

GARANTÍAS FRENTE A LA MODERACIÓN DE CONTENIDOS EN LA PROPUESTA DE REGLAMENTO ÚNICO DE SERVICIOS DIGITALES*

Mario Santisteban Galarza

Investigador predoctoral

Universidad del País Vasco (UPV/EHU)

Resumen: Con la propuesta de Reglamento relativo a un mercado único de servicios digitales la Comisión Europea pretende garantizar un entorno en línea transparente y seguro. Consciente de que los prestadores intermediarios tienen un importante rol en la forma en que los consumidores de productos digitales interactúan en el ciberespacio, la propuesta crea una serie de obligaciones que limitan sus poderes de supresión y remoción de contenidos compartidos por usuarios. La propuesta viene a establecer así una serie de garantías contra la tan sonada censura privada de las plataformas, que se espera que revertirá en la mayor protección de los derechos fundamentales en línea de los consumidores.

Palabras clave: Ley de servicios digitales, moderación de contenidos, derechos fundamentales, consumidores, Comisión Europea,

Title: Safeguards regarding Content Moderation in the proposal for a Digital Service Act.

Abstract: The European Commission's proposal for a Digital Service Act aims to ensure a transparent and secure online environment. Conscious that intermediary providers play an important role in the way consumers of digital products interact in cyberspace, the proposal creates a series of obligations that limit their powers of deletion and removal of content that users share. Thus, the proposal establishes a series of safeguards against the much-touted private censorship of platforms, which is expected to result in greater protection of consumers' fundamental rights online.

Key Words: Digital Service Act, content moderation, fundamental rights, consumers, European Commission.

SUMARIO. 1. INTRODUCCIÓN. 2. ASPECTOS GENERALES DE LA PROPUESTA. 3. LOS DEBERES DE LOS PRESTADORES INTERMEDIARIOS CUANDO REALIZAN TAREAS DE MODERACIÓN DE CONTENIDOS. 3.1. El deber de exponer el motivo de la remoción.

* Este trabajo es posible gracias a la financiación de la Universidad del País Vasco (UPV/EHU) derivada de la convocatoria PIF 2020.

3.2. El deber de proporcionar un procedimiento interno de revisión. 3.3. Deberes de transparencia. 4. CONCLUSIONES.

1. INTRODUCCIÓN

El 15 de diciembre de 2020, la Comisión Europea dio a conocer su propuesta¹ de Reglamento del Parlamento Europeo y del Consejo relativo a un mercado único de servicios digitales, conocida como Ley de Servicios Digitales (en lo que sigue PLSD) y por el que se modifica la Directiva de Comercio Electrónico² (en lo que sigue DCE). La propuesta se enmarca en un "paquete legislativo", que abarca tanto la PLSD, como otras normas que pretenden incidir en la regulación de las plataformas en línea, como la propuesta de Reglamento sobre mercados disputables y equitativos en el sector digital (Ley de Mercados Digitales).

La nueva legislación no sustituirá, sino que complementará la DCE y otros actos legislativos y normas sectoriales que afectan a la lucha contra el contenido ilícito en línea³. En este punto de la negociación, parece consensuado que los principios fundamentales que introdujo la DCE deben mantenerse vigentes⁴. El Parlamento Europeo antes de conocer la propuesta de la Comisión ya señaló que dichos principios⁵ "siguen siendo válidos" y "deben utilizarse para elaborar propuestas futuras"⁶.

Si bien el "espíritu del texto" se mantiene, la propuesta no desconoce que desde la promulgación de la DCE han aparecido nuevos e innovadores modelos de negocio, como las redes sociales y los mercados en línea, que han permitido "comunicar información y acceder a ella, y efectuar transacciones de formas novedosas". Acepta también que el creciente uso de estos servicios entraña nuevos riesgos y desafíos (considerando primero de la PLSD).

¹ El texto de la propuesta puede encontrarse disponible en: file:///C:/Users/usuario/Desktop/CELEX_52020PC0825_ES_TXT.pdf (última consulta 3 de enero de 2022).

² 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, DOCE núm. 178, de 17 de julio de 2000.

³ Cabe citar aquí la Directiva 2019/790 sobre los derechos de autor y derechos afines en el mercado único digital, el Reglamento 2021/784 sobre la lucha contra la difusión de contenidos terroristas en línea, la Recomendación 2018/334 sobre medidas para combatir eficazmente los contenidos ilícitos en línea y otras medidas como el Código de Buenas prácticas en materia de desinformación o el Código de conducta para frenar el discurso ilegal de odio.

⁴ En el momento en que se escriben estas líneas el Consejo ya ha consensuado una posición general sobre la norma y diversos comités del Parlamento Europeo han emitido sus informes sobre la propuesta. El inicio de las negociaciones entre ambos órganos se espera para el año 2022.

⁵ Siguiendo así la percepción del "éxito" que supuso la Directiva 2000/31/CE la propuesta de la Comisión no altera aspectos fundamentales del modelo de la Directiva como la "cláusula del mercado interior" (art 3.1), las exenciones de responsabilidad de los proveedores de servicios de la sociedad de información que tengan el carácter de intermediarios (art 12 a 14) y la prohibición de imponer una obligación general de monitorización a los prestadores intermediarios (art 15).

⁶ Resolución del Parlamento Europeo de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre la Ley de servicios digitales: una mejora del funcionamiento del mercado único (2020/2018) apdo B. Disponible en: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0272_ES.html#title1 (última consulta 17 de diciembre de 2021).

Como ya señaló la Comisión Europea ⁷, los servicios de la sociedad de la información se han convertido en una parte importante de la economía de la Unión y de la vida cotidiana de sus ciudadanos. Estos servicios son objeto de frecuente abuso por terceros, que pretenden realizar actividades ilícitas, como la difusión de propaganda terrorista, incitación al odio o infracciones de la legislación sobre protección de los consumidores. A la vista de este problema, se pretende favorecer un “entorno en línea seguro”, tratando de que lo que sea ilegal en el mundo *offline* lo sea también en el mundo *online*⁸. Para lograr ese objetivo, el texto de la Comisión crea una serie de obligaciones de carácter gradual y asimétrico, que afectan a los prestadores de servicios intermediados en función de la naturaleza de su servicio y tamaño.

Dichas obligaciones tratan de ponderar varios intereses jurídicos contrapuestos “como la salvaguardia de los intereses legítimos de los destinatarios del servicio, la lucha contra las prácticas ilícitas y la protección de los derechos fundamentales en línea” (considerando treintaicincoavo de la PLSD). Refiriéndonos a este último objetivo, la PLSD pretende evitar que las plataformas acaben afectando con sus prácticas de moderación de contenidos a la circulación de contenido lícito. De esta forma, siguiendo las indicaciones del Parlamento Europeo, el texto trata de que exista una “mayor equidad, transparencia y rendición de cuentas en los procesos de moderación de contenidos de los servicios digitales, de modo que se asegure el respeto de los derechos fundamentales”⁹.

La moderación de contenidos se define como aquellas “actividades realizadas por los prestadores de servicios intermediarios destinadas a detectar, identificar y actuar contra contenidos ilícitos o información incompatible con sus condiciones, que los destinatarios del servicio hayan proporcionado”, y que pueden consistir tanto en la remoción de los contenidos ilícitos o la suspensión de la cuenta del usuario infractor. Frente a ella, se establecen una serie de obligaciones cuyos destinatarios son prestadores de servicios de la sociedad de la información que llevan a cabo tareas de moderación.

Caracterizada la red actual por el intercambio de información proporcionada por usuarios a través de los interfaces diseñados por grandes intermediarios¹⁰, las decisiones tomadas por dichos intermediarios respecto a qué es admisible en sus foros tienen una no desdeñable relevancia. De ahí que paulatinamente haya crecido

⁷ Recomendación 2018/334 de la Comisión de 1 de marzo de 2018 sobre medidas para combatir eficazmente los contenidos ilícitos en línea.

⁸ European Council, “What is illegal offline should be illegal online: Council agrees position on the Digital Services Act”, 25 de noviembre de 2021. Disponible en: <https://www.consilium.europa.eu/en/press/press-releases/2021/11/25/what-is-illegal-offline-should-be-illegal-online-council-agrees-on-position-on-the-digital-services-act/> (última consulta 21 de diciembre de 2021).

⁹ Resolución del Parlamento Europeo de 20 de octubre de 2020 “Recomendaciones destinadas a la Comisión sobre la Ley de servicios digitales: una mejora del funcionamiento del mercado único (2020/2018(INL))”. Disponible en: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0272_ES.html#title1 (última consulta 17 de diciembre de 2021).

