparada por la Primera Enmienda. El Tribunal de Distrito dictó una medida cautelar sosteniendo que estas disposiciones son inconstitucionales. No obstante, el Tribunal de Apelaciones en septiembre de 2022 falla a favor del Estado argumentando que la Primera Enmienda no protege el derecho a censurar reclamado por las plataformas y que las leyes que prohíben la censura están protegiendo el derecho a libertad de expresión de los usuarios. En la argumentación alude a que precisamente la sección 230 no consideró a las plataformas como editores tradicionales, y por lo tanto no tienen un discurso propio que proteger, sino una actuación sobre el discurso ajeno, el de los usuarios. El tribunal señaló expresamente:

*“La sección 230 refleja el juicio del Congreso de que las Plataformas no funcionan como editores tradicionales y no están «hablando» con el contenido enviado por los usuarios. El juicio del Congreso refuerza nuestra conclusión de que la censura de las plataformas no es un discurso bajo la Primera Enmienda.*

*[...]*

*La sección 230 socava los dos argumentos de las plataformas para sostener que su censura de usuarios es discurso protegido. Hay que recordar que se basan en dos argumentos clave: primero, sugieren que el contenido enviado por el usuario que alojan es su discurso; y segundo, argumentan que son editores similares a un periódico. La sección 230, sin embargo, instruye a los tribunales a no tratar las plataformas como «editores u oradores» del contenido enviado por el usuario que alojan.”<sup>32</sup>*

Por lo tanto, el Tribunal de Apelaciones del 5º Circuito que, aunque la 230 de la Decency Act ampare el control editorial, este ha de interpretarse restrictivamente y no

---

<sup>32</sup> NetChoice, L.L.C. v. Paxton, 49 F.4th 439 (5th Cir. 2022)

da derecho para censurar libremente, especialmente cuando se trata de contenido con carácter ideológico. En definitiva, se reconoció la facultad del Parlamento de Texas para impedir que las plataformas vulneren la libertad de sus usuarios en ese Estado.

El caso se llevó a la Corte Suprema que falló a favor de NetChoice. En la misma sentencia también resolvió el caso de *Moody v. NetChoice*, en el que se planteó una situación similar, pero en el Estado de Florida<sup>33</sup>. En la motivación se pueden destacar los siguientes argumentos<sup>34</sup>.

La Primera Enmienda protege a las entidades dedicadas a actividades expresivas, incluida la recopilación y conservación del discurso de otros, de verse obligadas a dar cabida a mensajes que prefieren excluir. El gobierno no puede justificar la interferencia con las opciones editoriales de un orador privado simplemente alegando un interés en mejorar o equilibrar el mercado de las ideas. Continúa argumentando que es probable que estos principios se apliquen las prácticas de moderación de contenidos de plataformas de redes sociales, lo que indica que leyes estatales como la HB 20 podrían vulnerar la protección otorgada por la Primera Enmienda.

La regulación de Texas de las políticas de moderación de contenidos tiene como objetivo alterar el discurso que se muestra en estas plataformas, lo que refleja la desaprobación del Estado de las prácticas actuales de selección y moderación de contenidos de las plataformas. En virtud de la Primera Enmienda, Texas no puede imponer sus preferencias sobre la forma en que las entidades privadas seleccionan y presentan el discurso, ya que esto equivaldría a un control gubernamental sobre la expresión de las ideas<sup>35</sup>.

---

<sup>33</sup> *Moody v. NetChoice, LLC*, 603 U.S. 707 (2024)

<sup>34</sup> <https://www.oyez.org/cases/2023/22-555>

<sup>35</sup> *NetChoice, LLC v. Paxton*. 1 julio 2024, 603 U. S. 707 (2024)

En comparación con esta amplia discrecionalidad de la que gozan las plataformas estadounidenses, en la Unión Europea se evolucionó hacia la vinculación de la función que desempeñan las plataformas al interés general. Este giro en la comprensión de la naturaleza jurídica de las grandes plataformas partió de un cambio en la concepción de las consecuencias que podía tener su uso con el transcurso del tiempo. De forma que, en los años 90, se consideraba que el uso de estos medios iba a posibilitar el desarrollo y acceso al “mercado de ideas”<sup>36</sup>, pudiendo acceder a opiniones diferentes a las propias, y por tanto ser conscientes de la falibilidad de las opiniones propias, y la deliberación basada en valores democráticos<sup>37</sup>. Después, en Europa se pasó a tomar conciencia de que la realidad es que existe una tendencia a acudir a la red para consolidar nuestros propios prejuicios, lo cual redundaba en un incremento de la polarización política<sup>38</sup>. Además, esto es un efecto artificial ya que las redes sociales están diseñadas para incentivar su uso y enganchar al usuario, por lo que los sistemas algorítmicos de recomendación priorizan los gustos y preferencias del usuario, en este caso preferencias políticas.

Otra idea preconcebida en Europa era la de que en Internet no habría barreras de entrada para la innovación y por lo tanto habría una oferta amplia, variada y competitiva. Sin embargo, esto ha resultado ser al revés, se han forjado monopolios transnacionales, dotados de una capacidad extraordinaria para analizar de forma automatizada las pautas del comportamiento humano y las dinámicas sociopolíticas. Lo que ha demostrado la evidencia, es que se ha provocado una concentración en los medios de comunicación y que además estos tienen un carácter parcial.

Este cambio en la comprensión de las dinámicas del ejercicio de la libertad de expresión e información ha llevado a Europa a redefinir su marco normativo, desde la

---

<sup>36</sup> Traducción literal de la conocida formulación académica y jurisprudencial norteamericana del marketplace of ideas inspirada en *On Liberty*, el ensayo de John Stuart Mill.

<sup>37</sup> Vázquez Alonso, V. *Op. Cit.* p.113

<sup>38</sup> Sunstein, C. 2018. *Divided Democracy in the Age of social media*. Princeton University Press.

aprobación de la Directiva de Comercio Electrónico de 2000, pasando por legislaciones nacionales que han introducido matices y modificaciones al texto original, hasta la adopción de la Ley de Servicios Digitales (Digital Services Act).

En contraste, el derecho en Estados Unidos se mantiene fiel a los principios del derecho privado, que rigen las obligaciones y responsabilidades de las plataformas digitales. En EEUU se da más prioridad que en Europa al derecho a la libertad de expresión de las plataformas o libertad editorial, así como también, a la libertad de empresa.

Un ejemplo del cambio en la concepción de la responsabilidad de los intermediarios en Europa, que supuso además un punto de inflexión en la jurisprudencia del TJUE, fue el caso de Google y el derecho al olvido digital. No sólo se abrió paso a la responsabilidad de Google por contenido aportado por los usuarios, sino que, además, se le insta a ejercer la función de tutela en primera instancia del derecho a la intimidad y a la protección de datos<sup>39</sup>. Cada vez es más la responsabilidad exigida a las grandes corporaciones en el ejercicio de funciones de interés público.

Por otra parte, en Europa se permite que las autoridades independientes controlen los sistemas algorítmicos de las grandes plataformas, a diferencia de EEUU donde las recomendaciones algorítmicas son “simplemente parte de la infraestructura a través de la cual todo el contenido de sus plataformas es filtrado”. Por tanto, en Europa los Estados miembros pueden conceder a las autoridades la facultad de sancionar a aquellas plataformas cuyos sistemas algorítmicos favorezcan el discurso de odio<sup>40</sup>.

---

<sup>39</sup> Guichot, E. 2019, *El reconocimiento y desarrollo del derecho al olvido en el derecho europeo y español*. Revista de Derecho Político, 106,pp, 45-92.

