



OVERBLOCKING EN REDES SOCIALES Y OTRAS
PLATAFORMAS: RIESGOS Y REMEDIOS PARA LOS
USUARIOS PROVEEDORES DE CONTENIDOS

POL OLIVÉ BUSQUETS

ESTUDIANTE DEL GRADO DE DERECHO

UNIVERSITAT DE BARCELONA*

WORKING PAPER 1/2020

**WORKING PAPERS
JEAN MONNET CHAIR**



**EUROPEAN
PRIVATE LAW**

 UNIVERSITAT DE
BARCELONA

Resumen: Cada día, plataformas como las redes sociales bloquean una gran cantidad de contenidos subidos por sus usuarios, ya sea a través de mecanismos de inteligencia artificial pensados para detectar ilícitos, o mediante revisión humana, tras la notificación de terceros. Muy frecuentemente, los contenidos eliminados no son ilícitos y están amparados en la libertad de expresión, pero los mecanismos de los que gozan los usuarios para apelar las decisiones de las plataformas son ineficaces, incomprensibles, oscuros e incluso muchas veces inexistentes. En un momento en el que se plantea imponer nuevas obligaciones a las plataformas para la detección y bloqueo de contenidos ilícitos (vid. el art. 17.4 de la Directiva CDSM), se hace más necesaria que nunca una regulación que garantice los derechos y libertades fundamentales de los usuarios.

Título: *Overblocking* en redes sociales y otras plataformas: riesgos y remedios para los usuarios proveedores de contenidos

Palabras clave: overblocking, usuarios, contenido generado por los usuarios, proveedor de contenidos, notificación y retirada, transparencia

Abstract: Every day, online platforms such as social media networks block a large amount of user generated content after being reviewed by artificial intelligence or claimed by third parties. Very often, this contents prove to be lawful (because they constitute freedom of expression), but the remedies available for users to appeal or challenge these decisions are usually ineffective, incomprehensible, obscure and even many times nonexistent. At a time when it is proposed to establish new obligations to the platforms for the detection and blocking of illegal content (such as the art. 17 of the CDSM Directive), a regulation that guarantees the fundamental rights and freedoms of users becomes more necessary than ever.

Title: Overblocking in social media and other platforms: risks and remedies for content providers

Keywords: overblocking, users, user generated content, content providers, notice and take down, counternotification, transparency

* Trabajo realizado en el marco del seminario “La responsabilidad civil en internet. Plataformas y otros intermediarios prestadores de servicios online”, organizado por la Càtedra Jean Monnet de Dret Privat Europeu de la Universitat de Barcelona, en la Facultat de Derecho de la misma Universidad, los días 25/10/2019, 7/11/2019 y 29/11/2019.

Índice

I. INTRODUCCIÓN	4
II. LA DIFICULTAD PARA DETECTAR ILÍCITOS EN LÍNEA.....	5
1. Pornografía infantil, terrorismo y discurso del odio	6
2. Derechos de la personalidad y libertad de expresión	6
3. Las prácticas creativas en internet. Derecho de autor y sus límites	8
III. EL FENÓMENO DEL OVERBLOCKING.....	11
1. Caso especial: <i>shadow banning</i>	12
2. Los sistemas de actuación de las plataformas para la eliminación de contenido ilegal.....	13
a) La Directiva e-commerce y el sistema de <i>notice and take down</i>	13
b) Sistemas derivados: <i>notice and stay down</i> y <i>notice and notice</i>	14
c) Sistemas de filtrado <i>ex ante</i> . El art. 17 de la Directiva CDSM	16
IV. SALVAGUARDAS PARA LOS DERECHOS DE LOS USUARIOS	19
1. La notificación al proveedor	20
a) Prevención contra el abuso y la temeridad de los denunciantes.....	21
b) Legitimación y anonimato	21
2. La comunicación de las medidas adoptadas	22
3. Mecanismos de oposición, apelación y reinstauración de contenidos.....	23
4. Autorregulación y arbitraje	25
V. CONCLUSIONES.....	26
VI. BIBLIOGRAFÍA.....	27

I. INTRODUCCIÓN

El problema entorno al control y filtrado por parte de los proveedores de servicios de internet de contenidos subidos por los usuarios no es nuevo. Es una manifestación más del conflicto entre distintos derechos; en especial, entre la libertad de expresión y los derechos de la personalidad. Además, suscita el debate sobre el grado de responsabilidad que deberían asumir las plataformas de prestación de servicios en línea, atendidas estas tres circunstancias: en primer lugar, la trascendencia social y pública que han adquirido las plataformas digitales, que hacen casi imposible desarrollar actividades relacionadas con la vida social, política o cultural si no es a través de ellas. En segundo lugar, la posición privilegiada que tienen los prestadores de servicios para prevenir, impedir y eliminar contenidos, sumado a su fácil identificación y solvencia a la hora de responder ante las víctimas de estas infracciones. Y en último lugar, el rol cada vez más activo que desarrollan estas plataformas, cuyos enormes beneficios provienen precisamente de esta actividad.

Como el volumen de contenidos que cargan los usuarios actualmente es ingente, proceder al filtrado mediante supervisión humana no parece una opción viable, y el riesgo de censura injustificada de contenidos lícitos mediante técnicas automatizadas es cada vez mayor. La generalización de los sistemas de detección y filtrado de ilícitos basados en la inteligencia artificial que consagra, por ejemplo, el artículo 17 de Directiva 2019/790/UE (en adelante Directiva CDSM), puede agravar la situación, ya preocupante, de los usuarios, frente a abusos y extralimitaciones en la remoción de contenidos legítimos cargados por ellos en las plataformas. Para evitar una situación de indefensión de los usuarios debe perseguirse una actuación transparente, proporcional, neutral y adecuada a derecho de las plataformas en línea, y garantizar la habilitación de mecanismos eficaces a través de los cuales los usuarios puedan hacer valer sus derechos de forma no abusiva.

Este artículo tiene por objeto el análisis del fenómeno de la remoción injustificada de contenidos subidos en línea por parte de usuarios, las dificultades que presenta la correcta detección de ilícitos en internet (en especial a través de sistemas de inteligencia artificial), y qué mecanismos podrían mitigar los efectos negativos de estos sistemas de detección y de los abusos de terceros, algunos ya mencionados en el citado artículo de la Directiva CDSM y pendientes de desarrollo en su transposición al derecho nacional.

II. LA DIFICULTAD PARA DETECTAR ILÍCITOS EN LÍNEA

Como señala la Comisión Europea¹, lo que es ilegal *offline* es ilegal *online*. Sin embargo, trazar el límite a partir del cual un contenido vulnera un derecho no es una tarea nada sencilla. En cada sector del ordenamiento jurídico el equilibrio y la ponderación entre los distintos derechos involucrados tiene matices, y muy a menudo su calificación jurídica precisará de un pronunciamiento judicial. Es una realidad muy presente que la detección de ilícitos llevada a cabo por las plataformas suelen llevar a eliminaciones arbitrarias de contenidos, en gran medida porque la inteligencia artificial aplicada a los mismos no es capaz —¿todavía?— de distinguir entre contenidos protegidos por el derecho de autor y otros que no lo están; o entre el insulto y la parodia. En algunos casos, puede tratarse de un contenido exactamente igual o muy parecido y tener un efecto radicalmente diferente por estar ubicado en otro contexto, o simplemente tratarse de un contenido vulgar o desagradable pero amparado por la libertad de expresión o de información.

Más difícil todavía es la determinación de la ilicitud de aquellos contenidos similares a aquellos ya declarados ilegales. Como ha señalado recientemente el TJUE, es compatible con el derecho europeo la obligación impuesta a un prestador de servicios de eliminar no solamente contenidos declarados ilegales, sino también aquellos subidos con posterioridad que estén ligeramente modificados,² lo que puede derivarse en una gran cantidad de errores de valoración. Es pues sorprendente esta decisión, mucho más si se tiene en cuenta que legitima el control mediante mecanismos de inteligencia artificial.