¹⁰ Así, la llamada Web.2.0 se caracteriza por la proliferación de ciertos servicios “que han cambiado la forma en la que los usuarios interactúan en la red, pasando de un comportamiento pasivo (en el que los usuarios consultan la información disponible en la web o la comparten) a un papel activo en el que los propios usuarios generan contenidos para su difusión pública” DE MIGUEL ASENSO, P M., *Derecho privado de Internet*, Civitas, Madrid, 2015, p. 36.

la preocupación social por la llamada "censura privada"¹¹ de las grandes plataformas, en aplicación de sus términos y condiciones del servicio. Acontecimientos recientes¹² parecen sugerir que las grandes plataformas no son neutrales a la hora de actuar como vigilantes del entorno digital y una regulación de esta actividad puede ser beneficiosa para los consumidores de servicios digitales. Esa preocupación social no se ha plasmado hasta el momento a nivel de la Unión, ni tampoco en el de los Estados miembros, en una normativa que proteja los derechos de los consumidores de servicios digitales frente a estas prácticas¹³. La PLSD pretende ofrecer una protección jurídica en esta materia, colmando el vacío normativo preexistente.

2. ASPECTOS GENERALES DE LA PROPUESTA

El Reglamento único de servicios digitales se propone la difícil empresa de crear las condiciones para un entorno en línea seguro, así como contribuir al correcto funcionamiento del mercado interior de servicios intermediarios¹⁴, por lo que aborda en su articulado muchas cuestiones sobre la regulación de los intermediarios en línea de cuyo estudio no puede ocuparse esta modesta aproximación¹⁵. Aquí, únicamente nos referimos a ciertos aspectos generales, con el fin de sentar unas nociones básicas de la norma sin las cuales sería difícil adentrarnos en las garantías que prevé frente a la moderación de contenidos.

La Propuesta se divide en cuatro capítulos. El primero de ellos nos ofrece ciertas nociones fundamentales, entre las que se encuentra el ámbito de aplicación de la norma. Este se circunscribe subjetiva y territorialmente "a aquellos servicios intermediarios¹⁶ prestados a destinatarios del servicio que tengan su lugar de establecimiento o residencia en la Unión, con independencia del lugar de establecimiento de los prestadores de dichos servicios".

¹¹ Evidentemente hablamos aquí de censura en un sentido distinto al término constitucional. Véase GARCÍA MORALES, M. J., "La prohibición de censura en la era digital" *Teoría y Realidad Constitucional*, Núm. 31, 2013, pp. 254-255.

¹² HORWITZ, J., "Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt", *New York Times*, 14 de septiembre de 2021. Disponible en: <https://www.nytimes.com/es/2021/10/26/espanol/los-papeles-de-facebook.html> (última consulta 17 de diciembre de 2021).

¹³ "Regarding moderation of content by platform operators, there are scarcely any specific legal provisions on the European level. Whereas the EU has enacted several specific directives or regulations aiming at fostering control of activities on platforms (such as copyright, anti-terrorism etc.), there are no provisions that assign to users' rights to upload content or have access to content on platforms. Neither the ECD nor the DSM-D provides such mandatory rights for users (even though Art. 17 (9) DSM-Directive (2019/790)³⁹ can be read as implicitly conferring rights to the user to upload non-infringing content); also, EU provisions regarding contract law such as the directive on unfair terms do not refer to standard terms and conditions of platforms" European Parliamentary Research Service, "Digital Service Act. European added value assessment", octubre de 2020, p. 195.

¹⁴ La propuesta, al amparo del art 114 del Tratado de Funcionamiento de la Unión Europea pretende evitar la fragmentación legislativa en el mercado de servicios intermediarios. La norma se dicta teniendo en mente pues, ciertos actos legislativos de Estados Miembros, como la Network Enforcement Act, que imponen obligaciones adicionales a los prestadores intermediarios en la lucha contra determinados ilícitos.

¹⁵ Para una aproximación más general a la propuesta y a la Ley de Mercados Digitales se recomienda la consulta de DE MIGUEL ASENSIO, P.A., "Servicios y mercados digitales: modernización del régimen de responsabilidad y nuevas obligaciones de los intermediarios" *La Ley Unión Europea*, Núm. 88, 2021.

¹⁶ Por servicio intermediario la propuesta identifica a los prestadores que ofertan servicios de mera transmisión, memoria tampón o almacenamiento de datos, identificándolos así con los tres tipos de servicios de la sociedad de la información a los que se refería la DCE a efectos de responsabilidad.

Asimismo, cuenta con un importante listado de definiciones (art 2). Entre ellas destacamos la noción de plataforma digital, como subtipo de prestador intermediario de hosting y que se caracteriza por almacenar y difundir al público información, es decir, “poner información a disposición de un número potencialmente ilimitado de terceros a petición del destinatario del servicio que ha facilitado dicha información”. También merece la pena resaltar el amplio concepto de contenido ilícito con el que trabaja el reglamento, “toda información que, por sí sola o en referencia a una actividad, incluida la venta de productos o la prestación de servicios, incumpla las leyes de la Unión o las leyes de un Estado miembro, sea cual sea el objeto o carácter concreto de esas leyes”¹⁷.

Su Capítulo II regula la responsabilidad de los prestadores de servicios intermediarios, sustituyendo los art. 12 a 15 de la DCE. Si bien la redacción de los citados preceptos no es muy diferente a la que encontramos en la DCE, se introducen previsiones adicionales que vienen a perfilar el régimen de responsabilidad de los prestadores intermediarios. El art. 6 de la propuesta contribuye a definir con mayor precisión la noción de conocimiento efectivo del ilícito a efectos de aplicación de las exenciones de responsabilidad, particularmente en relación con las actuaciones voluntarias de moderación que llevan a cabo los intermediarios. Por su parte, los artículos 8 y 9 delimitan los requisitos que deben de observar las órdenes de actuación contra contenidos ilícitos o de entrega de información que dicten las autoridades nacionales competentes y el modo de proceder de los prestadores intermediarios.

El grueso de la propuesta de la Comisión lo constituye el Capítulo III, que se rubrica como “obligaciones de diligencia debida para crear un entorno en línea transparente y seguro”. Este se divide en cinco secciones en las que la propuesta va elevando el número y alcance de las obligaciones de los prestadores intermediarios en función de dos factores: el tipo de prestador, y su tamaño.

Entre las previsiones que no abordaremos más detenidamente *infra* requiere especial mención el establecimiento de mecanismos de notificación y acción que permitan a los usuarios comunicar la presencia de contenidos ilícitos (art 14). El hecho de que la notificación al prestador a través de este mecanismo suponga que el intermediario adquiera “conocimiento efectivo” del ilícito viene a colmar un déficit fuertemente

¹⁷ El considerando duodécimo de la propuesta añade que “debe entenderse que dicho concepto se refiere a información, sea cual sea su forma, que sea de por sí ilícita en virtud de la legislación aplicable, como los delitos de incitación al odio o los contenidos terroristas y los contenidos discriminatorios ilícitos, o que tengan relación con actividades ilícitas, como el intercambio de imágenes que representen abusos sexuales de menores, la publicación ilícita no consentida de imágenes privadas, el acoso en línea, la venta de productos no conformes o falsificados, el uso no autorizado de material protegido por derechos de autor o actividades que conlleven infracciones de la legislación de protección de los consumidores. En este sentido, es irrelevante si la ilicitud de la información o actividad tiene su origen en el Derecho de la Unión o en legislación nacional coherente con el Derecho de la Unión y cuál es la naturaleza o materia precisa de la legislación en cuestión”.

criticado de la DCE¹⁸, y que revertirá en la seguridad jurídica de las relaciones entre profesionales y usuarios en el entorno digital¹⁹.

Además de los mecanismos que deben implementar para mejorar sus sistemas de moderación de contenidos, destacan varios deberes que se imponen a las plataformas en línea, como la obligación de suspender cuentas de usuarios que hagan un uso indebido de los servicios (art 20), la notificación a las autoridades competentes de ciertos hechos delictivos (art 21) o la adopción de medidas que favorezcan la transparencia de los servicios como la trazabilidad e identificación de los comerciantes que operan en ellos (art 22).

Más demandante es aún la propuesta con aquellos prestadores que tengan la consideración de grandes plataformas en línea, definidas como "aquellas plataformas en línea que presten sus servicios a un número medio mensual de destinatarios del servicio activos en la Unión igual o superior a cuarenta y cinco millones" (art 25.1). Estos deberes adicionales consisten en la gestión de los "riesgos sistémicos" que se derivan de su funcionamiento, que la propuesta identifica con la difusión de contenido ilícito a través de sus servicios, efectos negativos para el ejercicio de los derechos fundamentales o la manipulación de sus servicios con un "efecto negativo real o previsible sobre la protección de la salud pública, los menores, el discurso cívico o efectos reales o previsibles relacionados con procesos electorales y con la seguridad pública"²⁰. Dicha gestión consistirá en una evaluación anual de esos peligros en sus espacios (art 26), seguida de medidas de reducción del concreto riesgo (art 27) que en último término son supervisadas a través de una auditoría independiente (art 28).