<sup>40</sup> SÁENZ ROYO, E. *Op. Cit.* p. 9

#### **4. LAS POLÍTICAS DE MODERACIÓN DE CONTENIDOS EN INTERNET**

El concepto “moderación de contenidos” se refiere a las prácticas dirigidas a la evaluación, categorización y posible eliminación de contenido considerado inadecuado según los criterios establecidos por los administradores de una plataforma en internet.<sup>41</sup>

##### **4.1. La legitimidad de las políticas de moderación de contenidos en su doble dimensión pública y privada**

La legitimidad de la facultad de moderación de contenidos se basa por una parte (en su dimensión privada) en la relación legal entre las empresas propietarias de las redes sociales y los usuarios de estas. Cuando los usuarios se registran en una red social (requisito que es normalmente exigido para su uso) aceptan las condiciones impuestas en un contrato de adhesión. Estos contratos se caracterizan por el hecho de que una de las partes, en este caso la empresa propietaria de la red social impone las condiciones del contrato. De esta forma, el usuario sólo tiene la posibilidad de rechazarlas y no poder acceder al uso de la red social o aceptarlas en su totalidad. Por lo tanto, cualquier usuario que crea un perfil está consintiendo plenamente la aceptación de sus políticas y condiciones de uso.<sup>42</sup>

Con frecuencia, estas “condiciones de uso” o “normas de la comunidad” no definen con concreción los límites a los que se va a someter la actividad de los usuarios, sino que son disposiciones amplias que dejan mucho margen de discreción a la propia plataforma. Las medidas de moderación de contenidos pueden estar motivadas por su carácter ilícito (incumplimiento de la ley) o por su incompatibilidad con las normas

---

<sup>41</sup> VALIENTE MARTÍNEZ, F. (2023). La libertad de expresión y las redes sociales: de la doctrina de los puertos seguros a la moderación de contenidos. *Revista de Filosofía del Derecho y derechos humanos* (48), 182.

<sup>42</sup> B. RODRÍGUEZ ARROYO, “Las redes sociales: usuarios, términos y condiciones”, EOB, Fashion, Luxury and Retail, 2020.

impuestas por las plataformas (incumplimiento contractual)<sup>43</sup>. Es decir, existe un doble control, público y privado<sup>44</sup>. Si bien es cierto, que muchas disposiciones contractuales aluden a preceptos legales, existen otras disposiciones que tiene su origen en la autonomía privada. Por lo tanto, uno de los retos es el de dibujar la frontera entre el discurso ilícito y el considerado antisocial, inapropiado o incluso peligroso para la seguridad ciudadana pero no ilícito.

Por otra parte, la legitimación de las plataformas digitales para moderar los contenidos que circulan por ellas también proviene (en su dimensión pública) de la atribución normativa que la ley les concede.

Así, hemos asistido a una progresiva autorregulación de los proveedores a través de un régimen jurídico privado. El riesgo de estas prácticas reside en delegar la regulación del ejercicio de un derecho fundamental, la libertad de expresión, en manos de la gestión privada. Actualmente, el Reglamento sobre una internet abierta prohíbe el bloqueo de contenidos no exigido por ley y el bloqueo voluntario de contenido lícito.<sup>45</sup>

En la Directiva 2019/790, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital, permite utilizar tecnología de reconocimiento de contenidos, tanto para filtrar como para bloquear. Uno de los objetivos de esta directiva es armonizar las reglas del juego en la Unión Europea para operar en un marco común en especial en el ámbito de la propiedad intelectual.

#### **4.2. Las obligaciones que impone la Digital Service Act a los intermediarios**

Por otro lado, las obligaciones que impone la DSA afectan en gran medida al desarrollo de la moderación de contenidos de las plataformas. Entre ellas tiene especial

---

<sup>43</sup> Art. 14 DSA

<sup>44</sup> MORETÓN TOQUERO, M. ARANZAZU. *Op. Cit.* p.350

<sup>45</sup> ARROYO AMAZUELAS, E. *Op. Cit.*, p.834

relevancia la obligación de transparencia informativa, que exige la publicación de informes anuales sobre las actividades de moderación de contenidos (art. 15 DSA). Así como la publicación de una base de datos que contenga las decisiones con relación a la restricción de la accesibilidad a determinados contenidos. Asimismo, los servicios de alojamiento de datos, incluidas las plataformas en línea, deben disponer de mecanismos de denuncia que permitan a cualquier persona o entidad notificar la presencia de contenidos ilícitos de manera concreta y motivada (art. 16 DSA).

Los intermediarios tienen la obligación de informar de forma detallada y justificada sobre las medidas restrictivas que adopten en relación con los servicios que prestan, estas medidas pueden ser la eliminación de contenidos, la suspensión de cuentas y la restricción de pagos monetarios (art. 17 DSA).

Para garantizar la protección de los usuarios, la DSA introduce la obligación de que las plataformas en línea ofrezcan un mecanismo interno eficaz para la gestión de reclamaciones contra decisiones de moderación de contenidos (art. 20 DSA). Estos procedimientos deben contar con la supervisión de personal cualificado y no depender exclusivamente de medios automatizados.

Además, los usuarios podrán recurrir decisiones restrictivas cuando no hayan sido resueltas a través del sistema interno de tramitación de reclamaciones mediante un sistema de resolución extrajudicial de conflictos. (art. 21 DSA). Las plataformas también deben implementar un sistema de alertadores fiables o *trusted flaggers*, cuyas notificaciones tendrán prioridad en la detección de contenidos ilícitos (art. 22 DSA). Para garantizar un uso adecuado de los mecanismos de denuncia se prevé la imposición de medidas contra los usuarios que, de forma reiterada, difundan contenido ilícito o realicen denuncias manifiestamente infundadas (art. 23 DSA).

La DSA establece un régimen especial para las plataformas en línea y los motores de búsqueda de muy gran tamaño definidas en el art. 33 como aquellos que tengan un

número de consumidores europeos igual o superior a 45 millones (que es aproximadamente un 10% de los ciudadanos de la UE). Para estas plataformas se establece la obligación de llevar a cabo evaluaciones de riesgos sistémicos como la difusión de contenido ilícito, la desinformación o la afectación de procesos democráticos (art. 34 DSA) y poner en funcionamiento sistemas internos para mitigar estos riesgos de forma razonable, proporcionada y eficaz (art. 35 DSA). Además, se impone la realización de auditorías anuales independientes (art. 37 DSA) y mayor transparencia sobre el funcionamiento de sus sistemas de recomendación y publicidad en línea (art. 38 DSA).

Esta compleja regulación de la DSA es el fruto de la existente preocupación en Europa por la difusión de contenidos ilícitos, la desconfianza respecto a la autorregulación del discurso en la Red o la realización de conductas engañosas que puedan potencialmente limitar derechos fundamentales o afectar al desarrollo de procesos democráticos como campañas de desinformación.<sup>46</sup>

La DSA define la moderación de contenidos como *“las actividades, estén o no automatizadas, realizadas por los prestadores de servicios intermediarios, que estén destinadas, en particular, a detectar, identificar y actuar contra contenidos ilícitos o información incompatible con sus condiciones generales, que los destinatarios del servicio hayan proporcionado”*. Algunos ejemplos de esto son la restricción de visibilidad o la relegación de contenidos, la desmonetización o incluso la supresión o suspensión de la cuenta del usuario.

Esta actividad puede llevarse a cabo de forma automatizada con sistemas algorítmicos de detección de contenidos o mediante profesionales contratados a tal efecto. Por los riesgos que tanto uno como otro método pueden comportar, la DSA pone énfasis en la transparencia de los modelos de adopción de medidas restrictivas y la motivación que subyace tras ellas. Así como, en caso de que existan errores en las

---

<sup>46</sup> MORETÓN TOQUERO, M. ARANZAZU. *Op. Cit.* p.347

restricciones o discrepancias de los usuarios respecto a ellas, la ley prevé la existencia de mecanismos de reclamación. Como en una sentencia, la motivación deberá incluir el fundamento jurídico al que se hace referencia, así como el encaje de los hechos (la creación y/o difusión de contenido) en esa norma jurídica y qué vías de recurso están disponibles para el usuario.<sup>47</sup>

La DSA se posiciona a favor del empleo de recursos humanos para la moderación de contenidos. Por ejemplo, entre los deberes de transparencia de las plataformas de gran tamaño incluye informar de qué volumen de recursos humanos se dedica a esta tarea.