Si bien es verdad que los problemas que pueden derivarse de una mala calificación de un ilícito en internet son bastante generales, existen diferencias sectoriales sustanciales, tanto en lo que respecta a los matices del ámbito concreto como en el régimen de responsabilidad que tienen los proveedores de servicios sobre el contenido infractor, que ha tenido un desarrollo desigual en distintos ámbitos. En términos generales, podemos clasificar los ilícitos en internet según resultan afectaciones al orden público, vulneraciones de los derechos de la personalidad o infracciones a los derechos inmateriales como los derechos de autor.

¹ COM(2017) 555 final, p. 2.

² STJUE C-18/18 (Glawischnig-Piesczek), ap. 41.

1. Pornografía infantil, terrorismo y discurso del odio

La pornografía infantil, la apología del terrorismo o el discurso del odio se presentan, en general, de forma más claramente identificable que el resto. Son especialmente graves, y por esto despiertan una sensibilidad mayor entre la ciudadanía; además, no es fácil imaginar usos amparados por la legislación que puedan confundirse, por ejemplo, con la pornografía infantil. Se trata pues de contenidos cuya ilegalidad es, en algunos casos, manifiesta y reconocible al entendimiento normal de una persona.³ Además, existe una armonización parcial o total en la definición de este tipo de contenidos en el ámbito europeo,⁴ e incluso se han adoptado códigos de conducta específicos para combatirlos.⁵ También los proveedores de servicios suelen atender más rápida y eficazmente las reclamaciones de retirada cuando se trata de este tipo de ilícitos.⁶

A pesar de todo ello, puede darse también, por ejemplo, el uso de discursos ilícitos de forma irónica. En muchos casos, utilizados precisamente para criticar a quien los emite originariamente o que, simplemente, son una burla sin intención de incitar al odio o a la violencia (ver Imagen 1). La necesidad de valorar estos matices hace mucho más complicada una correcta determinación de la efectiva ilegalidad de estos contenidos.

2. Derechos de la personalidad y libertad de expresión

Mucho se ha hablado de internet como mecanismo inmediato de comunicación entre de gente de todo el planeta. También, como contrapartida, de la enorme capacidad lesiva de la hiperconectividad en lo que respecta al derecho al honor y, en especial, a la intimidad personal. La entrada a las redes sociales de figuras políticas internacionales, celebridades y, en general, todo el mundo que aspira a tener un mínimo de representación pública comporta que todo el mundo puede interactuar con cualquier personalidad en cualquier momento; no son siempre interacciones inocuas. Además, la

³ Ejemplo paradigmático, el tweet manifiestamente racista y xenófobo de la política de extrema derecha Beatrix von Storch, que contenía el siguiente texto: “¿Qué demonios pasa en este país? ¿Por qué una página oficial de la policía de Nordrhein-Westfalen tuitea en árabe? ¿Creen que así calmarán a la manada de bárbaros y violadores que son los musulmanes? (Traducción propia). Más información: <https://www.zeit.de/digital/internet/2018-01/beatrix-von-storch-twitter-titanic-gesperrt-netzdg-afd>

⁴ Ver las Directivas 2011/93/EU, sobre pornografía infantil, la 2017/541/UE, sobre terrorismo, o la Decisión Marco 2008/913/JAI, sobre la lucha contra determinadas manifestaciones racistas o xenófobas.

⁵ Ver por ejemplo el Código de conducta para la lucha contra la incitación ilegal al odio en internet: http://ec.europa.eu/justice/fundamental-rights/files/hate_speech_code_of_conduct_en.pdf

⁶ COM(2017) 555 final, p. 4.

potencial promoción y viralidad de un determinado contenido se ve potenciado por el uso de algoritmos que reaccionan a la atracción que siente la gente por este contenido y multiplican exponencialmente su repercusión pública. El insulto, la difamación, la exposición íntima o incluso el linchamiento son una realidad muy presente en cualquier rincón de internet.

Encontramos en un primer grupo aquellos contenidos que son especialmente delicados porque pueden afectar gravemente los derechos fundamentales, como las imágenes de violencia o sexuales. Estos constituyen las vulneraciones más graves a la privacidad y a la protección de datos y, como tales, relativamente fáciles de detectar.⁷ En un segundo grupo, tenemos la difamación, el insulto o determinadas manifestaciones de la esfera íntima menos intrusivas. Al no ser los derechos de la personalidad derechos absolutos, determinar la ilicitud de un contenido en este ámbito no es en absoluto una tarea sencilla. Señalar exactamente dónde se encuentra el límite entre los derechos de la personalidad y el derecho a la libertad de expresión o información es una labor que ocupa los tribunales y que está lejos de resolverse definitivamente.⁸

En relación con el insulto o la injuria, se necesita un razonable juicio de ponderación y sentido común para determinar qué lo es o no realmente. El propio Tribunal Constitucional ha dicho reiteradamente que el derecho al honor en relación con la libertad de expresión es un “concepto jurídico normativo cuya precisión depende de las normas, valores e ideas sociales vigentes en cada momento”.⁹ Prueba de ello es que muchas veces incluso los propios humanos somos incapaces de distinguir la ironía o el carácter humorístico de un contenido en ausencia del contexto sociocultural en el que se ha emitido.¹⁰

⁷ Además, existen protocolos específicos para la lucha contra las vulneraciones de derechos de la personalidad de especial riesgo o especialmente sensibles. Por ejemplo, el Canal Prioritario de la AEPD <https://sedeagpd.gob.es/sede-electronica-web/vistas/formNuevaReclamacion/canalprioritario.jsf>

⁸ Ver STS 297/2016 (ECLI:ES:TS:2016:297), sobre si el “derecho al insulto” es una manifestación del derecho a la libertad de expresión. También STC 14/2003: el derecho al honor impide “la difusión de expresiones o mensajes insultantes, insidias infamantes o vejaciones que provoquen objetivamente el descrédito”.

⁹ SSTC 180/1999, 52/2002 y 51/2008.

¹⁰ Ver, por ejemplo, las páginas satíricas como El Mundo Today <https://www.elmundotoday.com/> o Der Postillon <https://www.der-postillon.com/>, donde absolutamente todos los contenidos son falsos a propósito y con una voluntad humorística, y que han resultado en multitud de confusiones por parte de algunos lectores, interpretando que pretendían dar una información veraz.

A las dificultades de apreciación puramente sustantivas se le añade la incertidumbre respecto a los hechos y, con especial relevancia, su veracidad. Esta veracidad es determinante para una correcta calificación de la legalidad un contenido ofensivo. Cabe dudar de que existan vulneraciones al derecho al honor cuando la supuesta difamación no expresa sino hechos objetivamente probados (*exceptio veritatis*). Por ejemplo, no podremos hablar de calumnias o de injurias en un comentario en una red social llamando estafador a un empresario sobre el que se ha dictado sentencia condenatoria por el delito de estafa.¹¹

Otra cuestión que podría plantearse es si el baremo para enjuiciar la afectación al derecho al honor en internet debería ser más laxa que en el mundo material. Esto puede encontrar su justificación, en primer lugar, en el hecho de que, en general, tendemos a ser más impulsivos en internet. También por la presencia de la mayoría de celebridades de ámbito mundial en las redes sociales, que pueden dar una sensación de proximidad irreal y motivar rápidamente a una actuación poco meditada. En este sentido, el TEDH recoge la importancia del contexto en el que se ubica un contenido (p. ej. un comentario ofensivo ubicado en internet) para hacer una apreciación proporcional sobre la ilicitud de un contenido.¹²

3. Las prácticas creativas en internet. Derecho de autor y sus límites

El fenómeno de la hiperconectividad y la facilidad para reproducir y compartir contenidos en línea ha transformado las prácticas creativas de forma crucial. Durante las últimas décadas la producción cultural ha vivido progresivamente un cambio de paradigma y de procesos, y ha llegado a dar la espalda a la noción de creación — clásica, singular, originaria— para abrir paso a nuevas formas de producción de contenidos basadas en la selección, reordenación, compilación, descontextualización e incluso apropiación de contenidos ya existentes. El creciente prestigio de la figura del comisario de arte o el DJ son buen ejemplo de ello.¹³ El contenido en línea generado

¹¹ Ver el art. 207 del Código Penal y la STS 1001/2001 (ECLI:ES:TS:2001:1001), F.J. nº 5 y 6.