Finalmente, el Capítulo IV de la propuesta establece una extensa lista de disposiciones que garanticen el cumplimiento de la norma. Los Estados miembros deberán de nombrar autoridades competentes para vigilar el cumplimiento del Reglamento, entre las cuales una de ellas deberá de tener la condición de coordinador de servicios digitales (art 38.1.) La propuesta otorga a dicha figura las competencias para exigir a los prestadores intermediarios información, realizar investigaciones, ordenar que cesen infracciones e imponer multas, entre otras funciones (art 41).

Estas capacidades de los coordinadores de servicios digitales se compaginan con unos amplios poderes que recibe la Comisión Europea para fiscalizar el cumplimiento del Reglamento por parte de las grandes plataformas en línea. La Comisión podrá iniciar solicitudes de información, realizar entrevistas e inspecciones, adoptar medidas provisionales en caso de riesgo de daños graves, y adoptar decisiones que constaten

¹⁸ Véase, por ejemplo, LÓPEZ RICHART, J., "Difamación en la Web 2.0 y responsabilidad civil de los prestadores de servicios de alojamiento", *Derecho Privado y Constitución*, Núm. 26, 2012, pp. 143-201.

¹⁹ A pesar de ello se ha criticado que es posible que no consiga evitar uno de los riesgos fundamentales que pretendía mitigar su implementación, como es que el intermediario no se encuentre en la delicada posición de decidir sobre si un contenido es lícito o no. DIGITALEUROPE, "Digital Service Act position paper", 31 de marzo de 2021, p. 8. Disponible en: <https://www.digitaleurope.org/resources/digital-services-act-position-paper/> (última consulta 27 de diciembre de 2021).

²⁰ La Comisión ha venido tratando de mitigar estos riesgos mediante la creación de códigos de buenas prácticas como el Código de conducta relativo a la lucha contra el discurso de odio en línea y el Código de buenas prácticas contra la desinformación.

el incumplimiento de las grandes plataformas, las cuales pueden llevar aparejadas multas.

Paralelamente, la PLSD crea una Junta de Servicios Digitales como grupo consultivo que contribuye a la aplicación coherente del Reglamento, a coordinar a la Comisión y a los coordinadores de servicios digitales sobre aspectos relacionados con el Reglamento, "y a asistir a los coordinadores de servicios digitales y a la Comisión en la supervisión de las plataformas en línea de muy gran tamaño" (art 47).

3. LOS DEBERES DE LOS PRESTADORES INTERMEDIARIOS CUANDO REALIZAN TAREAS DE MODERACIÓN DE CONTENIDOS

La PLSD no aborda frontalmente los posibles conflictos de derechos que puedan surgir entre las libertades que disponen los consumidores de servicios digitales y el ejercicio de la moderación de contenidos. La norma no determina cuando el resultado de las actividades de moderación (la eliminación de un comentario o la suspensión de una cuenta) es lícito. Por otra parte, tampoco cierra la puerta a la revisión judicial de dichas decisiones a la luz de derechos fundamentales como la libertad de expresión.

En el ámbito de la Unión Europea, puede decirse que tal posibilidad se encuentra más abierta si lo comparamos con el modelo norteamericano. Como es conocido, en este sistema la sección 230 de la *Communications Decency Act* apartado b) (CDA), exime de responsabilidad a los prestadores intermediarios por cualquier acción voluntaria tomada con buena fe que suponga la restricción del acceso o la disponibilidad de un contenido que el prestador intermediario considere controvertido²¹. Ello ha supuesto que los tribunales inadmitan las demandas de revisión judicial de decisiones de moderación de contenidos, tanto las fundamentadas en la lesión del derecho a la libertad de expresión como las que se basan en el incumplimiento contractual del intermediario al interrumpirse el servicio²².

El ordenamiento jurídico comunitario se ha inspirado en algunos de los principios que asentó la normativa norteamericana, estando las exenciones de responsabilidad reguladas en la DCE inspiradas en gran parte en las previstas en la CDA, y principalmente en la *Digital Millennium Copyright Act* (DMCA). Como decíamos, el régimen de responsabilidad de los intermediarios en el entorno digital respecto a los contenidos publicados por los destinatarios del servicio es similar tanto en el ámbito europeo como en el norteamericano. En ambos casos, se apuesta por una exención de responsabilidad que trata de evitar que los intermediarios se han considerados editores de la información (*Publisher*)²³. A pesar de ello, entre las múltiples

²¹ Nos referimos al apartado c (2), que señala "No provider or user of an interactive computer service shall be held liable on account of any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected" el texto legal puede encontrarse aquí: <https://www.law.cornell.edu/uscode/text/47/230>

²² Véase, GOLDMAN, E., "Online User Account Termination and 47 U.S.C. §230(c)(2)", *Santa Clara law Digital Commons*, Vol. 2(659), 2011, pp. 659-673 y VÁZQUEZ 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 Deusto*, Núm. 68/1, 2020, pp. 475-508.

²³ Si bien en el ámbito de la DCE esta se condiciona a que el intermediario no tenga conocimiento efectivo del ilícito y en caso de que lo obtuviese de la retirada del contenido, inspirándose en el modelo de la DMCA.

diferencias entre los sistemas, es destacable la inexistencia de una previsión similar en el modelo europeo al de la sección 230 apartado b), que blinde la moderación de contenidos.

En el sentido inverso, en el ordenamiento jurídico comunitario tampoco se esclarece si los intermediarios pueden llegar a ser responsables en último término de dichas decisiones. Es cierto que la DCE trata de tutelar la libertad de expresión a través de varios de sus preceptos, frenando la remoción de contenidos lícitos. Como ha aclarado el Tribunal de Justicia de la Unión Europea, el art 15 de la DCE, que establece la prohibición de imponer una obligación general de monitorización a los prestadores intermediarios²⁴, persigue evitar que se impongan sistemas de filtrado de contenidos que acaben afectando al derecho a comunicar y recibir información²⁵. No obstante, esta normativa no se posiciona en ningún momento sobre la posibilidad de la revisión judicial de una decisión de remoción de contenidos que parta de un intermediario²⁶. Esta omisión persiste en la PLSD, si bien con ciertos matices.

La posibilidad de acudir a la vía judicial como consecuencia de las decisiones de moderación es mencionada en el texto en varias ocasiones. No obstante, la cita a este medio de resolución de conflictos se realiza únicamente a efectos de categorizar los remedios que presenta el Reglamento como alternativos, y por tanto no excluyentes. No creo que pueda inferirse del texto de la norma que la revisión judicial de las decisiones de moderación deba de ser considerada como obligatoria.

Más esperanzador desde la perspectiva de la revisión judicial de estas decisiones es lo dispuesto en el art 12.2, que trata de integrar el respeto de los derechos fundamentales en la ejecución de los términos y condiciones contractuales de los prestadores intermediarios. Refiriéndose a las actividades de moderación de contenidos que partan de la aplicación de sus condiciones de uso este precepto demanda que "Los prestadores de servicios intermediarios actuarán de manera diligente, objetiva y proporcionada" teniendo en cuenta "la debida consideración de los derechos e intereses legítimos de todas las partes implicadas, incluidos los derechos fundamentales aplicables de los destinatarios del servicio consagrados en la Carta"²⁷.

²⁴ Es de especial interés resaltar que actualmente se ha presentado ante el TJUE un recurso de anulación sobre el art 17 de la Directiva 2019/790 (Asunto C-401/19), que cuestiona los mecanismos de filtrado de contenidos para prevenir infracciones de copyright que indirectamente introduce la normativa a la luz del derecho fundamental a la libertad de expresión. Véase la opinión del abogado general aquí: <https://curia.europa.eu/juris/liste.jsf?language=es&td=ALL&num=C-401/19>

²⁵ STJUE de 24 de noviembre de 2011, asunto C-70/10 (Scarlet Extended SA), apdo.50.

²⁶ En España, la sentencia de 26 de marzo de 2019 de la Sección Tercera de la Audiencia Provincial de Palma de Mallorca se ha planteado el problema de la revisión judicial de una moderación de contenido. Resolución disponible en <https://www.poderjudicial.es/search/AN/openDocument/5e74eb9a2f3b070e/20200526>

²⁷ El apartado segundo parece así indicar que los prestadores intermediarios deben de tener en cuenta los derechos fundamentales de los destinatarios a la hora de aplicar sus condiciones, siendo un precepto en el que hipotéticamente podría ampararse un órgano jurisdiccional cuando tenga ocasión de revisar una decisión de moderación de contenidos que sea contraria a tales libertades.