#### **4.3. El caso Twitter vs Vox**

Respecto a la doble dimensión pública y privada de las plataformas digitales y la autorregulación de esta, surge el siguiente dilema: “¿Quién controla al controlador?”. Si desde la perspectiva pública esta facultad de control tiene su legitimidad original en el Estado y la legitimidad de la facultad de control de las plataformas es una legitimidad derivada, ¿No debería el Estado vigilar la actividad de moderación de contenidos o eso supondría una intromisión en la libertad de empresa y en un contrato privado?

La STS 735/2022, de 28 de febrero, arroja luz sobre esta cuestión. Trae causa de la suspensión de la cuenta de Twitter del partido político Vox, Twitter argumentó que un tuit en el que se afirmaba que «[...] son los musulmanes o las personas de origen magrebí quienes cometen el 93 % de los delitos en Cataluña» constituía una violación de la política relativa a las conductas de incitación al odio vigente en esta red social.

---

<sup>47</sup> MORETÓN TOQUERO, M. ARANZAZU. *Op. Cit.* p.348

La Junta Electoral Central por una parte, considera que Twitter es un “[...] canal de comunicación que permite a sus usuarios el envío e intercambio de mensajes de texto breves a grupos reducidos de seguidores o, de forma abierta, a todo el público” por lo tanto encaja en el concepto de “medios de comunicación privados” contenido en el artículo 66.2 de la Ley Orgánica 5/1985, del Régimen Electoral General (LOREG) sobre la garantía de pluralismo político y social: *“Durante el periodo electoral las emisoras de titularidad privada [...] deberán respetar también los principios de proporcionalidad y neutralidad informativa en los debates y entrevistas electorales...”*

No obstante, en aras del principio de proporcionalidad, la Junta considera que la suspensión de la cuenta tuvo lugar durante un período corto, se debió a un incumplimiento reiterado de las políticas internas de Twitter y los candidatos de Vox pudieron seguir utilizando sus cuentas individuales en esa red social. Por lo tanto, concluye que la medida de Twitter no pudo considerarse desproporcionada. En este caso, la decisión de la Junta se fundamenta también en que los criterios de Twitter son equiparables a las disposiciones contenidas en la legislación nacional respecto al discurso de odio.<sup>48</sup>

Por su parte, el Tribunal Supremo, admitió que la relación entre Twitter y Vox está sujeta a Derecho privado, pero cuestiona hasta dónde puede llegar la Administración electoral a fiscalizar una relación entre particulares pero con incidencia

---

<sup>48</sup> En contra, el voto particular que formuló el vocal don Lázaro al Acuerdo 146/2021, de 25 de febrero, de la JEC. Este voto particular entendió que la acción de Twitter de suspender durante ocho días la cuenta de Vox durante la campaña electoral en Cataluña priva de acceso al conjunto de la actividad de esa cuenta. Añade que la sanción se impone inaudita parte y sin explicitar las razones que llevan a considerar que el tuit incita expresamente a la discriminación, la xenofobia o el racismo. De haber excesos en la libertad de expresión, lo procedente es ir a un procedimiento judicial con garantía de contradicción y defensa. Considera que lo hecho por Twitter es un caso de censura previa proscrita por el art. 20 de la Constitución, «un verdadero secuestro sin resolución judicial y una decisión desproporcionada». En definitiva, que consolida el poder de censura unilateral por las redes sociales y el recurso ante la Administración electoral no soluciona esa carencia de regulación.

en el proceso electoral. Finalmente, la sentencia concluyó que la actuación de la administración fue legítima y proporcionada.<sup>49</sup>

#### **4.4. Los riesgos de la co-regulación**

Dado que la vigilancia total y constante del Estado de la actividad de moderación resulta muy difícil, si no imposible, (y este es el principal motivo por el que se delega la facultad de moderación, una cuestión de practicidad.) Parece que la solución pasa por un diseño de co-regulación (Considerando 3 DSA), para compatibilizar la creación de un “entorno en línea seguro, predecible y digno de confianza”, las libertades de expresión e información, la libertad de empresa, la protección frente a la discriminación, la seguridad y la protección de consumidores y usuarios. Así, las plataformas digitales son más idóneas para ejercer una influencia efectiva en el comportamiento de los usuarios.

No obstante, de esta colaboración entre el Estado y entes privados afloran dos riesgos:

- Que las plataformas abusen de la facultad de moderación, limitando en exceso los derechos y libertades de los usuarios.
- Que los intermediarios se conviertan en un brazo privado del Estado que sea aprovechado por este para limitar “el libre mercado de ideas”, excluyendo del foro público digital discursos políticamente incorrectos.

En otras palabras, el riesgo de que las plataformas puedan limitar la libertad de expresión, y el riesgo de que sea el Estado el que limite esta libertad utilizando a las grandes plataformas como brazo ejecutor.

En cualquiera de los casos, se acerca al concepto de censura. No obstante, en sentido estricto no encaja en la definición del art 20.2 CE que caracteriza la censura como

---

<sup>49</sup> MORETÓN TOQUERO, M. ARANZAZU. *Op. Cit.* p.351

un control previo y la moderación de contenidos tiene un carácter ex post. Sin embargo, sí que podría encontrar un límite en el art. 20.5 CE que establece que “Sólo podrá acordarse el secuestro de publicaciones, grabaciones y otros medios de información en virtud de resolución judicial.”

En cuanto al segundo riesgo, es especialmente sensible la moderación de contenidos no ilícitos, sino que contravengan las condiciones de uso de la plataforma. Esta facultad tiene una legitimación privada basada en el contrato de adhesión que todos los usuarios han aceptado. De forma que, alegando una contravención de estas condiciones podrían retirarse contenidos no ilícitos imponiéndose una “específica militancia democrática” excluyendo discursos en RRSS que en el ámbito analógico estarían permitidos o incluso protegidos<sup>50</sup>. Las RRSS se convertirían por tanto en espacios más restrictivos que las interacciones presenciales<sup>51</sup>.

Esto es agravado por la falta de correspondencia entre las consecuencias derivadas del exceso cancelatorio y la responsabilidad procedente de la inactividad frente a los contenidos ilícitos, previsiblemente favorece que las plataformas prefieran “pecar por exceso” antes que “por defecto”, y que eventualmente tiendan a eliminar cualquier contenido que suscite la mínima sospecha de ilegalidad<sup>52</sup>.

Respecto al problema de la influencia sobre la opinión pública en temas de política, muchas organizaciones han advertido que los intermediarios promueven cierto tipo de contenido con el fin de maximizar la participación de los usuarios, estimular la adicción y fomentar la polarización. Esto no es difícil si se tiene en cuenta que los algoritmos de recomendación de contenidos se basan, entre otras cosas, en las preferencias de cada usuario que a menudo tienen un sesgo político, de tal forma que

---

<sup>50</sup> MORETÓN TOQUERO, M. ARANZAZU. *Op. Cit.* p.353

<sup>51</sup> VÁZQUEZ ALONSO, V.: “La censura “privada” de las grandes corporaciones digitales...”, *op. cit.*, p. 118.

<sup>52</sup> BARATA MIR, J.: *Op. Cit.* p. 101.

un usuario con una determinada ideología recibe más contenido de esa ideología de forma que se retroalimenta y se reafirma en ella. Y esto, puede tener un impacto importante en los derechos fundamentales, sobre todo en la consolidación de sesgos y conductas discriminatorias.