¹² SSTEDH Delfi AS v. Estonia, 64569/09, n. 142 y 143, Magyar Tartalomszolgáltatók Egyesülete and Index.hu ZRT v. Hungary 22947/13, n. 69, Høiness v. Norway 43624/14, n. 67. Sobre el particular, ARROYO, Esther. “La responsabilidad de los intermediarios en internet. ¿Puertos seguros a prueba de futuro?” Cuadernos de Derecho Transnacional (CDT), Vol. 12, nº 1. 2020, nº 39.

¹³ Sobre este tema apunté un par de ideas en: “De la creación a la reordenación: el derecho de autor ante la postproducción cultural”. Disponible en: <https://www.linkedin.com/pulse/de-la-creaci%C3%B3n-reordenaci%C3%B3n-el-derecho-autor-ante-pol-oliv%C3%A9-busquets/> También: FROSIO, Giancarlo. “From the oral-formulaic tradition to digital remix”, en *Reconciling Copyright with Cumulative Creativity*. Edwar Elgar. 2018, pp. 279 y ss.

por los usuarios (*user generated content*), o lo que se ha llamado “cultura de internet” o “creatividad social”,¹⁴ está basado en gran medida en la práctica del *remix*, el *found footage* o el *sampling*, con los *memes* como ejemplo más paradigmático. Sin embargo, el amparo de estas prácticas bajo alguno de los límites al derecho de autor no siempre es posible.¹⁵ En estos casos, determinar cuándo, en efecto, nos encontramos ante un uso legítimo de una obra protegida, puede resultar extremadamente complicado.

Además, en este terreno existe también un problema de diferencias nacionales. En los países de tradición jurídica anglosajona o de *copyright*, tiende a analizarse si concurre o no una infracción bajo el baremo del *fair use* o *fair dealing* (uso justo o razonable). A modo de resumen, este sistema permite interpretar de forma abierta y flexible la concurrencia de un límite o excepción al derecho de autor, en atención a los requisitos de la regla de los tres pasos (Art. 9(2) del Convenio de Berna y art. 13 del Acuerdo ADPIC): esto es, que se trate de supuestos especiales, que no se afecte a la explotación normal de la obra y que no se produzca un perjuicio injustificado a los legítimos intereses del titular. YouTube, plataforma estadounidense, parte de este sistema y lo indica a sus usuarios como referencia para determinar si un tipo de uso es o no legítimo, afirmando, por ejemplo, que el uso de pequeños fragmentos audiovisuales de obras de terceros puede constituir un uso legítimo si de ellos nace una nueva obra.¹⁶

Sin embargo, en el derecho continental europeo (y en España en particular) el sistema de límites es un *numerus clausus*,¹⁷ que no permite el amparo en un límite si no es en relación con los previstos en la ley¹⁸ y a sus requisitos.¹⁹ Así, el uso no autorizado de

¹⁴ FROSIO. *Reconciling...* p. 283.

¹⁵ Algunos autores incluso se han planteado si el sistema de límites en la UE es demasiado rígido en relación con la tipología de contenidos que se produce actualmente. HILTY, Reto - KÖKLÜ, Kaya. “Limitations and Exceptions to Copyright in the Digital Age. Four Cornerstones for a Future-Proof Legal Framework in the EU” en STAMATOUDI, Irini (ed.) *New Developments in EU and International Copyright Law*. Wolters Kluwer, 2016, pp. 285 a 287.

¹⁶ Ver “¿Qué es el uso legítimo?” en las directrices de YouTube: <https://www.youtube.com/intl/es-419/about/copyright/fair-use/#yt-copyright-four-factors>

¹⁷ También en relación con los límites posibles que establece la Directiva 2001/29/CE (InfoSoc) en su artículo 5, que no permite a los estados miembros establecer otros diferentes. Ver SSTJUE C-301/15 (Soulier) ap. 34, C-161/17 (Renckhoff) ap. 16, C-476/17 (Pelham) aps. 58 y 65.

¹⁸ Ver arts. 31 y ss. LPI.

¹⁹ GARROTE FERNÁNDEZ-DIEZ, Ignacio. “¿Puede crearse un nuevo límite en la Ley de Propiedad Intelectual española para dar cobertura a los contenidos generados por los usuarios?”. En MARÍN LÓPEZ,

pequeños fragmentos audiovisuales de obras de terceros para constituir una nueva obra, solamente será legítimo en la medida en que encaje en un límite como el de cita (art. 32 de la Ley de Propiedad Intelectual, en adelante LPI) o parodia (art. 39 LPI), por lo que muchos de los contenidos generados por los usuarios mediante técnicas como el *remix* podrían resultar infracciones de propiedad intelectual.

Una dificultad añadida en la determinación del ilícito en el derecho de autor la encontramos en el propio carácter de obra. De acuerdo el art. 10 LPI, una obra estará protegida siempre y cuando sea original. El derecho de autor, a diferencia de otros derechos inmateriales como las marcas o los diseños, no nace con el registro, sino con la creación (art. 1 LPI). Por esto, como en muchas ocasiones no existirá un registro donde consultar, será difícil valorar si —por concurrir el requisito de originalidad—²⁰ estamos efectivamente ante una obra protegida o si, por el contrario, se trata de una prestación protegida por un derecho conexo²¹ o un elemento en dominio público. Como resultado de esta distinción podemos encontrar sustanciales diferencias en la duración de la protección o en el baremo de la infracción:

- a) El plazo general de protección de los derechos de autor es de 70 años después de la muerte del autor (art. 26 LPI). Sin embargo, el plazo de protección de los derechos conexos es mucho más corto. Además, algunos derechos conexos como la mera fotografía no existen en todos los países y, en los que sí, el plazo de protección puede variar. A modo de ejemplo ver art. 128 LPI, que establece un plazo de protección de 25 años frente a los 50 que establece el § 72 (3) UrhG en Alemania.
- b) El TJUE descartó en el caso C-476/17 (Pelham) la aplicación del límite del § 24 UrhG (*Freie Benutzung* o “uso libre” en la Ley alemana de derecho de autor), en la utilización de 2 segundos de la canción *Metall auf Metall*, del famoso grupo Kraftwerk, para la producción de otra canción (i.e. la técnica del *Sampling*); entre otros motivos, por tratarse de un derecho conexo (productor de fonogramas). En el caso de una obra, a diferencia del caso de las prestaciones protegidas por derechos conexos, cuando la fracción utilizada de una obra no constituye un mínimo de expresión debemos concluir

J. – CASAS VALLES, Ramon (coords.) *Estudios sobre la Ley de Propiedad Intelectual. Últimas reformas y materias pendientes*, pp. 262 y 263.

²⁰ Según la reciente jurisprudencia del TJUE, en el sentido de que sea una creación intelectual que refleje la personalidad de su autor, expresada mediante decisiones libres y creativas. Ver. SSTJUE C-5/08 (Infopaq), C-393/09 (BSA), C-145/10 (Painer), C-310/17 (Cofemel), C-683/17 (Levola Hangelò).

²¹ En el derecho español regulados en el Libro segundo de la LPI. En especial, fonogramas, producciones audiovisuales y meras fotografías.

que no hay un uso efectivo de la misma, porque se trata de un elemento que, en tanto que no es original por sí mismo (ver art. 10 LPI), es de dominio público (pongamos por ejemplo dos palabras al azar de un libro de Harry Potter). Sin embargo, dijo el Tribunal que, en atención al hecho que el derecho del productor de fonogramas no pretende proteger la creación sino la inversión realizada por el productor²², “el derecho exclusivo conferido por esta disposición al productor de fonogramas a autorizar o prohibir la reproducción de su fonograma le permite oponerse a que un tercero use una muestra sonora, incluso muy breve, de su fonograma con el fin de insertarla en otro fonograma, a menos que esa muestra sea incorporada de forma modificada y que no resulte reconocible al escucharla.”²³

III. EL FENÓMENO DEL OVERBLOCKING

En términos generales, entendemos el *overblocking* (sobrebloqueo), como la eliminación o censura injustificados de unos determinados contenidos por la aplicación abusiva o incorrecta de uno de los derechos que hemos visto en el apartado anterior, particularmente a través de un sistema automatizado de bloqueo o filtrado. En especial y por lo que nos interesa aquí, hace referencia a aquellos contenidos en línea publicados por los usuarios que son eliminados por la propia plataforma donde se alojan; en ocasiones, porque la legislación relevante impone a la plataforma un deber de control o actuación y, en otras, por su propia iniciativa de acuerdo con su política empresarial. Algunos ejemplos de casos discutibles recientes:



Imagen 1 (izquierda):

eliminación de una imagen de “Les Tres Bessones”, claramente irónica, por supuesta apología de la violencia u organizaciones peligrosas, al coincidir las iniciales de las protagonistas con un grupo terrorista.