Sin duda la salvaguarda de los derechos fundamentales en línea es un objetivo que la PLSD tiene muy presente²⁸. No obstante, parece evidente que la norma no busca frenar la moderación, señalando en qué supuestos esta es lícita o contraria a derechos subjetivos insertos en el patrimonio jurídico del usuario²⁹. La técnica elegida es más bien el reconocimiento de una serie de obligaciones que el prestador intermediario debe de observar, y que revierten en beneficio del usuario, permitiendo corregir decisiones erróneas o arbitrarias.

El presente apartado se centrará en analizar estos deberes. Como ya hemos tenido ocasión de señalar en el apartado anterior, en el Capítulo III, en el que se integran dichas obligaciones, se establecen otros preceptos que no suponen en puridad garantías frente a la moderación de contenidos. Es por ello por lo que, en vez de seguir la sistemática del Reglamento, aglutinaremos las garantías en tres grandes bloques que vienen a identificarse con los principios de Santa Clara³⁰. Creemos que esto facilitará la claridad de la exposición teniendo en cuenta que las garantías de la moderación se regulan junto a otras disposiciones que no serán objeto de este trabajo.

3.1. El deber de exponer el motivo de la remoción

Conforme al art. 15 de la propuesta el prestador de alojamiento de datos que efectúe la retirada del contenido “comunicará la decisión al destinatario del servicio, a más tardar en el momento de la retirada o inhabilitación del acceso, y aportará una exposición clara y específica de los motivos de tal decisión”. Siguiendo el enfoque gradual y asimétrico de las obligaciones que sienta la PLSD, el deber de motivación se aplica a los prestadores intermediarios que tengan la consideración de prestadores de un servicio de alojamiento de datos (*hosting*), con independencia de su tamaño. Veamos el alcance de dicho deber de notificación.

²⁸ En su considerando tercero la PLSD señala que “Es esencial que los prestadores de servicios intermediarios se comporten de modo responsable y diligente para crear un entorno en línea seguro, predecible y confiable, y para que los ciudadanos de la Unión y otras personas puedan ejercer los derechos garantizados por la Carta de los Derechos Fundamentales de la Unión Europea (en lo sucesivo, «la Carta»), en particular la libertad de expresión e información y la libertad de empresa, así como el derecho a la no discriminación”

²⁹ Algún informe presentado ante la propuesta pretende ir más allá de este “enfoque formal”. Resaltamos aquí la opinión del Committee on Civil Liberties, Justice and Home Affairs que apunta “Online platforms’ terms and conditions shall respect fundamental rights and permit interferences with the free exchange of lawful information only where it is incompatible with the declared purpose of the service”. COMMITTEE ON CIVIL LIBERTIES, JUSTICE AND HOME AFFAIRS, “Opinion of the Committee on Civil Liberties, Justice and Home Affairs for the Committee on the Internal Market and Consumer on the proposal for a regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC”, 28 de julio de 2021, Disponible en: https://www.europarl.europa.eu/doceo/document/LIBE-AD-692898_EN.pdf

³⁰ Los principios de Santa Clara, esbozados como consecuencia del trabajo en una serie de conferencias en la universidad del mismo nombre fijan tres pilares fundamentales que deben sustentar la moderación privada de las plataformas. Estos son la transparencia de las actividades de moderación, el derecho a conocer los razonamientos detrás de las decisiones individualizadas de remoción de contenidos o suspensión de cuentas y el derecho a que existan procedimientos para recurrir dichas decisiones. La Declaración de los principios puede encontrarse aquí: <https://santaclaraprinciples.org/> (última consulta 22 de diciembre de 2021).

En primer lugar, el prestador deberá de notificar "si la decisión conlleva la retirada de la información, o la inhabilitación del acceso a la misma, y, cuando proceda, el ámbito territorial de la inhabilitación del acceso". Este último estribo es particularmente relevante en Internet, medio de comunicación de carácter transfronterizo en el que los ilícitos que se cometen pueden no estar reprimidos por la totalidad de los ordenamientos jurídicos. Así, es posible, que en base a una reclamación de una autoridad judicial o gubernamental el prestador decida remover el contenido únicamente en el Estado del que parte la reclamación; piénsese en el supuesto de un delito que ofenda los sentimientos religiosos o símbolos de un Estado.

Asimismo, la notificación de remoción deberá de contener los hechos y circunstancias en que se ha basado la adopción de la decisión, que incluirán, en su caso, si la decisión se ha adoptado en respuesta a un aviso enviado de conformidad con el artículo 14. En relación con este requisito, el prestador también debe de notificar "información sobre el uso de medios automatizados para adoptar la decisión, que incluirá si la decisión se ha adoptado al respecto de contenidos detectados o identificados utilizando medios automatizados". Esto puede ser de especial interés para el usuario perjudicado pues el uso de medios automatizados podría implicar, según el supuesto, la activación del art 22 del Reglamento General de Protección de Datos. Este protege al interesado contra decisiones basadas únicamente en el tratamiento automatizado, que produzcan efectos jurídicos en él o le afecten significativamente. El conocimiento de que ese tratamiento automatizado se está produciendo en el seno de la moderación puede ser relevante a la hora de explorar las posibilidades de recurso del usuario, más sabiendo que es posible que este deba basarse en su consentimiento previo³¹. En las condiciones y términos del servicio de se define con poca precisión cuando se utiliza la Inteligencia Artificial (IA) para detectar y remover contenidos, si bien es conocido que es utilizada por plataformas como Facebook o YouTube como filtro previo para enfrentarse a determinados ilícitos³².

Dentro de la notificación debe integrarse el fundamento legal o contractual en el que el prestador del servicio se ampara para remover o suspender la cuenta del usuario. Sobre el contenido considerado presuntamente ilícito, el prestador debe comunicar "una referencia al fundamento legal utilizado y explicaciones de por qué la información se considera contenido ilícito conforme a tal fundamento". En este punto cabe preguntarse cuál debe de ser la extensión de dicha justificación a la vista de las condiciones en las que se ejerce la labor moderadora de las plataformas. Cuando la remoción o suspensión se base en un aviso, ciertamente podrá procurarse al perjudicado la explicación de los motivos por los que la persona que inició el

³¹ El art 22.1 del RGPD señala que "Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar". No obstante, dicho tratamiento automatizado será posible cuando "sea necesario para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento; está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o se basa en el consentimiento explícito del interesado".

³² Véase GORWA, R., BINNS, R., KATZENBACH, C., "Algorithmic content moderation: Technical and political challenges in the automation of platform governance" *Big Data & Society*, 2020, pp. 1-15.

procedimiento consideró que el contenido compartido era ilícito. En la misma línea, cuando la remoción se deba a una orden judicial o emanada de otra autoridad competente, en esta ya se deberán de especificar los fundamentos legales en los que se basa (art. 8.2 a), lo que hará más sencillo satisfacer ese deber de información.

Fuera de esos supuestos, concretamente cuando la remoción se realice como consecuencia de una investigación propia del prestador, la referencia a la base legal presenta ciertos problemas. Dicha referencia, que conlleva un cierto análisis jurídico del contenido se integra con dificultad en las condiciones en las que se inserta la moderación, ejercida por trabajadores que frecuentemente tienen nula o escasa formación jurídica y que toman la decisión con cierta rapidez³³. A ello hay que sumarle que este deber de notificación se aplica a todos los prestadores de alojamiento de datos, incluidos aquellos que tengan la consideración de microempresas o pequeñas empresas en el sentido del anexo de la Recomendación 2003/361/CE.

En la práctica, la exigencia de este deber de fundamentación legal puede llevar a las plataformas a optar por remover el contenido en base a sus normas comunitarias, prescindiendo así de la referencia a un ordenamiento jurídico concreto. Materialmente, las normas comunitarias o términos del servicio de las plataformas traen de los ordenamientos jurídicos principios comunes, existiendo cierta similitud entre lo que se encuentra reprimido legalmente, ya en el ámbito penal (piénsese en el caso de la pornografía infantil o la incitación al odio) o civil (comentarios que no amparados por la libertad de expresión o información vulneran el derecho al honor). De ahí que sea sencillo eliminar cualquier tipo de contenido presuntamente ilegal en base a sus propios términos, pudiendo estos abarcar un amplio abanico de conductas reprimidas por los ordenamientos jurídicos de los Estados Miembros³⁴.