Un ejemplo de esto es lo ocurrido en el conocido como *Facebookgate* que ha sacado a la luz prácticas bastante cuestionables llevadas a cabo por esta red social, en concreto, la relacionadas con el uso 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. No es una excepción, ya que las grandes corporaciones como Google, Amazon, Apple, Microsoft y Meta analizan sistemáticamente las conductas de los usuarios y las convierten en datos con los que obtienen beneficios e incluso comercializan<sup>53</sup>. Esta es una realidad que cada vez genera más preocupación haciendo patente la necesidad de una regulación.

## **5. IMPACTO EN LOS DERECHOS FUNDAMENTALES**

### **5.1. El derecho a la libertad de expresión**

La práctica y la protección de la libertad de expresión se han ido adaptando a lo largo de la historia a los contextos sociales. En nuestros días es un elemento esencial para garantizar una sociedad democrática, plural y libre.

---

<sup>53</sup> VALIENTE MARTÍNEZ, F. (2023). *La libertad de expresión y las redes sociales: de la doctrina de los puertos seguros a la moderación de contenidos*. Revista de Filosofía del Derecho y derechos humanos (48), 193.

La libertad de expresión es un derecho fundamental recogido en el artículo 20 de nuestra Constitución, situado en el Título I, Capítulo II, Sección 1ª *“De los derechos fundamentales y de las libertades públicas”* <sup>54</sup>.

La condición de derecho fundamental implica que es directamente aplicable y exigible ante los tribunales. No es necesario un desarrollo legislativo anterior a su aplicación y exigibilidad. Además, su contenido esencial está sujeto a reserva de ley, es decir, sólo se puede regular su ejercicio por ley, que además, en virtud del art.81 CE deberá ser ley orgánica.

El contenido del derecho a la libertad de expresión del art.20 CE es el siguiente:

*“1. Se reconocen y protegen los derechos:*

*a) A expresar y difundir libremente los pensamientos, ideas y opiniones mediante la palabra, el escrito o cualquier otro medio de reproducción.”*

Este es el contenido de libertad de expresión en sentido estricto, además, la jurisprudencia ha interpretado que están incluidas las creencias y los juicios de valor<sup>55</sup>.

*“b) A la producción y creación literaria, artística, científica y técnica.*

*c) A la libertad de cátedra.*

*d) A comunicar o recibir libremente información veraz por cualquier medio de difusión. La ley regulará el derecho a la cláusula de conciencia y al secreto profesional en el ejercicio de estas libertades.”*

---

<sup>54</sup> En el sistema europeo se encuentra reconocido en el artículo 11 de la Carta de los Derechos Fundamentales de la Unión Europea y en el 10 del Convenio Europeo de Derechos Humanos (en adelante: CEDH). Además, está previsto en el artículo 19 de Declaración Universal de los Derechos Humanos.

<sup>55</sup> STC 6/1988, de 21 de enero.

El apartado d) se refiere al derecho a libertad de información, difiere de la libertad de expresión en que en la libertad de información se exige que esta sea veraz. El TC ha delimitado el concepto de veracidad: no se exige que la información sea verdadera sino *“un deber de diligencia sobre el informador a quien se le puede y debe exigir que lo que transmite como hechos haya sido objeto de previo contraste con datos objetivos”*<sup>56</sup>. Este requisito no se exige respecto de la libertad de expresión puesto que las creencias u opiniones no tienen por qué ser veraces. Por otra parte, la libertad de información también requiere del carácter de interés público, interpretado por el TC como aquella información que es relativa a un asunto público y por lo tanto atañe a la ciudadanía en su conjunto, ya sea por el tema sobre el que versa, ya sea por los sujetos intervinientes.

*“2. El ejercicio de estos derechos no puede restringirse mediante ningún tipo de censura previa.”*

La prohibición de la censura es un apartado importante en el contexto de la moderación de contenidos. Si bien es cierto que esta supone un control a posteriori, cabe preguntarse si una censura posterior tendría cabida en esta previsión constitucional. La respuesta la encontramos en la STC 187/1999, de 25 de octubre, en la que el Tribunal Constitucional interpretó que la “censura” prohibida constitucionalmente es:

*“[...] la intervención preventiva de los poderes públicos para prohibir o modular la publicación o emisión de mensajes escritos o audiovisuales. La presión de ciudadanos o grupos de ellos para impedir esa difusión, aunque consiga obtener el mismo resultado, puede llegar a ser una intromisión en un derecho ajeno, con relevancia penal en más de un caso y desde más de un aspecto, pero no «censura» en el sentido que le da la Constitución”*

---

<sup>56</sup> STC 1/2005, de 17 de enero.

Por lo que se puede concluir que la moderación de contenidos no encaja en este precepto. En cuanto a la llamada “autocensura”, ya que son las propias plataformas las que limitan los contenidos que circulan por ellas, tradicionalmente este término se había usado en sectores como la cinematografía o la prensa. La autocensura está aún más alejada del límite constitucional, ya que a los editores y directores se les permite e impone el examen de sus contenidos previo a su difusión. Por lo que el derecho de veto del editor está reconocido en la Ley de Prensa e Imprenta de 18 de marzo de 1996 <sup>57</sup>.

El apartado 4 versa sobre los límites a los que está sujeto el ejercicio de la libertad de expresión:

*“... 4. Estas libertades tienen su límite en el respeto a los derechos reconocidos en este Título, en los preceptos de las leyes que lo desarrollen y, especialmente, en el derecho al honor, a la intimidad, a la propia imagen y a la protección de la juventud y de la infancia.”*

Dado que los poderes públicos han de salvaguardar otros derechos como el derecho al honor o la no discriminación, la libertad de expresión no puede gozar de una protección absoluta, está sujeta a unos límites. Los tribunales controlan que estas restricciones sean proporcionales, justificadas, adecuadas, eficaces para cumplir el fin legítimo que persiguen y no vulneren el contenido esencial del derecho. Estas restricciones deberán además estar fijadas por ley.

Ya que la propia Constitución no establece una jerarquía entre derechos, cuando se produce una colisión entre distintos derechos fundamentales ha de hacerse un “juicio de ponderación” definido como un examen de la intensidad y trascendencia con la que

---

<sup>57</sup> SÁENZ ROYO, E. *Op. Cit.* p. 15

cada uno de ellos resulta afectado, con el fin de elaborar una regla que permita, dando preferencia a uno a otro, la resolución del caso mediante su subsunción en ella”<sup>58</sup>

Resolviendo una cuestión de inconstitucionalidad, el TC estableció que el art. 20.1 CE no garantiza *“el derecho a expresar (...) con el deliberado ánimo de menospreciar y discriminar, al tiempo de formularlo, a personas o grupos por razón de cualquier condición o circunstancia personal, étnica o social”*<sup>59</sup>. En este sentido, el TC ha dejado claro que no existe un derecho al insulto, pero sí a las críticas, aunque ofendan, pues es presupuesto de las sociedades democráticas<sup>60</sup>.

En cuanto a los límites penales de la libertad de expresión resulta relevante en el contexto de internet mencionar la reforma que hace la LO 1/2015, de 30 de marzo, del artículo 510 CP que prohíbe la incitación al odio. La reforma añade en el apartado 3 la circunstancia agravante *“cuando los hechos se hubieran llevado a cabo a través de un medio de comunicación social, por medio de internet o mediante el uso de tecnologías de la información, de modo que, aquel se hiciera accesible a un elevado número de personas.”*

Por último, el apartado 5 del artículo 20 CE prohíbe el secuestro administrativo de la información:

*“5. Sólo podrá acordarse el secuestro de publicaciones, grabaciones y otros medios de información en virtud de resolución judicial.”*

Este mecanismo de garantía establece que sólo podrá restringirse esta libertad en virtud de resolución judicial. Cabe plantearse si tiene sentido entonces que la

---

<sup>58</sup> STS 297/2016, de 5 de mayo, FJ 3.