Imagen 2 (derecha): eliminación de una entrada en la que se hacía burla del aspecto de los policías, por considerar que era equivalente al discurso del odio.

²² Nº 30 de la sentencia.

²³ Nº 39

1. Caso especial: *shadow banning*

El *shadow banning* o censura encubierta hace referencia al fenómeno que ocurre cuando en una plataforma de contenidos se da menos promoción pública a un determinado tipo de contenidos por la aplicación de un algoritmo opaco. De esta forma, si bien el contenido de un determinado usuario no resulta eliminado de la plataforma, aparece cada vez menos en el *feed* de sus seguidores o en los resultados de los motores de búsqueda. Esto, además de afectar a los derechos fundamentales de los usuarios, puede tener una trascendencia económica en el posicionamiento en los resultados de los motores de búsqueda. Gran manifestación de esta trascendencia es el ascenso de la figura profesional del SEO (*Search Engine Optimizer*). El SEO es aquella persona experta en los criterios de posicionamiento de un determinado motor de búsqueda, que realiza las modificaciones necesarias para que un determinado contenido (de una empresa, etc.) obtenga buenos resultados y se muestre preferentemente en las búsquedas de los usuarios.

Aunque no exista propiamente una eliminación de contenido —este se mantiene accesible en el perfil del usuario—, el efecto generado, es decir la pérdida de repercusión pública, es muy similar al bloqueo. Además, la aplicación de este tipo de inteligencia artificial puede resultar en un sesgo muy grave por motivo de ideología, género, raza, etc.²⁴ En este supuesto, el elemento de más peso es la opacidad de los algoritmos, que impide que sean cuestionados o recurridos. En el recientemente publicado Libro Blanco sobre inteligencia artificial, la Comisión se hace eco del riesgo que presentan la opacidad, la complejidad y la falta de predictibilidad de los algoritmos a los efectos comprobar si efectivamente cumplen con la legalidad vigente en materia de derechos fundamentales, especialmente la prohibición de discriminación.²⁵ Sin embargo, no se propone ninguna medida concreta para garantizar el ejercicio de estos derechos, más allá de mencionar que “puede ser necesario ajustar o clarificar la legislación existente en ciertos ámbitos.”²⁶

²⁴ “FaceApp apologizes for building a racist AI” en TechCrunch: https://techcrunch.com/2017/04/25/faceapp-apologises-for-building-a-racist-ai/?_ga=2.249941637.2084318103.1574535912-1162879941.1574535912. También “Algorithmic accountability” <https://techcrunch.com/2017/04/30/algorithmic-accountability/>.

²⁵ COM(2020) 65 final, “WHITE PAPER on Artificial Intelligence - A European approach to excellence and trust”, p. 10 y 11.

²⁶ *Ibid.* p. 14.

2. Los sistemas de actuación de las plataformas para la eliminación de contenido ilegal

Cuando hablamos en general de bloqueo o filtrado por parte de las plataformas, hacemos referencia a mecanismos que, desde el punto de vista del usuario, causan el efecto que un contenido alojado en internet deviene inaccesible.²⁷ Este efecto puede delimitarse en distintos niveles geográficos: a escala nacional, sobre un proveedor de servicios concreto, una red local o un dispositivo individual, y producirse por medio de diferentes mecanismos técnicos como el filtrado de búsquedas, el bloqueo de DNS, de direcciones IP, o de hiperenlaces.²⁸ Sin embargo, es relevante la distinción según el mecanismo de respuesta o actuación exigida a un proveedor de servicios a la hora de retirar contenidos ilícitos. De ello depende también la necesidad y el grado de automatización que requerirán los sistemas de detección y control de ilícitos.

a) La Directiva e-commerce y el sistema de *notice and take down*

El actual régimen de responsabilidad de los proveedores de servicios de internet se enmarca en la Directiva 2000/31/CE (en adelante Directiva e-commerce), que es esencialmente una norma de exención de responsabilidad con ciertos requisitos. En el régimen general de esta Directiva, la exención de responsabilidad para ciertas plataformas —que la Directiva define como prestadores de servicios de alojamiento de datos—²⁹ viene doblemente condicionada, bien por el desconocimiento del ilícito, bien por su retirada diligente una vez adquirido conocimiento del mismo.³⁰ A este sistema de respuesta se le ha venido llamando *notice and take down* (notificación y retirada), por ser la notificación una de las formas de puesta en conocimiento efectiva al prestador de servicios de un eventual ilícito.

Precisamente porque en este paradigma la decisión sobre la remoción del contenido corresponde a las plataformas, el riesgo de *overblocking* deriva de la especial tendencia a atender sistemáticamente las solicitudes de retirada de terceros interesados, sin hacer demasiadas valoraciones o ninguna sobre la ilicitud de los

²⁷ INTERNET SOCIETY. “Perspectives: Internet Society, Perspectives on Internet Content Blocking: An Overview”. 2017, p. 5. Disponible en: <https://www.internetsociety.org/resources/doc/2017/internet-content-blocking/>.

²⁸ *Ibid.* pp. 8 a 10.

²⁹ Art. 14 de la Directiva e-commerce.

³⁰ Art. 14.1 de la Directiva e-commerce en relación con el art. 16.1 de la Ley de Servicios de la Sociedad de la Información y Comercio Electrónico (en adelante LSSI).

contenidos eliminados. Este comportamiento se fundamenta, además de en el deseo de no perder la exención de responsabilidad que le confiere la Directiva, en la inclusión de cláusulas en las condiciones generales de la contratación³¹ que prevén en muchos casos limitaciones de responsabilidad en casos de remoción errónea de contenidos,³² por lo que las plataformas tienen más bien poco que perder en los casos de remoción excesiva de contenidos y ningún incentivo para velar por la libertad de expresión de los usuarios.³³ A esta situación se le suma el riesgo de notificación abusiva de los terceros (como titulares de derechos de autor) cuando, a sabiendas de que la plataforma lo eliminará casi automáticamente, notifican un contenido lícito. Para tratar de evitar esta situación de mala fe o temeridad, la DMCA americana (Sec. 512(g)(2)) prevé expresamente la responsabilidad por el uso conscientemente fraudulento o abusivo del sistema de notificación y retirada cuando de este resulta una eliminación de contenidos lícitos.³⁴ Los tribunales estadounidenses, además, han vinculado este “conscientemente” a la creencia de buena fe que el contenido es ilícito. Como resultado, podrá ser también un uso fraudulento (*misrepresentation*) cuando no exista ningún tipo de revisión de los contenidos y simplemente se confíe “ciegamente” en un rastreador automático de contenidos infractores.³⁵ Sin embargo, la Directiva e-commerce no prevé nada al respecto.

b) Sistemas derivados: *notice and stay down* y *notice and notice*

Del sistema de notificación y retirada se pueden derivar otros sistemas según se imponga a las plataformas un deber de control más o menos estricto. La característica común de estos sistemas es que parten de una notificación inicial que señala el

³¹ Ver, a modo de ejemplo, las Condiciones de Uso de Instagram: “Eliminación de contenido e inhabilitación o cancelación de la cuenta: Podemos eliminar cualquier contenido o información que compartas en el Servicio si consideramos que infringe estas Condiciones de uso o nuestras políticas (incluidas nuestras Normas comunitarias de Instagram), o si la ley lo exige. Podemos negarnos a proporcionarte o dejar de proporcionarte la totalidad o una parte del Servicio (incluso inhabilitar o cancelar tu cuenta) de inmediato si: infringes estas Condiciones de uso o nuestras políticas de forma clara, grave o reiterada, incluidas nuestras Normas comunitarias de Instagram, infringes reiteradamente los derechos de propiedad intelectual de otras personas o nos lo exige la ley.” Disponible en: <https://help.instagram.com/581066165581870>

³² ARROYO. “La responsabilidad...” nº 55.