Finalmente, la propuesta señala que las exposiciones de motivos deberán de contener "información sobre las posibilidades de recurso disponibles para el destinatario del servicio al respecto de la decisión, en particular a través de los mecanismos internos de tramitación de reclamaciones, resolución extrajudicial de litigios y recurso judicial". A las primeras vías de reclamación nos referimos a posteriormente. Sobre la posibilidad del recurso judicial, no suele ser frecuente en la práctica de la remoción aludir a dicho mecanismo de solución de controversias. Ciertamente redes sociales

³³ Sobre las condiciones de los moderadores véase ROBERTS, T. S, *Behind the screen. Content moderation in the shadows of the social media*, Yale University Press, USA, 2018.

³⁴ Ejemplo de este fenómeno es la aplicación de la Network Enforcement Act por la red social Facebook. La normativa alemana obliga a plataformas en línea de cierto volumen a establecer mecanismos de denuncia para eliminar determinado contenido que es sancionado por el Código Penal alemán. Asimismo, establece una serie de obligaciones adicionales a las plataformas, entre ellas un deber de remitir informes cada seis meses sobre sus prácticas de moderación. Contrastando con los informes remitidos por otras plataformas, se hizo notar que el volumen de retiradas en Facebook era claramente inferior. Esto se debía a que cuando un contenido era reportado por el procedimiento interno que hacía imperativo la Ley alemana se establecía un filtro previo en el que dicho contenido era analizado a la luz de sus términos y servicios. De esta forma, únicamente el contenido que no violaba sus términos era posteriormente analizado en base a la Network Enforcement Act, cosa que disminuía considerablemente el volumen de trabajo de los grupos de moderadores expertos en la normativa. Véase el informe de Facebook aquí: https://about.fb.com/wp-content/uploads/2018/07/facebook_netzdg_july_2018_english-1.pdf (última consulta 11 de noviembre de 2021).

como Facebook, sí que se preocupan de referirse a las vías de recurso, si bien lo hacen en un apartado de sus condiciones generales y no en la descripción de sus “normas comunitarias”, donde se delimitan qué comportamientos no son admisibles en su foro³⁵.

Cuando nos encontramos ante grandes plataformas que prestan servicios de intermediación de contenidos compartidos por los usuarios, lo más habitual es que los usuarios accedan a esos medios para realizar actividades ajenas a su ámbito profesional. En este supuesto habrá que atenerse a lo dispuesto en el art. 18 del Reglamento 1215/2012³⁶, pudiendo el consumidor optar entre demandar a la plataforma tanto en el Estado Miembro donde esté domiciliada la misma o el lugar donde esté domiciliado el propio consumidor. No obstante, la moderación de contenidos puede afectar especialmente a profesionales que utilicen las plataformas digitales para promover sus productos³⁷. En estos supuestos deberá valorarse lo dispuesto en el art. 7 de la citada disposición, y dependiendo de dónde se entienda que se presta el servicio de la sociedad de la información el usuario podrá demandar al prestador de alojamiento en el Estado donde se encuentra domiciliado o por el contrario deberá de entablar su reclamación en el Estado miembro donde se encuentre domiciliado el prestador.

3.2. El deber de proporcionar un procedimiento interno de revisión

Seguramente una de las innovaciones más importantes de la PLSD es la obligatoriedad de establecer procedimientos internos de revisión ante determinadas decisiones que impliquen moderación de contenidos (art 17)³⁸. Que la controversia

³⁵ En los términos de Facebook se señala “Si eres consumidor y tu residencia habitual se encuentra en un estado miembro de la Unión Europea, las leyes de dicho país se aplicarán a cualquier reclamación, causa o disputa que inicies contra nosotros y que surja como consecuencia de estas Condiciones o los Productos de Facebook, o en relación con ellos (“reclamación”). Asimismo, puedes resolver la reclamación en cualquier tribunal competente del país que tenga jurisdicción. En todos los demás casos, aceptas que la reclamación debe resolverse en un tribunal competente en la República de Irlanda y que las leyes de dicho país regirán estas Condiciones y cualquier reclamación (independientemente de las disposiciones relativas a conflictos de derecho)”. Facebook, “Términos del servicio”, Disponible en: <https://www.facebook.com/legal/terms> (última consulta 11 de noviembre de 2021).

³⁶ Reglamento 1215/2012 del Parlamento Europeo y del Consejo de 12 de diciembre de 2012 relativo a la competencia judicial, el reconocimiento y la ejecución de resoluciones judiciales en materia civil y mercantil.

³⁷ Además de periódicos o empresas que utilicen las redes sociales como canal para ofrecer al público general sus productos, otra serie de usuarios como cuentas de partidos políticos o los propios perfiles de representantes políticos pueden verse afectados por la moderación privada. En España Twitter así a suspendido cuentas de partidos políticos con representación parlamentaria CARVAJAL, A., “Twitter suspende la cuenta de Vox al inicio de la campaña en Cataluña por relacionar inmigración con delincuencia” *El Mundo*, 28 de enero 2021 <https://www.elmundo.es/espana/2021/01/28/6013039ffdddf85a98b4668.html> (última consulta 11 de noviembre de 2021).

En Alemania e Italia controversias similares han llegado a judicializarse como tendremos ocasión de señalar más abajo.

³⁸ Recientemente, actos de la Unión Europea han venido a imponer límites a esta facultad de moderación, si bien en el ámbito de ciertas relaciones contractuales, como son las reguladas por el Reglamento 2019/1150 que impone normas de derecho imperativo en los contratos de intermediación en línea. Entre dichas disposiciones se encuentran normas relativas a la terminación unilateral del contrato a instancia del servicio de intermediación, preceptuándose que “Si un proveedor de servicios de intermediación en línea decide restringir o suspender para una determinado usuario profesional la prestación de sus servicios relacionados con bienes o servicios concretos ofrecidos por dicho usuario profesional, proporcionará a este,

surgida como consecuencia de la labor moderadora no se judicialice ni derive a órganos de resolución alternativos a la vía judicial puede ser de especial interés para el consumidor de servicios digitales. En la mayoría de los casos, cuando no medien intereses económicos de cierta entidad, el usuario no asumirá la carga de acudir a la vía judicial por la suspensión de su cuenta ni mucho menos por la remoción de un contenido concreto³⁹. Al mismo tiempo debe advertirse que la suspensión de las facultades de uso no es indefinida en la mayoría de los casos, y no suele prolongarse demasiado en el tiempo. Así, es beneficioso que sea el intermediario quién actúe pues es él quien puede reintegrar al usuario afectado en su posición previa lo más rápido posible.

El procedimiento interno de revisión no debe ser implementado por todos los prestadores de alojamiento de datos. Este únicamente se aplica a las plataformas en línea, esto es, "un prestador de un servicio de alojamiento de datos que, a petición de un destinatario del servicio, almacena y difunde al público información". Asimismo, también quedan excluidas ciertas plataformas que sean microempresas o pequeñas empresas.

El objeto del procedimiento de revisión son una serie de decisiones que adopta la plataforma respecto al contenido que comparte el usuario, o que afectan al estatus de su cuenta. La PLSD en este punto ha acertado al positivizar tres supuestos que vienen a abarcar todos los escenarios en los que el usuario puede verse perjudicado por las prácticas de moderación. El primero de ellos abarca aquellas decisiones que impliquen la retirada de información compartida por el usuario o inhabiliten el acceso a la misma. Por tanto, son susceptibles de fiscalización también aquellas decisiones que recaigan sobre contenidos concretos y no supongan una alteración de la situación de la cuenta del usuario. Estas últimas también se encuentran incluidas dentro del procedimiento, si bien se diferencian en dos grupos.

El primero de ellos se refiere a aquellas que tienen un carácter temporal o parcial, esto es, "las decisiones de suspender o cesar la prestación del servicio, en todo o en parte, a los destinatarios". Quedaría así tutelado un supuesto habitual en el que la plataforma no permite compartir información al usuario, limitando así las posibilidades de publicar en dicho espacio, pero en el que sí que puede acceder a las publicaciones de terceros y seguir usando otras funcionalidades del servicio. De ellas se diferencian las que tienen una vocación de terminar la relación entre usuario y plataforma, es decir las decisiones de suspender o eliminar la cuenta de los destinatarios. Como decíamos, debe calificarse de acierto el amplio abanico de supuestos que abarca la norma, que permite corregir una práctica previa en la que

antes o en el momento en que surta efecto la restricción o suspensión, una motivación de la decisión en un soporte duradero" (art 4) . Junto a ese deber de motivación, se suman otros deberes específicos como es el de ofrecer al profesional afectado por la medida la oportunidad de aclarar los hechos y circunstancias en el marco de un procedimiento interno de tramitación de reclamaciones. En nuestra doctrina ha tratado la cuestión ÁLVAREZ MORENO en, ÁLVAREZ MORENO, M., *La contratación electrónica mediante plataformas en línea: Modelo comercial (B2C), régimen jurídico y protección de los contratantes (proveedores y consumidores)*, REUS, 2021.