<sup>59</sup> STC 235/2007, de 7 de noviembre, FJ 4.

<sup>60</sup> STC 174/2006, de 5 de junio.

administración necesite una resolución judicial para retirar ciertos contenidos, pero encomiende la tarea de moderar contenidos a plataformas privadas.

## **5.2. El derecho al honor, a la intimidad y a la propia imagen**

El derecho al honor, intimidad y propia imagen forman parte de los derechos de la personalidad ligados a la dignidad humana. Son bienes jurídicos y sociales constituidos como derechos subjetivos y limitados por los usos sociales.<sup>61</sup> Están reconocidos como derechos fundamentales por la Constitución Española en el Título I, Capítulo II, Sección 1ª en su artículo 18, por lo tanto, la Constitución les confiere una especial protección.

Los titulares de estos derechos son las personas físicas, las personas jurídicas y las entidades sin personalidad jurídica. Estos últimos pueden ser grupos étnicos o religiosos, entre otros, los cuales sufren vulneraciones en sus derechos de manera colectiva, por lo que se interpreta que cuentan con la capacidad de defenderse de una manera común, con razón de la personalidad que tienen, ya sea, por ejemplo, histórica o religiosa<sup>62</sup>.

En el ámbito del que se ocupa este trabajo tiene especial relevancia la previsión del constituyente de 1978 en el apartado 4 del art. 18CE ya establece que: *“La ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos.”*

El derecho al honor, intimidad y propia imagen son derechos independientes y autónomos<sup>63</sup>:

---

<sup>61</sup> Encabo Vera, M.A. (2012). *Derechos de la personalidad*. Marcial Pons, Ediciones Jurídicas y Sociales.

<sup>62</sup> Grimalt Servera, P., *La Protección Civil de los derechos al Honor, a la Intimidad y a la Propia Imagen*, Iustel, p. 43-29.

<sup>63</sup> Se desarrolla así en la STS de 17 de diciembre de 1997:

### 5.2.1. El derecho al honor

El honor es un concepto indeterminado, si no se delimita puede tener mucho alcance y es difícil definirlo con exactitud. A esto se le suma el hecho de que la interpretación de su alcance varía en función del contexto social, la presencia pública de la persona, la colisión con otros derechos fundamentales, así como otras circunstancias. Así lo establece la STC 180/1999, de 11 de octubre, que expresa que el concepto de honor “*depende de las normas, valores e ideas sociales en cada momento*”. De forma que expresiones que hace siglos no afectarían negativamente al honor de un sujeto actualmente podrían hacerlo y viceversa.

En la STC 49/2001, de 26 de febrero, el TC define este derecho como aquel que existe para que otro no convenga de una forma negativa la opinión que otras personas tienen sobre el sujeto titular de este derecho, siendo su bien jurídico protegido el aprecio social y la buena fama, entre otros.

Por otra parte, la STC 9/2007, de 15 de enero, en su FJ 3º estableció que este derecho al honor se declara como honra, siendo parte de la moral de la persona y actuando directamente en la reputación del mismo, vista esta como la opinión o estima que tiene la sociedad frente a él como individuo<sup>64</sup>. En esta misma sentencia el TC entendió que dentro del concepto de honor se incluía el prestigio profesional que puede tener un impacto muy intenso en el caso de las personas jurídicas.

---

«Los derechos de personalidad son derechos subjetivos que recaen sobre aspectos o manifestaciones inherentes a la persona, como ser humano, y no constituyen un solo derecho con varios aspectos «ius in se ipsum» sino un conjunto de derechos; entre ellos los del honor, intimidad e imagen; no se trata de un derecho tricéfalo sino de tres derechos».

<sup>64</sup> Contreras Navidad, S., *La Protección del Honor, la Intimidad y la Propia Imagen en Internet*, Aranzadi, p. 20.

### 5.2.2. *Derecho a la intimidad*

El concepto de intimidad tampoco cuenta con una definición en la ley por lo que para su delimitación se puede acudir al diccionario panhispánico del español jurídico, según el cual es *“el derecho a disfrutar de un ámbito propio y reservado para desarrollar una vida personal y familiar plena y libre, excluido tanto del conocimiento como de las intromisiones de terceros”*. Por su parte, la jurisprudencia ha ido definiendo el concepto de intimidad como un espacio propio, ya sea físico o mental, que se encuentra apartado del resto de las personas y busca ser protegido. Fue la STC 196/2004, de 15 de noviembre (Rec. 1322/2000), la que marcó que este derecho otorgaba a su titular la posibilidad de evitar que terceros o poderes públicos vulneraran este mismo derecho fundamental, dando así una facultad negativa.

### 5.2.3. *Derecho a la propia imagen*

La propia imagen es todo aquello que deja visible a la persona de cara a la sociedad, su aspecto exterior. Se reconoce el derecho a decidir quién puede o no representarla o usarla y transmitirla de cara al público. El TC especifica el contenido de su protección al impedir que una persona obtenga y haga uso de la imagen sin autorización, incluso aunque la finalidad de la obtención de la imagen no sea negativa.<sup>65</sup>

Partiendo del análisis de la jurisprudencia, podría discutirse hasta qué punto la comunicación a través de Internet ha provocado un cambio cultural en el contenido

---

<sup>65</sup> STC 81/2001, de 26 de marzo (Rec. 922/1998); STC 139/2001, de 18 de junio (Rec. 4824/1997); y STC 83/2002, de 24 de abril

protegido de estos derechos<sup>66</sup>, o si se pueden definir límites específicos a la libertad de expresión o de información cuando se ejerza en la red<sup>67</sup>.

### **5.3. Colisión de derechos fundamentales en internet**

Estos derechos no tienen un carácter absoluto. De hecho, no es infrecuente la colisión entre el derecho a la libertad de expresión y de información y el derecho al honor, la intimidad y la propia imagen. De tal forma que el derecho al honor o la intimidad de una persona puede suponer un límite al ejercicio de la libertad de expresión de otra.

Además, del mismo modo que se tienen en cuenta las circunstancias en el espacio analógico, también deberán tenerse en cuenta las características específicas de la red o la cuenta en la que se ha publicado el contenido. No es lo mismo incurrir en una vulneración del derecho al honor o la intimidad en una cuenta individual de una red social que en una cuenta con muchos más seguidores. Esta última tiene un mayor alcance, influencia y por tanto puede ocasionar un mayor perjuicio como apunta María Luisa Balaguer Callejón en su voto particular en la STC 93/2021, de 10 de mayo de 2021 (sobre los mensajes difundidos a través de una red social en los que tilda de asesino y se expresa alivio por la muerte de un torero en el curso de la lidia).

De acuerdo con la jurisprudencia del TEDH, para determinar la legalidad de las limitaciones a la libertad de expresión ha de realizarse un examen sobre su necesidad. Esto se ha llevado a cabo mediante el test de Estrasburgo que está compuesto por un triple análisis:

---

<sup>66</sup> Boix Palop, A. 2016. *La construcción de los límites a la libertad de expresión en las redes sociales*. Revista de Estudios Políticos

<sup>67</sup> De Miguel Bárcena, J. 2016. *Las transformaciones del derecho de la información en el contexto del ciberperiodismo*. Revista de Estudios Políticos

- Previsión legal: Se debe comprobar que la limitación a la libertad de expresión esté prevista legalmente.
- Objetivo legítimo: Se debe comprobar que la limitación sea necesaria para perseguir un fin legítimo de los previsto en el apartado 2 del art.10 CEDH:

*“la seguridad nacional, la integridad territorial o la seguridad pública, la defensa del orden y la prevención del delito, la protección de la salud o de la moral, la protección de la reputación o de los derechos ajenos, para impedir la divulgación de informaciones confidenciales o para garantizar la autoridad y la imparcialidad del poder judicial”*

- Necesidad en una sociedad democrática: Que la limitación sea realmente necesaria implica comprobar que sea proporcional al objetivo pretendido considerando los recursos de los que se disponen y no habiendo medidas menos restrictivas de la libertad de expresión que persigan el mismo fin<sup>68</sup>.