³³ ANGELOPOULUS, Christina - SMET, Stijn. “Notice-and-Fair-Balance: How to Reach a Compromise between Fundamental Rights in European Intermediary Liability”. *Journal of Media Law*, vol. 8, 2016, p. 24.

³⁴ *Misrepresentations*, Sec. 512 (f) DMCA. Ver también el caso *Online Policy v Diebold* (N.D. California 2004).

³⁵ Ver el caso *Lenz v. Universal Music Corp* (9th Circuit 2016).

contenido supuestamente ilegal que se aloja en la plataforma. A partir de aquí, los efectos varían.

De acuerdo con el sistema de *notice and stay down*, la plataforma deberá no solamente velar por la eliminación el contenido notificado, sino también evitar que en el futuro se vuelva a subir. Esta obligación ha sido determinada en algunos casos por la jurisprudencia europea y nacional³⁶ como consecuencia de una interpretación amplia del régimen de responsabilidad derivado de la Directiva y el sistema de *notice and take down*.³⁷ Sin embargo, esta alternativa ha sido criticada por ser desproporcionada y posiblemente incompatible con el art. 15 de la Directiva e-commerce, que prohíbe imponer a los proveedores de servicios una obligación de supervisión general de contenidos,³⁸ además de imposibilitar a los usuarios futuros apelar las decisiones ya tomadas, a pesar de afectar claramente a sus derechos de expresión.³⁹ En la práctica, con el volumen de contenidos que se cargan en la actualidad, impedir que un contenido concreto se suba en el futuro a una plataforma requiere hacer uso de un mecanismo de bloqueo automatizado y permanente, con los riesgos y dificultades añadidas para la correcta determinación de la ilegalidad de los contenidos.

Otros sistemas, como el de *notice and notice*, proponen que la única obligación de las plataformas derivada de la notificación sea trasladar esta, a modo de advertencia, al usuario que ha subido el contenido. La relevancia de este sistema reside precisamente en que la plataforma no está obligada a retirar el contenido, por lo que resultaría bastante ineficaz para la lucha contra los ilícitos más graves o de más difícil reparación,

³⁶ Ver SSTJUE C-324/09 (eBay), C-18/18 (Glawischnig-Piesczek), y a modo de ejemplo, la doctrina del BGH (Tribunal Supremo alemán) sobre responsabilidad indirecta (llamada *Störerhaftung*, por su procedencia de la protección de la propiedad §1004 BGB), por la que el que, cuando sin ser el actor directo de una infracción, se está en una posición adecuada para prevenirla o cuando ocurre en el área de dominio (como el caso de algunas plataformas), se deben tomar las medidas necesarias para evitar que vuelva a ocurrir. Sobre la *Störerhaftung* ver WOLLIN, Sören. *Störerhaftung im Immaterialgüter- und Persönlichkeitsrecht: Zustandshaftung analog § 1004 I BGB*. Nomos. Baden-Baden, 2018 y las sentencias BGH GRUR 2004, 560 (Internet-Versteigerung I) GRUR 2007, 708 (Internet-Versteigerung II), GRUR 2008, 702 (Internet-Versteigerung III), GRUR 2013, 370 (RapidShare I), GRUR 2013, 1030 (RapidShare III).

³⁷ KUCZERAWY, Aleksandra. *Intermediary Liability and Freedom of Expression in the EU: from Concepts to Safeguards*. Intersentia, KU Leuven Centre for IT & IP Law Series, 2019, p. 233. KUCZERAWY, Aleksandra. "From 'Notice and Take Down' to 'Notice and Stay Down': Risks and Safeguards for Freedom of Expression". En FROSIO, Giancarlo (ed). *The Oxford Handbook of Intermediary Liability Online* (próxima publicación, 2019), pp. 14 y 15.

³⁸ KUCZERAWY. *Intermediary Liability...* p. 245 y 246.

³⁹ ANGELOPOULUS – SMET. *Notice-and-Fair-Balance...* p. 19. KUCZERAWY. *Intermediary Liability...* p. 246.

como aquellos que afecten la intimidad o, por ejemplo, la pornografía infantil. Sin embargo, algunas voces han defendido sus beneficios en el ámbito del derecho de autor, porque los titulares de derechos mantienen la posibilidad de acceder a la vía jurisdiccional para la eliminación de infracciones (con las garantías que ello supone para los derechos de los usuarios) y se incentiva la autoregulación, garantizando un papel neutral de las plataformas.⁴⁰ Este es, por ejemplo, el modelo previsto en la legislación de propiedad intelectual vigente en Canadá.⁴¹

En el apartado IV.3. se tratará el sistema de *notice and counter-notice*, en el cual se ofrece la facultad de réplica del usuario proveedor de contenidos, permitiendo a la plataforma decidir sobre el bloqueo y brindando la posibilidad de volver a habilitar el contenido ya bloqueado.

c) Sistemas de filtrado *ex ante*. El art. 17 de la Directiva CDSM

No cabe duda de que las plataformas digitales de contenidos generados por los usuarios tienen en esta actividad su negocio. Es una evidencia que YouTube, Instagram, TikTok, entre muchos otros, obtienen sus nada modestos ingresos gracias a la explotación indirecta de estos contenidos. En el ámbito del derecho de autor, los titulares de derechos —grandes y pequeños—, además de utilizar estas plataformas como un medio más para la explotación comercial de sus obras y prestaciones, siempre han recelado de las mismas por ser el medio idóneo para la subida no autorizada de contenidos protegidos por parte de terceros usuarios. Muchas veces, siendo imposible cobrarse los daños directamente de los usuarios infractores, por su anonimato o estado de insolvencia. Sobran ejemplos de conflictividad judicial entre Google (la empresa titular de YouTube) y productoras de TV, cine, musicales, etc. sobre contenidos protegidos subidos por los usuarios.

La plasmación de estos celos la encontramos en la aprobación del art. 17 de la Directiva CDSM que aclara la interpretación jurisprudencial que el TJUE había hecho anteriormente en relación con el grado de responsabilidad de las plataformas sobre las infracciones de propiedad intelectual en internet.⁴² La directiva da una redefinición de

⁴⁰ ANGELOPOULUS – SMET. *Notice-and-Fair-Balance...* p. 25.

⁴¹ Ver sec. 47 de la Canadian Copyright Modernization Act.

⁴² Todavía a día de hoy existen como mínimo dos cuestiones prejudiciales en el TJUE sobre la cuestión de si las plataformas —con sus particulares características— realizan actos de comunicación al público por los contenidos subidos por sus usuarios: C-682/18 (YouTube), C-683/18 (Elsevier). Además, está pendiente otra cuestión prejudicial (también sobre YouTube) sobre la pérdida de la exención de responsabilidad del art. 14.1 de la Directiva e-commerce por realizar actividades mucho más activas que

proveedor de servicios específica para aquellos que almacenen y publiquen contenido subido por sus usuarios (“prestadores de servicios para compartir contenidos en línea”) y establece que siempre se considerará que estos realizan actos de comunicación al público —es decir, serán considerados responsables directos en caso de infracción— cuando ofrezcan al público “el acceso a obras protegidas por derechos de autor u otras prestaciones protegidas que hayan sido cargadas por sus usuarios.”⁴³ Además, se excluye de forma expresa la aplicación de la exención de responsabilidad del art. 14 de la Directiva e-commerce para estos casos.⁴⁴ En particular, se exige como condición para gozar de la exención de responsabilidad haber realizado los máximos esfuerzos para obtener una autorización del titular (art. 17.4.b) de la Directiva CDSM) y actuar “de modo expeditivo al recibir una notificación [...] de los titulares de derechos, para inhabilitar el acceso a las obras”, además de hacer “los mayores esfuerzos por evitar que se carguen en el futuro” (art. 17.4.c)), algo cuyo alcance deberá precisar el TJUE en el futuro.