³⁹ CAUFFMAN, C., GOANTA, C., "A New Order: The Digital Services Act and Consumer Protection", *European Journal of Risk Regulation*, 2021, p. 14.

en algunos sistemas de resolución de controversias no se permiten apelar decisiones como la eliminación de la cuenta⁴⁰.

A pesar de ello, algunos actores en el procedimiento legislativo han sugerido que también deberían de ser objeto del procedimiento de revisión aquellas decisiones que denegasen la remoción de contenido⁴¹. Por tanto, de aceptarse esta modificación al texto de Comisión, las decisiones de las plataformas adoptadas como resultado del procedimiento de aviso del art 14 que fueran en contra del usuario que reporta el contenido también entrarían dentro de su ámbito de aplicación.

La PLSD establece determinados requisitos adicionales, que se refieren a aspectos cualitativos que deben observar estos procedimientos. Estos deberán de ser de fácil acceso y manejo, facilitando el envío de reclamaciones suficientemente precisas y adecuadamente fundamentadas. Asimismo, el procedimiento debe de ser gratuito para el usuario y debe poder tramitarse de forma electrónica. A ello hay que sumarle que este debe estar disponible durante un período mínimo de seis meses desde la decisión a que se refiere este apartado. Dicho plazo es bastante generoso teniendo en cuenta que los intereses del usuario requieren que la decisión sea revertida con la mayor prontitud.

Las plataformas deben resolver las reclamaciones de manera oportuna, diligente y objetiva (art 17.3). Ello conlleva la existencia de un deber de revertir la decisión que afecta al usuario cuando la plataforma tenga motivos suficientes para creer que la información removida no es ilícita ni incompatible con sus condiciones. Ese deber alcanzará también la revisión de la decisión de suspender la cuenta o una de sus funcionalidades cuando la calificación de la información que dio lugar a dicha suspensión sea revisada.

La PLSD no establece un plazo⁴² en el que la plataforma debe resolver la solicitud; tampoco para revocar los efectos de la decisión contra la que se dirige el procedimiento, si bien en este caso señala que deberá de realizarse "sin dilaciones indebidas" (art 17.3). El texto del Reglamento únicamente alude a un deber de diligencia en la tramitación de la solicitud, el cual necesariamente debe conllevar que deba ser resuelta en un plazo en el que el usuario pueda beneficiarse de los efectos de la revisión. Dicho esto, debido al carácter potestativo del uso de estos procedimientos, en aquellos casos en los que la dilación pueda causar un grave

⁴⁰ Por ejemplo, para acudir al proceso de apelación del Facebook *Oversight Board* (Consejo asesor de contenido) las apelaciones deben provenir del titular de una cuenta activa, pues es necesario que "la cuenta no pueda estar inhabilitada ni haber sido eliminada, y el propietario debe poder iniciar sesión en ella" Información adicional del procedimiento disponible en: <https://www.oversightboard.com/> (última consulta 18 de diciembre de 2021).

⁴¹ Nos referimos a la enmienda 98 del Informe preliminar del Comité del mercado interior y la protección de los Consumidores del Parlamento Europeo. El texto del informe se encuentra disponible aquí: https://www.europarl.europa.eu/doceo/document/IMCO-PR-693594_EN.pdf (última consulta 18 de diciembre de 2021).

⁴² La cuarta evaluación del Código de conducta contra el discurso de odio en línea señaló que, en 2019, las compañías resolvían en un 89% de los casos las solicitudes de remoción de contenidos en las primeras 24 horas. Parece razonable que se maneje con un plazo similar a la hora de responder ante las peticiones de revisión.

perjuicio al interesado la decisión de reactivar la cuenta podría resolverse en sede judicial, particularmente a través de su solicitud como medida cautelar. Este ha sido el modo de proceder en varios Estados miembros, en los que tribunales han concedido la reactivación cautelar de las cuentas de partidos políticos a la vista del posible perjuicio causado por la dilación del proceso⁴³.

Finalmente, la PLSD señala que “las plataformas en línea velarán por que las decisiones a que se refiere el apartado 4 no se adopten exclusivamente por medios automatizados”. Como decíamos más arriba, es frecuente que las plataformas en línea se valgan de la Inteligencia Artificial para ejercer sus prácticas de moderación. Si bien estos medios automatizados pueden ser de utilidad para retirar contenido previamente categorizado como ilícito, no lo son para realizar juicios autónomos. Fuera de los casos en los que la remoción de contenidos puede basarse en la detección de caracteres que impliquen lenguaje malsonante es dudoso que la IA de la que disponen las plataformas esté preparada para acometer semejante labor. La apreciación de si una determinada expresión es contraria al ordenamiento jurídico o a las condiciones de la plataforma requiere analizar el contexto en el que es vertida, y en su caso, la finalidad del autor. Ello requiere de un juicio humano, que entienda el significado social de las expresiones controvertidas y pueda atribuirle una calificación derivada del sentido global de lo comunicado y que es difícil de extraer de cada carácter o extracto de contenido⁴⁴. Teniendo en cuenta estas circunstancias, si ya es discutible que las plataformas utilicen la IA para remover contenidos o suspender cuentas, lo es más que sean estos aparatos quienes se ocupen de la revisión de sus decisiones previas⁴⁵.

3.3. Deberes de transparencia

La propuesta establece una serie de deberes de transferencia informativa que tratan de hacer frente a la frecuente opacidad con los que las plataformas ejecutan sus políticas de moderación de contenidos. Estas obligaciones tienen distinto alcance en función del tipo de prestador de servicios de la sociedad de la información y también su volumen⁴⁶.

⁴³ Por ejemplo, Decisión del Tribunal Regional Superior de Munich del 24 de agosto de 2018, 18 W 1294/18 Resolución disponible en: <https://openjur.de/u/2111165.html> (última consulta 2 de febrero de 2021) o caso *Casa Pound Italia e Davide Di Stefano v. Facebook* Resolución disponible en: <https://www.ilprimatonazionale.it/wp-content/uploads/2019/12/6374196s.pdf>. Véase TUCHTFELD, E., “Marktplätze, soziale Netzwerke und die BVerfG-Entscheidung zum “III. Weg””, *VerfBlog*, 2019/5/26, Artículo online disponible en: <https://verfassungsblog.de/marktplaetze-soziale-netzwerke-und-die-bverfg-entscheidung-zum-iii-weg>

⁴⁴ Véase, European Parliamentary Research Service “Regulating disinformation with artificial intelligence, Effects of disinformation initiatives on freedom of expression and media pluralism”, European Parliamentary Research Service, Brussels, 2019

⁴⁵ No obstante, tampoco pueden ignorarse las condiciones materiales en las que se efectúa la supervisión humana de los contenidos, que como ha relatado SARA ROBERTS, no pueden dar lugar a un juicio respetuoso con los derechos de los usuarios ni tampoco revertir en la salud de los moderadores. Si tales condiciones no mejoran el recurso a la IA para ejecutar tareas de moderación podría considerarse un mal menor, particularmente si se tiene en cuenta que la exposición continua al contenido ilícito que sufren los moderadores es perjudicial para su salud mental. ROBERTS, T.S., *Behind the screen. Contend moderation in the shadows of the social media*, Yale University Press, USA, 2018.

⁴⁶ Hasta el momento los esfuerzos de autoregulación que habían precedido al enfoque regulatorio de la PLSD justamente han coincidido en una colaboración arbitraria y fragmentada de las plataformas con la

El primer deber, aplicable en este caso a todos los prestadores intermediarios, y con independencia de su volumen de negocio, es el de especificar en sus términos del servicio determinados aspectos sobre la forma en que ejecutan su política de moderación de contenidos (art 12.1.). Así, deberán de publicar “información sobre cualquier restricción que impongan en relación con el uso de su servicio al respecto de la información proporcionada por los destinatarios del servicio”, información que debe abarcar “todo tipo de políticas, procedimientos, medidas y herramientas que se utilicen con fines de moderación de contenidos, incluidas las decisiones algorítmicas y la revisión humana.”

En la actualidad los términos y condiciones de plataformas en línea calificadas como “redes sociales” no son ciertamente prolijas a la hora de especificar los términos en los que se ejecuta la moderación de contenidos. Progresivamente estas condiciones han sido ampliadas, tratando de especificar con mayor precisión los supuestos en los que un usuario puede ver su contenido removido o limitadas las funciones de su cuenta⁴⁷. No obstante, las condiciones en las que se ejercen la moderación, número de moderadores, su localización geográfica y formación; el uso de inteligencia artificial para detectar o actuar contra el contenido; los criterios que sigue la plataforma para optar por un tipo de sanción (apercibimiento, suspensión de cuenta o inhabilitación definitiva) son por lo general aspectos que tienen especial repercusión en la moderación y no son conocidos⁴⁸.