Es decir, el tribunal debe interpretar no solamente el contenido textual de los mensajes, sino también el objetivo que se persigue con ellos y el contexto en el que se realizan, que incluye considerar quién emite los mensajes, por qué medio y qué alcance tienen.

#### 5.3.1. El caso *Delfi AS vs Estonia*

El Tribunal Europeo de Derechos Humanos en el asunto *Delfi AS v. Estonia* estableció que para determinar el cumplimiento de la diligencia debida deben tenerse en cuenta el contexto de los comentarios, las personas implicadas, las medidas aplicadas

---

<sup>68</sup> MARTÍN HERRERA, D (Coordinador). QUESADA ALCALÁ, C (Autor); *La libertad de expresión desde un enfoque global y transversal en la era de los objetivos de desarrollo sostenible*, Thomson Reuters Aranzadi, 1ª ed., febrero 2022 pp. 7.

por la plataforma para impedir o eliminar los comentarios difamatorios y, si fuera el caso, la responsabilidad de los autores reales de los comentarios<sup>69</sup>

El supuesto de hecho en este caso fue que el titular de un portal web fue declarado civilmente responsable, según los tribunales de Estonia, por los comentarios difamatorios publicados por terceros en su plataforma. La página era un famoso portal de noticias en Estonia, en el que se había habilitado un apartado de comentarios que debían pasar un filtro. Este filtro consistía en un sistema automatizado que impedía la publicación de comentarios con ciertas palabras consideradas inapropiadas. Además, cualquier usuario podía marcar cualquier comentario para ser eliminado de forma automática, o bien ponerse en contacto con la dirección de la página web para denunciar un mensaje ofensivo.

La responsabilidad civil del titular del portal de noticias se derivaba de una vulneración del derecho al honor por los comentarios injuriosos de algunos usuarios contra una importante empresa de transporte marítimo. Dicha empresa solicitó al portal Delfi.ee la retirada de los comentarios, así como la compensación civil por daños y perjuicios. Delfi.ee retiró los comentarios, pero se negó al pago. La sentencia de los tribunales estonios fue recurrida ante el TEDH alegando haberse vulnerado su derecho a la libertad de expresión. El TEDH ratifica la decisión de los Tribunales de Estonia resolviendo que no se había vulnerado el artículo décimo del Convenio Europeo de Derechos Humanos (CEDH) y que las medidas interpuestas por el titular del portal eran insuficientes.

Para el TEDH resultó evidente que los mensajes eran difamatorios por lo que no quedan amparados por el art.10 CEDH, se trata aquí de determinar el grado de responsabilidad de la plataforma.

---

<sup>69</sup> STEDH de 16 de junio de 2015, asunto Delfi AS v. Estonia

En relación con las circunstancias en las que se difundieron los mensajes, el TEDH destacó ciertos aspectos:

El primero de ellos es que los usuarios no podían eliminar los mensajes que habían publicado, sino que esta tarea correspondía al titular de la página web. Lo cual implica que Delfi AS tenía control sobre los mensajes, aunque no los escribiera ella misma.

Por otra parte, la plataforma no exigió a los usuarios que se identificasen para poder publicar mensajes por lo que los perjudicados sólo podían ejercitar acciones sobre la persona que conocían, es decir, la titular de la web, Delfi AS.

Después, las medidas para evitar la publicación de contenido difamatorio de la página web eran insuficientes ya que estos mensajes pasaron el filtrado automático. Tampoco fue suficiente la respuesta de la página web ante la petición de que se eliminaran los mensajes ya que estos quedaron expuestos durante seis meses desde su publicación.

Se trata de una sentencia muy relevante de cara al enjuiciamiento de este tipo de situaciones en otros Estados Miembros del CEDH, como España.

Según el art. 20.5 CE la competencia para determinar dónde se encuentra el límite pertenece al poder judicial. Así, la STS (Sala de lo Contencioso-Administrativo, Sección 4.ª), 1231/2022, de 3 de octubre de 2022 incluyó a las páginas web dentro de la categoría de “otros medios de información”. Como establece la propia sentencia, esta previsión seguramente se deba a una especial protección del constituyente frente a prácticas abusivas que se han dado en el pasado. De esta forma, las grandes plataformas no estarían legitimadas para determinar los límites de la libertad de expresión.

No obstante, no parece viable que cada cuestión de este tipo sea resuelta por jueces y tribunales debido a la inmensa cantidad de comentarios que se producen cada día que se encuentran en esta “zona gris”. A modo de ejemplo, el número de denuncias internas que recibió X (anteriormente Twitter) en el segundo semestre de 2024 fue de 56 millones por discurso de odio y casi 9 millones por contenido privado. Estos son sólo dos de las diez áreas de denuncia que tiene X entre las cuales están el contenido violento, contenido de abuso y acoso, pornografía infantil, contenido relacionado con el suicidio y la autolesión o la desnudez no consentida entre otros. Las denuncias totales suman más de 181 millones en seis meses.<sup>70</sup>

Ante esta situación, cabe plantearse si sería una buena opción delegar la decisión de eliminar o no un contenido en las plataformas antes de que exista una resolución judicial firme. De manera que, si la plataforma decide eliminarlo por considerar que vulnera el derecho al honor del perjudicado, el usuario cuyo contenido ha sido eliminado pudiera denunciar la acción de la plataforma ante los tribunales. O, sensu contrario, si la plataforma decide no tomar acción pudiera el que se considere ofendido o perjudicado denunciarlo judicialmente.

En el primero de los casos, el riesgo que se corre es que ciertos usuarios sean silenciados injustamente, aunque sea de forma temporal (hasta que hubiera una sentencia judicial a su favor), pudiendo esto tener un impacto negativo en la finalidad que se pretendía perseguir en el ejercicio de la libertad de expresión. Por otra parte, en el segundo caso, se corre el riesgo de que si eventualmente se dicta una sentencia a favor del que considera vulnerado su derecho al honor, intimidad o propia imagen, cuando el contenido ofensivo sea eliminado puede ser demasiado tarde y el perjuicio a su bien jurídico protegido ya haya sido ocasionado.

---

<sup>70</sup> [https://transparency.x.com/content/dam/transparency-twitter/2025/x-global-transparency-report\\_h2\\_2024.pdf](https://transparency.x.com/content/dam/transparency-twitter/2025/x-global-transparency-report_h2_2024.pdf)

En cualquier caso, los perjuicios ocasionados por la plataforma que actúe erróneamente cuando los perjudicados no decidieran denunciar serían, naturalmente, mucho mayores.

En ámbitos como el contenido violento, la pornografía infantil o contenido que vulnere derechos de propiedad intelectual o marcas comerciales registradas en los cuales parece más claro denotar su carácter ilícito, por consiguiente, existe menos margen de error. Por lo tanto, parece una buena política hacer responsables a las plataformas de retirar el contenido en cuanto tengan conocimiento efectivo de su existencia. Además, en estos ámbitos resulta más difícil vulnerar la libertad de expresión o información de los usuarios implicados.

El problema que aquí se plantea, es si se debe conceder o incluso encomendar a las plataformas la facultad de moderación respecto a aquellos comentarios que aunque sean susceptibles de considerarse discurso de odio, su eliminación vulnere la libertad de expresión de los usuarios. Cada vez son más los contenidos que se encuentran en esta zona gris, por eso, debemos preguntarnos hasta qué punto deben ser las plataformas las que decidan qué contenido sí y qué contenido no y qué riesgos podemos correr con ello. Es decir, el planteamiento es determinar el margen que deberían tener las plataformas digitales para moderar contenidos.