Esta regulación específica, equivale prácticamente a un régimen de responsabilidad objetiva de las plataformas por infracciones de propiedad intelectual, que son consideradas responsables incluso si no ha habido notificación inicial y retirada posterior, de forma que resultará imprescindible contar con un sistema de filtrado de contenidos automatizado que impida *ex ante* la subida de contenidos protegidos sin licencia.⁴⁵ Esto comporta riesgo de *overblocking*. Si bien es verdad que no es un problema exclusivo de los sistemas automatizados de control de contenidos —también pueden producirse casos de eliminación ilegítima de contenidos llevada a cabo por humanos— parece evidente que el uso de la tecnología puede agravar la situación, ya que el grado de sofisticación actual de los programas de detección automática de contenido ilegal todavía está lejos de lograr una exactitud del 100%.

El objetivo perseguido por el legislador europeo es un modelo de plataforma de contenido licenciado que los usuarios usan activa y libremente, como el modelo

el mero almacenamiento de contenidos ajenos: C-500/19 (Puls). Sobre actos de comunicación al público por parte de prestadores de servicios ver SSTJUE C-466/12 (Svensson), C-348/13 (Bestwater International), C-160/15 (GS Media), C-527/15 (Filmspeler), C-610/15 (The Pirate Bay).

⁴³ Art. 17.1 Directiva CDSM.

⁴⁴ Art. 17.3 Directiva CDSM.

⁴⁵ Así mismo estaba expresado en el redactado de la propuesta inicial de la Directiva (*effective content recognition technologies*), aunque posteriormente se eliminó la referencia por las críticas que estaba recibiendo durante el proceso de aprobación. Art. 13 COM(2016) 593 final.

“TikTok”. TikTok (ex musical.ly) es una plataforma de *user generated content* que permite de forma sencilla la edición y sincronización de música en vídeos breves (de un máximo de 15 segundos), principalmente enfocados al humor. A través de la plataforma, se puede interactuar con vídeos de otros usuarios con la música más actual, ya que esta cuenta con licencias de uso de los principales sellos discográficos.⁴⁶ Aun así, un resultado plausible de esta regulación será la implantación un modelo cada vez menos abierto a los usuarios y más similar a las plataformas clásicas de contenido licenciado, como las plataformas *on demand* Netflix o Spotify.⁴⁷ Especialmente, porque la posibilidad de obtener una licencia de uso de absolutamente todos los contenidos que puedan subir los usuarios es simplemente una utopía irrealizable en la forma que está configurado el derecho de autor. Es verdad que las plataformas siempre introducen una licencia a su favor (mundial, no exclusiva, gratuita y con facultad de cesión a terceros) en las condiciones generales de la contratación con sus usuarios, pero esto en absoluto es garantía de que el usuario es el legítimo titular de los derechos. Ante esta situación, parece mucho más realista para la viabilidad económica de una plataforma permitir solamente la subida de contenidos a operadores de acreditada solvencia que dar cumplimiento estricto a los deberes de diligencia contenidos en el art. 17.4 de la Directiva con el contenido subido espontáneamente por cualquiera, lo cual perjudica en última instancia a usuarios y creadores independientes.⁴⁸

De la aplicación de este artículo, además, se han señalado en alguna ocasión posibles efectos lesivos para la competencia:

- a) Por entorpecer la difusión de aquellos contenidos autoproducidos por usuarios sin una industria solvente y consolidada detrás. Plataformas como YouTube no son de mero consumo, sino de participación activa, y ofrecen —por lo menos idealmente— la posibilidad de construir con pocos recursos contenidos interesantes y alternativos a las productoras tradicionales y, al fin y al cabo, si el

⁴⁶WIRED. “TikTok Is a Short-Form Monetized Musical Meme Machine”. Disponible en: <https://www.wired.com/story/tiktok-musical-meme-machine-vine/>

⁴⁷ PEGUERA, Miquel. “La controvertida reforma del copyright en la UE”. Blog de los estudios de derecho y ciencia político de la UOC, 2019. <http://edcp.blogs.uoc.edu/20190220-la-controvertida-reforma-del-copyright-en-la-ue/>

⁴⁸ KREUTZER, Till. “The EU copyright directive and its potencial impact on cultural diversity on the internet”. Kluwer Copyright Blog. 2020.

derecho de autor tiene alguna función es la de estimular la creación y contribuir al desarrollo de la cultura en beneficio y disfrute de todos los ciudadanos.⁴⁹

- b) Por obligar tácitamente a la competencia⁵⁰ de las grandes plataformas como Google a licenciar la tecnología de filtrado que ellos han desarrollado previamente por la imposibilidad material de desarrollar una propia,⁵¹ reforzando la posición de dominio en el mercado que ya tienen estas empresas. Es cierto que la propia Directiva incluye en su art. 17.6 una exención a la obligación de filtrado para aquellas empresas que no lleguen a un determinado volumen de negocios, a saber, aquellas que lleven menos de tres años operando y cuyo volumen de negocios anual sea inferior a 10 millones de euros. Sin embargo, quedarán fuera de la exención muchísimas empresas en línea que en absoluto representan una competencia real a las grandes plataformas digitales como YouTube.⁵²

El legislador europeo, consciente de los riesgos que presenta este sistema y del incentivo que esto da a las plataformas para eliminar todo tipo de contenidos (lícitos e ilícitos) para evitar ser consideradas responsables, reitera la obligatoriedad de respetar los límites al derecho de autor (art. 17.7) y de establecer mecanismos efectivos de apelación (17.9), algo que se tratará en los apartados siguientes.

IV. SALVAGUARDAS PARA LOS DERECHOS DE LOS USUARIOS

Hemos visto hasta ahora la enorme dificultad que presenta la detección de contenidos ilícitos en internet y los riesgos de vulneración de los derechos de los usuarios por causa del *overblocking*. A continuación se propone una valoración de los elementos que podrían contribuir a un equilibrio efectivo de los derechos de las partes involucradas. En especial, la configuración de los mecanismos que deberían permitir a los usuarios a) conocer con total transparencia cómo aplican las plataformas sus

⁴⁹ Ver en esta dirección la exposición de motivos de la antigua Ley de Propiedad Intelectual de 1987 *in fine*.

⁵⁰ SPOERRI, Thomas. "On Upload-Filters and other Competitive Advantages for Big Tech Companies under Article 17 of the Directive on Copyright in the Digital Single Market". JIPITEC 10 (2) 2019, nº 52 y 56.

⁵¹ Un informe estima que Google ha invertido más de 60 millones de dólares en el desarrollo del sistema de reconocimiento y filtrado "Content ID": https://es.scribd.com/document/308605292/DMCA-Universal-Music-Group-First-Round-Comments-pdf?ad_group=66960X1514734Xd8e5e3075531624685c5d3c959f77b57&campaign=SkimbitLtd&keyword=660149026&medium=affiliate&source=hp_affiliate

⁵² SPOERRI. "On Upload-Filters..." nº 22.

políticas de bloqueo de contenidos⁵³ (sobre todo cuando las decisiones se basan en inteligencia artificial); b) apelar tales decisiones mediante un mecanismo sencillo e intuitivo, que dé garantías de obtener una resolución fundada e imparcial.⁵⁴ Un buen desarrollo de la regulación de estos elementos —algunos ya previstos, como en el art. 17.9 de la Directiva CDSM— debería facilitar al usuario la posibilidad de reinstaurar en poco tiempo el contenido lícito en aquellos casos de bloqueo o filtrado erróneo, así como ayudar a prevenir la denuncia abusiva o temeraria de terceros interesados.

1. La notificación al proveedor

Al hablar de notificación en los distintos sistemas de actuación de las plataformas, se hace referencia al mecanismo que tiene la persona víctima de la infracción cometida por quien pone a disposición del público contenidos ilegales, pero también cualquier tercero interesado en que estos desaparezcan, de comunicar la existencia de un contenido ilícito a una plataforma, de forma que esta no pueda ignorarlo sin faltar a sus deberes de actuación, por ejemplo de actuar con prontitud para retirarlo.⁵⁵ De otro modo, esperar a obtener una resolución judicial o administrativa sería en muchos casos demasiado lento para los afectados.⁵⁶

La sencillez para la denuncia efectiva de contenidos ilícitos —que debe valorarse positivamente—, puede resultar en situaciones de abuso o temeridad por parte de los denunciantes, especialmente cuando las plataformas cuentan con formularios automatizados para ello.⁵⁷

⁵³ En la misma línea, Recomendación de la Comisión 2018/334, nº 16 y 17.