Por su parte, el art 13 impone un régimen general aplicable a todos los prestadores de servicios de la sociedad de la información que presten servicios de *hosting*, y que no tengan la caracterización de microempresas (art 13.2.). Ese deber se concreta en la realización de un informe anual, que sea claro, detallado y fácilmente comprensible sobre cualquier actividad de moderación de contenidos que hayan realizado durante ese periodo.

El informe debe contener “el número de órdenes recibidas de las autoridades de los Estados miembros, categorizadas según el tipo de contenido ilícito de que se trate, incluidas las órdenes dictadas de conformidad con los artículos 8 y 9, y el tiempo medio necesario para llevar a cabo la actuación especificada en dichas órdenes”. Así, deberán de notificarse el número de ordenes de remoción de contenido o solicitud de información que se regulan en los mencionados artículos de la propuesta. Especial acertada es la inclusión del tiempo “medio necesario para llevar a cabo la actuación” que permitirá fiscalizar la diligencia del prestador a la hora de verificar su cumplimiento.

sociedad civil y los investigadores. Ejemplos de estas carencias pueden encontrarse en los primeros informes de aplicación de Código de Buenas Prácticas contra la desinformación.

⁴⁷ BUNI, C., CHEMALY, S., “The secret rules of the internet. The murky history of moderation, and how it’s shaping the future of free speech”, The Verge, 15 de abril de 2016, Disponible en: <https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech> (última consulta 27 de diciembre de 2021).

⁴⁸ Es frecuente además que los moderadores cuenten con instrucciones internas de trascendencia para las tareas de moderación pero que no sean de conocimiento público. Véase HOPKINS, N., “Revealed: Facebook’s internal rulebook on sex, terrorism and violence” *The Guardian*, 21 de mayo 2017. Disponible en: <https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence> (última consulta 27 de diciembre de 2021).

Asimismo, se deberán de exponer el número de órdenes recibidas conforme al art. 14, es decir, la cantidad de avisos que son remitidos por los destinatarios del servicio a través del procedimiento de reclamación y la decisión que tomó el prestador frente a ellos. En esa información deberán de distinguirse si dichos avisos versaban sobre contenido que era ilícito, o, por el contrario, contenido que no era conforme a las condiciones del servicio. También en este caso el prestador deberá de especificar el tiempo medio que le toma actuar contra el contenido reportado por los usuarios.

Junto al contenido removido a instancia de avisos realizados por usuarios, la PLSD obliga a las plataformas a informar sobre el contenido removido como resultado de investigaciones propias realizadas por la propia plataforma. Debe recordarse, que conforme al art. 15 de la DCE (ahora art 7 de la propuesta) los Estados miembros no pueden imponer un deber general de monitorización, que obligue a los prestadores a supervisar el contenido que circula en sus sitios webs. Eso no impide que los prestadores deban de responder a órdenes judiciales que afectan a contenido individualizado⁴⁹. Tampoco que los propios prestadores realicen investigaciones voluntarias para detectar contenido ilícito o contrario a sus términos, y actúen contra dicha información una vez que concluyan que merecen de tal calificativo.

Sobre estas medidas de remoción voluntaria, la PLSD señala que los prestadores deberán de notificar su número y el tipo de medidas adoptadas que afecten a la disponibilidad, visibilidad y accesibilidad de la información proporcionada por los destinatarios del servicio. A ello hay que sumarle "la capacidad de los destinatarios para proporcionar información, categorizada según el tipo de motivo y fundamento para adoptar esas medidas". Con este último inciso la PLSD parece hacer referencia al papel de los destinatarios de la medida respecto a la decisión tomada por el prestador. En particular, pretende esclarecer en qué medida los usuarios han podido realizar aclaraciones o remitir información que ofrezcan a la plataforma un contexto más detallado a la hora de tomar su decisión.

La última letra del art 13, se refiere al deber de publicar "el número de reclamaciones recibidas a través del sistema interno de tramitación de reclamaciones a que se refiere el artículo 17, el fundamento de dichas reclamaciones, las decisiones adoptadas al respecto de dichas reclamaciones, el tiempo medio necesario para adoptar dichas decisiones y el número de casos en que se revirtieron dichas decisiones." Debe recordarse que los deberes de información contemplados en el art 13 son aplicables a todos los prestadores de servicios de *hosting*, sean o no plataformas en línea, mientras que el procedimiento de recurso interno del art 17 es únicamente imperativo para los prestadores calificados como plataformas en línea. Una lectura armónica de ambos preceptos nos hace concluir que los prestadores de *hosting* que no tengan tal consideración no deben implementar dichos procedimientos, si bien, si así los introducen de forma voluntaria, deberán de recopilar la información relativa a ellos en sus informes anuales.

⁴⁹ STJUE de 3 de octubre de 2019, asunto C-18/18 (Eva Glawischnig-Piesczek v. Facebook Ireland Limited), apdo. 45.

Junto a estos deberes de transparencia aplicables a los prestadores de alojamiento de datos, el art 23 de la propuesta amplía los deberes de transparencia de los prestadores de servicios de la sociedad de la información que merecen la calificación de plataformas en línea. Estos deberes son correlativos a las obligaciones adicionales que se imponen a este tipo de prestadores⁵⁰, regulados en los artículos 17 a 22 de la propuesta, y que quedan fuera del objeto de este estudio.

Sin embargo, sí que es relevante a efectos de nuestra exposición lo dispuesto en el art 23.1. d), que señala que las plataformas en línea deberán de informar del “uso de medios automáticos con fines de moderación de contenidos, incluida una especificación de los fines precisos, los indicadores de la precisión de los medios automatizados para cumplir dichos fines y las salvaguardias aplicadas”. Contrastando con la información que debe de recopilarse en los términos y servicios sobre el uso de estos medios (art 12.1.), el informe que deben de remitir las plataformas en línea indudablemente deberá de incidir con mayor detalle en este aspecto.

Finalmente merece la pena reseñar que las plataformas en línea de muy gran tamaño se encuentran sujetas a mayores deberes de transparencia, precisamente por su papel en el entorno online y en el desenvolvimiento de los procesos democráticos (art 33). Como ya hemos apuntado en el punto segundo de este trabajo, las plataformas en línea de muy gran tamaño deben someterse “a su propia costa y al menos una vez al año”, a una auditoría independiente que supervise el cumplimiento de las obligaciones previstas en el Capítulo III de la propuesta⁵¹. Las organizaciones auditoras realizaran un informe en el que conste un dictamen que deberá de ser favorable o desfavorable, en el que se determine si la plataforma en línea de muy gran tamaño ha cumplido con las previsiones del Capítulo III. Dichos informes, además de otros documentos como las evaluaciones de riesgos a las que se refiere el art 26, y las medidas implementadas para gestionar los riesgos (art 27), deberán de hacerse públicos⁵². Asimismo, deberán de publicar los informes de transparencia a los que se refiere el mencionado art. 13 en un periodo de seis meses, contrastando con la periodicidad anual que es exigible al resto de prestadores.

Dentro de las obligaciones de transparencia se incluye un procedimiento contradictorio a través del cual el coordinador de servicios digitales y la Comisión pueden solicitar el “acceso a los datos que sean necesarios para vigilar y evaluar el cumplimiento del presente Reglamento” (art. 31.1). Este procedimiento se encuentra también abierto a aquellos investigadores autorizados por el coordinador de servicios

⁵⁰ Algunos de ellos ya los hemos tratado, como la necesidad de imponer sistemas de reclamación interno frente a las decisiones de moderación de contenidos. Otros pasan por someterse a órganos de resolución extrajudicial de litigios (art 18), priorizar los avisos de ciertas entidades consideradas como alertadores fiables (art 19), suspender las cuentas de destinatarios del servicio cuando se aprecien infracciones reiteradas (art 20) y alertar a las autoridades competentes de ciertos ilícitos de gravedad (art 21).

⁵¹ Este abarca todas las obligaciones de diligencia que tiene por fin crear un entorno en línea seguro.