#### **5.4. La eficacia entre particulares de los derechos fundamentales**

Mientras en EEUU se protege la libertad editorial de las plataformas al amparo de la Primera Enmienda, en Europa el poder de moderación debe estar limitado por los derechos fundamentales de los usuarios, aunque no existe una legislación específica que lo establezca.

En España se reconoce la eficacia entre particulares de los derechos fundamentales<sup>71</sup>. No obstante, sin legislación específica sobre la materia, las decisiones judiciales no siempre han sido coherentes con la tesis de la eficacia directa<sup>72</sup>.

Por ejemplo, la Sentencia de la Audiencia Provincial de Mallorca 124/2020, de 26 de marzo, para respaldar su decisión utiliza argumentos en línea con la jurisprudencia de EE.UU., concluye que el caso se trata de un contrato entre particulares, por lo que no puede invocarse una vulneración de libertad de expresión. El tribunal balear respaldó la decisión de Twitter de suspender la cuenta de un usuario que publicó comentarios homófobos de manera reiterada. Argumentó que esta decisión supone únicamente la limitación de utilizar un medio específico para ejercer la libertad de expresión, pero que puede hacerlo por cualquier otro medio. La sentencia, asimismo, rechazó que hubiera un abuso contractual en la cláusula que permite a Twitter eliminar cuentas que infrinjan sus normas. Se destacó que esta disposición no genera un desequilibrio entre las partes, ya que el usuario también tiene la libertad de cancelar su cuenta en cualquier momento y sin necesidad de justificarlo.

Lo interesante de esta sentencia no es la decisión de respaldar a Twitter sino el fundamento de que la decisión estaba basada en un contrato de derecho privado en el cual no existían, según el tribunal, cláusulas abusivas.

Los derechos fundamentales son derechos públicos subjetivos, es decir, se conciben, en principio, como derechos de los ciudadanos frente al Estado. Según la concepción tradicional de estos derechos, estos sólo vinculan al Estado<sup>73</sup>. En el origen de los Estados liberales, las Constituciones optan por configurar los derechos naturales como límites frente al poder público. Por otra parte, los derechos como la vida o la

---

<sup>71</sup> SÁENZ ROYO, E. *Op. Cit.* p. 16

<sup>72</sup> SÁENZ ROYO, E. *Op. Cit.* p. 16

<sup>73</sup> BILBAO UBILLOS, JM<sup>a</sup>, *La consolidación dogmática y jurisprudencial de la "drittwirkung": Una visión de conjunto*. AFDUAM. p.49

libertad en las relaciones entre particulares son garantizados por las normas de Derecho privado. Esta teoría parte de la premisa de que las relaciones entre particulares no necesitan tutela porque son relaciones entre iguales.

No obstante, la realidad es otra, son muchos los contratos privados en los cuales las condiciones son impuestas por una de las partes, y la otra no tiene el mismo poder de negociación por lo que se ve obligada a aceptarlas. En estos casos el poder se ejerce, aunque sea de carácter privado, con formas de coacción análogas a las de los poderes públicos, por ejemplo, el despido de un trabajador. Es por esto, que la eficacia de los derechos fundamentales frente a terceros ha sido reconocida por el Tribunal Constitucional, principalmente en dos sectores: las relaciones laborales y las relaciones internas de las asociaciones<sup>74</sup>. Por otra parte, la Constitución no es una norma autónoma entre los poderes públicos y los ciudadanos, sino que es la norma jurídica suprema y fuente principal de todo el ordenamiento jurídico, ejerce una influencia en el Derecho privado. En un Estado social deben ser materialmente constitucionales las normas y principios reguladoras de las relaciones sociales.

Los que se oponen a esta teoría, argumentan que la Constitución sólo delimita el ámbito de actuación del poder público estableciendo unos límites. Y que, por lo tanto, la Constitución no debe intervenir en la regulación de las relaciones jurídico-privadas ya que se pueden obtener los mismos resultados mediante la aplicación de las normas de Derecho civil. De lo contrario, el riesgo sería la degradación de la autonomía de la voluntad<sup>75</sup>

El Tribunal Constitucional español estableció, dando respuesta al dilema sobre la eficacia que tienen estos derechos en las relaciones privadas, que el derecho constitucional y el derecho privado no pueden concebirse como compartimentos

---

<sup>74</sup> BILBAO UBILLOS, JM<sup>a</sup>, *Op. Cit.* p. 51

<sup>75</sup> BILBAO UBILLOS, JM<sup>a</sup>, *Op. Cit.* p. 53

estancos<sup>76</sup>. Y, por tanto, las normas privadas deben interpretarse conforme a la Constitución. Es decir, la jurisprudencia constitucional española se posicionó a favor de la tesis de la eficacia directa frente a los particulares. Esta tesis se sostiene en el art.9.1 CE que reconoce expresamente la sujeción de los ciudadanos a la Constitución, aunque la vinculación de los poderes públicos y de los ciudadanos a los derechos fundamentales de los ciudadanos es distinta. Los ciudadanos tienen un deber general negativo de abstenerse de vulnerar los derechos ajenos mientras que los poderes públicos además de esto, tienen un deber general positivo de realizar funciones dirigidas a garantizar el ejercicio de los derechos fundamentales reconocidos por la Constitución.

Esta línea jurisprudencial permitiría a los tribunales, aún sin legislación específica, aplicar el derecho fundamental de libertad de expresión frente a las grandes plataformas a través de cláusulas de derecho privado<sup>77</sup>. Y así, los tribunales civiles podrían controlar la razonabilidad y proporcionalidad de las políticas de moderación.

Sin perjuicio de esto, hay quienes defienden que es tarea del legislador concretar el alcance de la eficacia de los derechos fundamentales entre particulares. No obstante, la previsión legislativa no es un requisito necesario que se reconozca la eficacia en los tribunales. Otros defienden que son los jueces los que han de tener en cuenta los derechos fundamentales a la hora de interpretar las normas de Derecho privado. En cualquiera de los dos casos, estaríamos ante un sistema de eficacia mediata o indirecta, en el que para que los derechos fundamentales sean eficaces en las relaciones privadas un órgano del Estado, ya sea el poder judicial o legislativo, ha de intervenir.

La solución de los Estados Unidos a este dilema es la *state-action doctrine* ya mencionada en el apartado tres de este trabajo. Esta doctrina se basa en que las

---

<sup>76</sup> BILBAO UBILLOS, JM<sup>a</sup>, *Op. Cit.* p. 51

<sup>77</sup> Están prohibidas las cláusulas que impliquen la renuncia de derechos art. 6.2 del Código Civil y el art. 82 del 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.

garantías constitucionales sólo suponen un límite, en principio, para la actuación del Estado. Ahora bien, esta doctrina ha sido utilizada en supuestos de ejercicio de una función pública y en aquellos en los que se detecta una “conexión” o “implicación” estatal en una actuación aparentemente privada<sup>78</sup>. Es aquí donde podríamos encontrar un encaje en las imposiciones de determinadas obligaciones a las plataformas digitales que pudieran tener como efecto colateral la vulneración de la libertad de expresión. Así, tendríamos una actuación aparentemente privada, ya que serían las plataformas las que moderan los contenidos, pero bajo un mandato legal impuesto por los poderes públicos.

## 6. CONCLUSIONES

El sistema de funcionamiento de Internet es muy complejo ya que existen multitud de actores con diferentes grados de intervención sobre el contenido que circula por la red. Además, estos actores intervienen en diferentes capas o niveles, lo cual supone, a su vez, una dificultad añadida a la regulación del ciberespacio.

Por otra parte, el alcance y los límites del ejercicio del derecho a la libertad de expresión, si bien teóricamente son los mismos que en el plano analógico, en la realidad están condicionados a las circunstancias que caracterizan Internet. También están condicionados por la regulación que recientemente se ha introducido, en el ámbito de la Unión Europea, la Digital Service Act que entró en vigor en 2022.