⁵⁴ En esta línea, Recomendación de la Comisión 2018/334, nº 11 a 13, HOUSE OF LORDS. Select Committee on Communications. 2nd Report of Session 2017-2019. “Regulating in a digital world”. Londres, 2019 y KUCZERAWY, Aleksandra. “The Power of Positive Thinking: Intermediary Liability and the Effective Enjoyment of the Right to Freedom of Expression”, 8 (2017) JIPITEC, nº 32 a 36.

⁵⁵ Ver el Considerando 46 y art. 14.1.a de la Directiva e-commerce.

⁵⁶ ARROYO. “La responsabilidad...”, nº 30.

⁵⁷ En general las plataformas suelen distinguir entre la denuncias por spam, por contenido abusivo u ofensivo, o por afectar a un derecho de propiedad intelectual o industrial. Ver, a modo de ejemplo, el procedimiento que ofrece Twitter <https://help.twitter.com/es/rules-and-policies/twitter-report-violation#specific-violations>

a) Prevención contra el abuso y la temeridad de los denunciantes

Prevenir el comportamiento abusivo o de mala fe en las denuncias de contenido lícito exige un sistema que permita sancionar este comportamiento.⁵⁸ Esta penalización puede tener carácter económico:⁵⁹ por ejemplo, una multa o el pago de las costas que se generen del procedimiento de restitución del contenido (sea en sede judicial o extrajudicial). También suele destacarse la necesidad de que el notificante de mala fe responda por los daños derivados de la remoción. Otra posibilidad que apunta la Comisión Europea es la postergación en atender las denuncias de aquellos que notifiquen un gran número de contenido lícito o que reciban un gran número de apelaciones,⁶⁰ es decir, que otras tengan prioridad. La concreción de estas medidas debería ir acompañada de un deber de transparencia de las plataformas sobre el funcionamiento de las mismas y sobre los datos de las notificaciones atendidas, que permita evaluar si son efectivas.⁶¹

Además, el notificante debería asumir en la formulación de la denuncia un cierto grado de concreción el fundamento que justifique la petición de retirada y bloqueo, así como la determinación exacta del contenido lesivo y su ubicación.⁶² La opción de pronunciarse expresamente sobre la necesidad de remoción del contenido debería venir aparejada con la penalización en caso de reclamación fraudulenta, y las plataformas así deberían comunicarlo en los formularios de notificación para prevenir un posible uso abusivo de los mismos. Es especialmente interesante, en este sentido, la propuesta de elaboración de una “hoja de notificación” armonizada en el ámbito europeo⁶³ que concrete estos deberes.

b) Legitimación y anonimato

Por último, está sobre la mesa la cuestión del quién debería poder notificar y si debería poder hacerse anónimamente. En principio, no parece que nada justifique exigir un

⁵⁸ En este sentido, la Recomendación 2018/334 de la Comisión, nº 21

⁵⁹ ARROYO. “La responsabilidad...”, nº 38.

⁶⁰ COM(2017) 555 final, p. 18.

⁶¹ *Ibid.*

⁶² Recomendación 2018/334, nº 6.

⁶³ ARROYO. “La responsabilidad...” nº 37.

especial interés personal presentar una notificación, así que cualquier tercero podría hacerlo.⁶⁴

También debe tenerse en cuenta que la carga de responsabilidad del notificante viene aparejada a la necesidad de determinación de la identidad de este. Los efectos del anonimato en la notificación equivaldrían en muchos casos a la imposibilidad de ejercer cualquier tipo de acción contra su uso abusivo. Sin embargo la Comisión⁶⁵ es del parecer que el requisito de identificación puede constituir un obstáculo a la hora de denunciar o notificar las infracciones. Una solución intermedia sería estimular la creación de entidades jurídicas que presenten en su propio nombre notificaciones de terceros (que se mantendrían en el anonimato). Esta función la podrían ejercer los notificantes fiables de los que habla la Recomendación 2018/334 de la Comisión.⁶⁶ Esto, si bien podría retrasar el procedimiento, añadiría una garantía de revisión del abuso por parte de la entidad, ya que en última instancia esta sería la responsable de su presentación.

2. La comunicación de las medidas adoptadas

Un momento de especial sensibilidad se da en la comunicación al usuario de la adopción, por parte de la plataforma, de unas determinadas medidas, que pueden consistir en la remoción de contenido, la suspensión de cuentas, una aparición menor en el *feed* de los seguidores, etc. En este sentido, es preciso de un deber de las plataformas para realizar una comunicación transparente que informe al sujeto restringido de los motivos por los cuales la decisión ha sido tomada.⁶⁷ Este deber de transparencia no puede limitarse a una remisión genérica a unos términos de uso poco concretos o incomprensibles. Por el contrario, debería permitir identificar sin equívoco cual es la condición vulnerada y qué elementos fundamentan la decisión (en caso de tratarse de una decisión automatizada, qué elementos ha considerado el sistema para determinarlo). El derecho europeo no es ajeno a esta noción de transparencia, otra manifestación de esta necesidad la podemos encontrar en el Considerando 71 del Reglamento 2016/679, General de Protección de Datos (RGPD), respecto al tratamiento transparente de datos personales, que se ha venido llamando *right to explanation*: “dicho tratamiento debe estar sujeto a las garantías apropiadas, entre las

⁶⁴ Ibid., nº35.

⁶⁵ Recomendación 2018/334, nº 7.

⁶⁶ Nº 4 g) y 25 y ss.

⁶⁷ Ver la Recomendación de la Comisión 2018/334, nº 9.

que se deben incluir la información específica al interesado y el derecho a obtener intervención humana, a expresar su punto de vista, a recibir una explicación de la decisión tomada después de tal evaluación y a impugnar la decisión”. Además, deberían procurar evitarse malas prácticas de diseño como los *dark patterns*.⁶⁸ Este deber de transparencia es especialmente relevante para posibilitar de forma efectiva un mecanismo de oposición a una decisión tomada por la plataforma.

3. Mecanismos de oposición, apelación y reinstauración de contenidos

Una regulación proporcional también debería exigir una explicación clara y transparente de los mecanismos que se ponen a disposición del usuario de la plataforma para apelar las decisiones tomadas por esta. La mayoría de ellas ya incluyen alguna forma más o menos formal de comunicarles los errores en su aplicación de medidas. Sin embargo, suele tratarse de meras informaciones sin ninguna relevancia práctica. Ninguna de ellas incluye información sobre si van a atender la petición de restaurar el contenido ni en qué plazos, ni con base en qué criterios, ni si va a resolver la oposición a la remoción un sistema de inteligencia artificial, ni qué otras vías existen para hacer valer los derechos del usuario. Sin toda esta información, la puesta a disposición de un mecanismo de apelación es papel mojado. Por otra parte, un aspecto fundamental a valorar a la hora de regular un mecanismo de estas características es la rapidez. Las prácticas en internet, sean creativas, recreativas o profesionales, tienen como característica común la inmediatez.⁶⁹ En un modelo social de comunicación donde las tendencias duran cada vez menos, es vital que un mecanismo de oposición a las decisiones tomadas por las plataformas sea también rápido y eficaz, de la misma forma que deben serlo los mecanismos para la protección del derecho al honor, la intimidad, la imagen o los derechos de autor.⁷⁰

⁶⁸ Los *dark patterns* son prácticas en el diseño de interfaces digitales y de experiencia de usuario (UI/UX) que llevan al usuario a realizar acciones que no le benefician o impiden que realice otras que perjudican a la plataforma. Más información en WAGNER, Ben - ROZGONYI, Krisztina - SEKWENZ, Marie-Therese - COBBE, Jennifer - SINGH, Jatinder. “Regulating Transparency? Facebook, Twitter and the German Network Enforcement Act”. FAT* 27-30, 2020. Pág. 268 y 269. Además, www.darkpatterns.org “How Dark Patterns Trick You Online” https://www.youtube.com/watch?time_continue=127&v=kxkrdLI6e6M&feature=emb_title

⁶⁹ SUZOR, Nicolas. *Lawless: The Secret Rules That Govern Our Digital Lives*. Cambridge University Press, 2019, p. 69, habla de unas dos semanas como máximo para la vigencia de la mayoría de contenidos digitales. Una reposición de un contenido más allá de este plazo podría dejarlo fuera de su vida útil, con los costes en beneficio de oportunidad que esto supone para sus creadores.