⁵² No obstante, la PLSD establece una excepción como es que la publicación suponga “la revelación de información confidencial de esa plataforma o de los destinatarios del servicio, las vulnerabilidades importantes para la seguridad de su servicio, menoscabe la seguridad pública o perjudique a los destinatarios”. En estos casos se permite a la plataforma retirar dicha información de los informes si bien deberá transmitir los informes completos al coordinador de servicios digitales de establecimiento y a la Comisión, “acompañados de una exposición de los motivos por los que se haya retirado la información de los informes públicos” (art 33. 3)

digitales o a la Comisión, si bien en este caso únicamente a efectos de estudiar los riesgos sistémicos a los que se refiere el art 26⁵³. La solicitud debe ser motivada, y debe aludir a un "periodo razonable" en el que la plataforma en línea debe de atenderla. El intermediario dispone, por el contrario, de un plazo concreto de 15 días para mostrar su objeción a la solicitud, pidiendo su modificación cuando no tenga acceso a los datos o bien cuando "el acceso a los datos implique vulnerabilidades importantes para la seguridad de su servicio o para la protección de información confidencial, en particular secretos comerciales". Cuando se propone esta modificación de la solicitud se deberá acompañar a ella una propuesta que contenga "uno o varios medios alternativos mediante los cuales pueda otorgarse el acceso a los datos solicitados u otros datos que sean apropiados y suficientes para la finalidad de la solicitud". Recibida la propuesta de modificación el coordinador de servicios digitales o la Comisión, según proceda, decidirá sobre ella en un plazo de quince días (art 31.7).

4. CONCLUSIONES

Es de esperar que la PLSD ofrezca a los consumidores de servicios digitales una protección sustancial contra las decisiones que impliquen la moderación de sus contenidos. Esto es así puesto que la propuesta de Reglamento supera el modelo de la DCE, que, aunque preocupada indirectamente por los efectos de la moderación, y de la existencia de una censura colateral⁵⁴, no estableció previsiones que protegieran a los usuarios de servicios digitales de prácticas arbitrarias en la gestión de sus contenidos. Con el enfoque de la PLSD, esa actividad de gestión, a la que se denomina moderación de contenidos es directamente supervisada, estableciéndose un cuerpo de obligaciones que los prestadores intermediarios deben observar.

Estas obligaciones no vedan directamente ningún tipo de moderación, es decir, no aclaran la relación entre la moderación de contenidos y el ejercicio de las libertades fundamentales en los espacios que gestionan las plataformas en línea. Lo que la PLSD trata de implementar son una serie de provisiones que afectan a la forma en que las plataformas moderan, imponiendo procedimientos y requisitos que deben de seguir cuando retiran un contenido o suspenden la cuenta de un destinatario del servicio.

Creo que la estrategia seguida por la PLSD es acertada, ya que consigue proteger los intereses de los usuarios de servicios digitales, sin impedir la moderación, pero tampoco blindándola. Particularmente consigue establecer límites a la moderación que deben entenderse complementarios al recurso a la vía judicial, el cual siempre está disponible a los usuarios que busquen remedios alternativos.

Las garantías son múltiples, y se han agrupado en tres grandes grupos que pueden resumirse en: el deber de notificar del motivo de la moderación, el acceso a

⁵³ Debe tenerse en cuenta que hasta el momento el acceso de los investigadores a datos relevantes para luchar contra estos fenómenos, por ejemplo, la desinformación ha sido "arbitrario y episódico", tal como criticó la Comisión en su primer informe anual sobre la aplicación del Código de buenas prácticas contra la desinformación.

⁵⁴ BALKIN, J. M., "Old School/New School Speech Regulation" Harvard Law Review, Forthcoming, Yale Law School, Public Law Research Paper No. 491, 2014, p. 2309

procedimientos internos para apelar las decisiones del intermediario y disposiciones que garantizan la transparencia de sus actividades. Estas son de un carácter asimétrico, modulándose su intensidad y alcance en función del tipo de prestador y su volumen de negocio. Este enfoque también merece calificarse acertado pues evita que compañías de pequeño tamaño deban de observar previsiones que puedan dificultar su crecimiento, siendo también respetuoso con los derechos del usuario ya que son los intermediarios de mayor tamaño, generalmente plataformas en línea, las que pueden afectar con mayor intensidad a sus intereses.

El alcance de muchas de estas obligaciones es en ocasiones, dudoso. Eso no debe de ser signo de alarma pues en el momento en el que se escriben estas líneas el procedimiento legislativo está lejos de encontrarse concluso, quedando un margen temporal suficiente para perfilar con mayor precisión los deberes de los intermediarios. El carácter abierto de dichas obligaciones también podrá ser mitigado en la práctica, principalmente a través de las opiniones que ofrezca la Junta de Servicios Digitales, y muchos instrumentos a caballo entre el "hard" y el "soft law" que se regulan en el propio reglamento, como pueden ser las normas sectoriales (art 35) y los Códigos de Conducta (art 35).

5. BIBLIOGRAFIA

ÁLVAREZ MORENO, M., *La contratación electrónica mediante plataformas en línea: Modelo negocial (B2C), régimen jurídico y protección de los contratantes (proveedores y consumidores)*, REUS, Madrid, 2021.

BALKIN, J. M., "Old School/New School Speech Regulation" *Harvard Law Review*, Forthcoming, Yale Law School, Public Law Research Paper No. 491, 2014.

BUNI, C., CHEMALY, S., "The secret rules of the internet. The murky history of moderation, and how it's shaping the future of free speech", *The Verge*, 15 de abril de 2016, Disponible en: <https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech> (última consulta 27 de diciembre de 2021).

CARVAJAL, A., "Twitter suspende la cuenta de Vox al inicio de la campaña en Cataluña por relacionar inmigración con delincuencia" *El Mundo*, 28 de enero 2021

CAUFFMAN, C., GOANTA, C., "A New Order: The Digital Services Act and Consumer Protection", *European Journal of Risk Regulation*, 2021.

COMMITTEE ON CIVIL LIBERTIES, JUSTICE AND HOME AFFAIRS, "Opinion of the Committee on Civil Liberties, Justice and Home Affairs for the Committee on the Internal Market and Consumer on the proposal for a regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC", 28 de julio de 2021, Disponible en: https://www.europarl.europa.eu/doceo/document/LIBE-AD-692898_EN.pdf

DE MIGUEL ASENSO, P M., *Derecho privado de Internet*, Civitas, Madrid, 2015.

DE MIGUEL ASENSIO, P.A., "Servicios y mercados digitales: modernización del régimen de responsabilidad y nuevas obligaciones de los intermediarios" *La Ley Unión Europea*, Núm. 88, 2021.

DIGITALEUROPE, "Digital Service Act position paper", 31 de marzo de 2021. Disponible en: <https://www.digitaleurope.org/resources/digital-services-act-position-paper> (última consulta 27 de diciembre de 2021).

EUROPEAN COUNCIL, "What is illegal offline should be illegal online: Council agrees position on the Digital Services Act", 25 de noviembre de 2021. Disponible en: <https://www.consilium.europa.eu/en/press/press-releases/2021/11/25/what-is-illegal-offline-should-be-illegal-online-council-agrees-on-position-on-the-digital-services-act/> (última consulta 21 de diciembre de 2021).

EUROPEAN PARLIAMENTARY RESEARCH SERVICE "Regulating disinformation with artificial intelligence, Effects of disinformation initiatives on freedom of expression and media pluralism", European Parliamentary Research Service, Brussels, 2019

HOPKINS, N., "Facebook's internal rulebook on sex, terrorism and violence" *The Guardian*, 21 de mayo 2017. Disponible en: <https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence> (última consulta 27 de diciembre de 2021).

GORWA, R., BINNS, R., KATZENBACH, C., "Algorithmic content moderation: Technical and political challenges in the automation of platform governance" *Big Data & Society*, 2020.

GOLDMAN, E., "Online User Account Termination and 47 U.S.C. §230(c)(2)", *Santa Clara Law Digital Commons*, Vol. 2(659), 2011.

LÓPEZ RICHART, J., "Difamación en la Web 2.0 y responsabilidad civil de los prestadores de servicios de alojamiento", *Derecho Privado y Constitución*, Núm. 26, 2012.

PARLAMENTO EUROPEO "Resolución del 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre la Ley de servicios digitales: una mejora del funcionamiento del mercado único (2020/2018) apdo B. Disponible en: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0272_ES.html#title1 (última consulta 17 de diciembre de 2021).

ROBERTS, T.S., *Behind the screen. Contend moderation in the shadows of the social media*, Yale University Press, USA, 2018.

TUCHTFELD, E., "Marktplätze, soziale Netzwerke und die BVerfG-Entscheidung zum "III. Weg"", *VerfBlog*, 2019/5/26, Artículo online disponible en: <https://verfassungsblog.de/marktplaetze-soziale-netzwerke-und-die-bverfg-entscheidung-zum-iii-weg>

VÁZQUEZ 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 Deusto*, Núm. 68/1, 2020.