Por otro lado, las redes sociales, en las cuales prolifera la difusión masiva de discursos de odio, se han reservado contractualmente el poder de aplicar políticas de moderación. Las políticas de moderación de contenidos pueden ocultar una vulneración de la libertad de expresión, esto se acentúa cuando los incentivos que tienen las redes

---

<sup>78</sup> BILBAO UBILLOS, JM<sup>a</sup>, *Op. Cit.* p. 56

sociales para moderar no están alineados con fines constitucionales sino de beneficio propio.

Entonces nos encontramos ante la necesidad de elegir en qué punto nos queremos situar dentro del triángulo cuyos vértices son:

- 1- Prohibir la moderación de contenidos para proteger el derecho a la libertad de expresión de los usuarios. El riesgo de esto sería convertir las redes sociales en una selva llena de contenidos ilícitos cuya eliminación sólo pudiera llevarse a cabo previa resolución judicial que lo autorice.
- 2- Responsabilizar a las grandes plataformas digitales de los contenidos que circulen por ellas. Esto supondría la imposición de una obligación de vigilancia que, si bien es en parte utópica, llevaría a las plataformas a pecar en exceso de ser restrictivas, ya que, al hacérseles responsables a ellas, eliminarían todo el contenido sobre el cuál tuvieran la mínima duda de que pueda ser ilícito mediante filtrados automáticos masivos. Esto supondría, sin duda alguna, una grave vulneración del derecho a la libertad de expresión.
- 3- Permitir la moderación de contenidos sin responsabilizar a las plataformas digitales por los contenidos que circulan por ellas. Lo cual implica la ausencia de responsabilidades ante ilícitos (ex post) y la ausencia de responsabilidades previas de transparencia (ex ante). En este caso, los criterios detrás de las políticas de moderación de contenidos serían aquellos que maximizan el beneficio de las corporaciones. En ocasiones, podrían estar alineados con los criterios que benefician a la sociedad (por ejemplo, una red social puede perder beneficios si está llena de contenidos ilícitos ya que número de usuarios que la utilizan podría bajar), pero en muchas otras ocasiones no lo estarían. No incentivarían el pluralismo político, ya que fomentarían la ideología de la cual estén recibiendo financiación. Es decir, se vulneraría la libertad de expresión,

pero directamente por las redes sociales, en vez de por parte del Estado utilizando a las corporaciones como ente vicarial.

Resulta evidente que ninguno de estos tres escenarios resulta óptimo para nuestra sociedad. Por consiguiente, es tarea del legislador encontrar el punto óptimo donde se maximicen los derechos de los ciudadanos.

## **7. BIBLIOGRAFÍA**

### **LIBROS Y ARTÍCULOS DE REVISTA**

ARRESE IRIONDO, M. N., «La responsabilidad de los diarios digitales por los comentarios de sus usuarios», RVAP, 99, 2014, pp. 387

ARROYO AMAZUELAS, E.: La responsabilidad de los intermediarios en internet ¿puertos seguros a prueba de futuro?, Cuadernos de Derecho Transnacional, vol. 12, Nº 1, marzo 2020

BARATA MIR, J.: “Libertad de expresión, regulación y moderación privada de contenidos”, Teoría y Derecho, n. 32, 2022,

BARRERO ORTEGA, A.: “Responsabilidad de los intermediarios de internet en el derecho de la UE”, en Revista Española de Derecho Constitucional, 123, 2021, pp. 109 y ss.

BILBAO UBILLOS, JM<sup>a</sup>, La consolidación dogmática y jurisprudencial de la “drittwirkung”: Una visión de conjunto. AFDUAM. P.43-74

BOIX PALOP, A.: “La construcción de los límites a la libertad de expresión en las redes sociales, Revista de Estudios Políticos, nº173, 2016

CONTRERAS NAVIDAD, S., La Protección del Honor, la Intimidad y la Propia Imagen en Internet, Aranzadi SA, Cizur Menor (Navarra), 2012, p. 20.

DE MIGUEL BÁRCENA, Josu (2016): «Las transformaciones del derecho de la información en el con texto del ciberperiodismo», Revista de estudios políticos, 173, 141-168.

MORETÓN TOQUERO, M. ARANZAZU: ENTRE SCILLA Y CARIBDIS: LOS INTERMEDIARIOS DIGITALES Y LA MODERACIÓN DE CONTENIDOS Rivista N°: 4/2023 Associazione Italiana Costituzionalisti

GUICHOT, Emilio (2019): «El reconocimiento y desarrollo del derecho al olvido en el derecho europeo y español», Revista de Administración Pública, 209, pp. 45-92.

GRIMALT SERVERA, P., La Protección Civil de los derechos al Honor, a la Intimidad y a la Propia Imagen, Iustel, Madrid, 2007, p. 43-49.

SUNSTEIN, Cass S. (2018): #Republic: Divided Democracy in the Age of social media, Princeton: Princeton University Press.

TERUEL LOZANO, G. M.: “Libertad de expresión, censura y pluralismo en las redes sociales: algoritmos y el nuevo paradigma regulatorio europeo”, en Derecho público de la inteligencia artificial, Balaguer Callejón, E. y Cotino Hueso, L. coords., Fundación Manuel Giménez Abad, 2023, pp. 197 y ss.

PERNAS ALONSO, JM.: La libertad de expresión en plataformas digitales amenazada en la UE: los casos Twitter y Telegram. International Journal of Digital Law. Pp 61-77.

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.

VALIENTE MARTÍNEZ, F. (2023). La libertad de expresión y las redes sociales: de la doctrina de los puertos seguros a la moderación de contenidos. Revista de Filosofía del Derecho y derechos humanos (48), 182- 202.

VÁZQUEZ ALONSO, V., «La censura “privada” de las grandes corporaciones digitales y el nuevo sistema de la libertad de expresión», TEORDER, 32, 2022. DOI: <https://doi.org/10.36151/td.2022.040>

VAZQUEZ ALONSO, V.: “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, enero-junio. 2020

VILLAVERDE MENÉNDEZ, I., Los poderes salvajes. Ciberespacio y responsabilidad por contenidos difamatorios, Marcial Pons, Madrid, 2020

## **JURISPRUDENCIA**

### **Del Tribunal Constitucional:**

STC 9/1981, de 31 de marzo.

STC 6/1988, de 21 de enero.

STC 180/1999, de 11 de octubre.

STC 187/1999, de 25 de octubre.

STC 49/2001, de 26 de febrero.

STC 81/2001, de 26 de marzo.

STC 139/2001, de 18 de junio.

STC 83/2002, de 24 de abril.

STC 196/2004, de 15 de noviembre.

STC 1/2005, de 17 de enero.

STC 174/2006, de 5 de junio.

STC 9/2007, de 15 de enero.

STC 235/2007, de 7 de noviembre.

STC 93/2021, de 10 de mayo de 2021.

### **Del Tribunal Supremo:**

SSTS de 9 de diciembre de 2009 (recurso 914/2006), de 10 de febrero de 2011 (recurso 1953/2008) y de 4 de diciembre de 2012 (recurso 1626/2011).

STS de 7 de enero de 2014.

STS 297/2016, de 5 de mayo.

STS 735/2022, de 28 de febrero.

STS 1231/2022, de 3 de octubre.

**De Audiencias Provinciales:**

SAP 124/2020, de 26 de marzo.

**Del Tribunal Europeo de Derechos Humanos:**

STEDH de 16 de junio de 2015, asunto Delfi AS v. Estonia

## **LEGISLACIÓN**

Communications Decency Act of 1996 (CDA), (Ley de Decencia de las Comunicaciones, EEUU)

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

Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y de comercio electrónico.

Directiva 2019/790, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital.

Reglamento (UE) 2022/2065 de Servicios Digitales, o DSA.