⁷⁰ Ver el Considerando 52 de la Directiva e-commerce.

Hasta la fecha, la mayoría de plataformas han seguido —por lo menos en relación con las reclamaciones de derechos de autor— el modelo de la contranotificación regulado de la Sec. 512 (g)(2) de la DMCA estadounidense.⁷¹ El mecanismo es relativamente sencillo: ante la petición de remoción del notificante, el afectado puede argumentar porqué cree que su contenido es lícito y debe permanecer accesible en la plataforma. La plataforma debe trasladar la apelación presentada al notificante, que tendrá 10 días para interponer la oportuna demanda judicial y presentarla para evitar que el contenido vuelva a ser habilitado.

La Recomendación 2018/334 de la Comisión⁷² adopta el modelo de contranotificación como garantía ante el riesgo de *overblocking*, en especial cuando se lleva a cabo a través de medios automatizados. Sin embargo, no precisa un plazo en concreto para la rehabilitación de contenidos:

19. Para evitar la retirada de contenidos lícitos, sin perjuicio de la posibilidad de que los prestadores de servicios de alojamiento establezca y aplique sus condiciones de prestación de servicios de conformidad con la normativa de la Unión y la legislación de los Estados miembros, deben existir salvaguardias eficaces y apropiadas para garantizar que los prestadores de servicios de alojamiento actúen con diligencia y de forma proporcionada respecto de los contenidos que almacenan, en particular al gestionar las notificaciones y contranotificaciones y al decidir sobre la posible retirada o el bloqueo de los contenidos considerados ilícitos.

Lo bueno de estos sistemas es que prevén de forma muy clara los plazos en los que deben resolverse las alegaciones de los usuarios, algo especialmente interesante cuando se trata de plazos muy breves. El problema es que los usuarios tienden a no presentar oposición por la posibilidad de tener que responder personalmente en el caso de no estar bien fundadas, de manera que el sistema es ineficaz en la mayoría de los casos.⁷³

También la Directiva CDSM prevé en el artículo 17.9.I que las plataformas “establezcan un mecanismo de reclamación y recurso ágil y eficaz que esté a disposición de los usuarios de sus servicios en caso de litigio sobre la inhabilitación del acceso a obras u otras prestaciones cargadas por ellos o sobre su retirada.” En este sentido, además de estudiar la posibilidad de extrapolar esta obligación al resto de ilícitos que no tienen

⁷¹ Ver por ejemplo el protocolo que utiliza YouTube respecto a la contranotificación: https://support.google.com/youtube/answer/2807684?hl=es&ref_topic=9282678

⁷² Y, en la misma línea, la Comunicación COM(2017) 555 final, pp. 16 y 17.

⁷³ KUCZERAWY. *Intermediary Liability...* p. 232.

que ver con la infracción de derechos de autor, corresponde al legislador nacional, a la hora de transponer la Directiva, hacer lo posible para concretar estos deberes de forma que no se conviertan en un mecanismo ineficaz.⁷⁴ Para esto, debería reservarse la remoción directa (es decir, sin escuchar al usuario) para aquellos casos en que la infracción resulte obvia o se trate de la obra protegida idéntica o entera. En el resto de casos, debería mantenerse el contenido en línea y dar la oportunidad al usuario que ha cargado estos contenidos de alegar las razones por las que conviene mantenerlos.⁷⁵ Además, dichos mecanismos de apelación deberían poder hacerse a través de un formulario accesible y fácil, que incluyera *check boxes* (casillas de verificación) con los motivos de oposición más frecuentes,⁷⁶ de forma que no fuera necesario un gran asesoramiento para ejercitarlos. La plataforma tampoco debería responder por la posible ilicitud de los contenidos que difunde durante el tiempo que dure este procedimiento una vez iniciado, hasta la efectiva determinación de su legalidad.⁷⁷

4. Autorregulación y arbitraje

Otra cuestión destacada en el contexto de los mecanismos de salvaguarda de los derechos fundamentales en internet es la autorregulación y los medios alternativos o extrajudiciales de resolución de conflictos. Esto incluye el fomento de la implantación de códigos de conducta en los principales agentes del sector y la promoción del arbitraje en la resolución de disputas derivadas de contenido en línea.⁷⁸

Llegados a este punto, conviene diferenciar, en primer lugar, la resolución de conflictos que puedan surgir del borrado de contenido derivado de una notificación de un tercero y, en segundo lugar, la resolución de conflictos derivados de decisiones adoptadas por las plataformas. En todos estos casos, la implantación efectiva de estos

⁷⁴ Grupo de Investigación DRET D'INTERNET (UOC). "Contribución a la consulta pública previa sobre la incorporación al ordenamiento jurídico español de la Directiva (UE) 2019/790 de 17 de abril de 2019", 2019, p. 8. Disponible en: <http://edcp.blogs.uoc.edu/wp-content/uploads/2019/12/Contribuci%C3%B3n-a-la-consulta-p%C3%BAblica-Directiva-DADMUD.pdf>

⁷⁵ QUINTAIS, João - FROSIO, Giancarlo - VAN GOMPEL, Stef - HUGENHOLZ, P. Bernt - HUSOVEC, Martin - JÜTTE, Bernd Justin - SENFTLEBEN, Martin. "Safeguarding User Freedoms in Implementing Article 17 of the Copyright in the Digital Single Market Directive: Recommendations From European Academics". 2019, p. 4.

⁷⁶ *Ibid.* p. 5.

⁷⁷ *Ibid.* p. 5.

⁷⁸ Ver el art. 17.9.II de la Directiva CDSM, la Recomendación 2018/334 de la Comisión, nº 14 y 15 y la Comunicación COM(2017) 555 final, p. 17 *in fine*.

recursos supondría un ahorro sustancial de recursos (tiempo y dinero) para las partes implicadas y en beneficio del sistema de justicia, ya muy colapsado. La rapidez en la resolución y el ahorro de tiempo cobra especial sentido en un contexto como el de internet, en el que la validez de los contenidos es muy breve y donde la inmediatez juega un papel esencial. Además, ofrecería garantías de imparcialidad. La independencia del órgano que resuelve este tipo de controversias es también un elemento de gran importancia. Parece claro que las propias plataformas —ya sea voluntariamente o sin querer— no suelen actuar de forma imparcial. En el ejemplo de los derechos de autor, es evidente la interdependencia vigente entre las plataformas (que necesitan que los contenidos de la gran industria estén presentes en sus páginas) y los titulares de derechos (que necesitan de los canales de máxima audiencia, sean los que sean). En otros casos en los que la ideología tiene un papel más importante, la capacidad de los proveedores de servicios para ser imparciales se ha demostrado todavía menor.

Una posibilidad en este sentido sería vincular el beneficio de exención responsabilidad a las plataformas que estuvieran suscritas a un sistema de resolución de conflictos por parte de un órgano independiente, que se ocupara de la resolución de las reclamaciones de los usuarios basándose en la ley y los códigos de conducta correspondientes, si los hubiera. Las plataformas deberían poder asumir la cuota de suscripción al organismo regulador o de resolución, mientras que podría penalizarse la reclamación u oposición temerarias de los titulares o usuarios con el pago del coste del procedimiento, siempre menor que el judicial. A modo de ejemplo, el procedimiento de reclamación de Autocontrol (una entidad autorreguladora en el ámbito de la publicidad), que tiene un coste gratuito para los consumidores y de 265€ para los socios. Además, sería recomendable un plazo de resolución máximo de entre 10 y 15 días dependiendo del ilícito de que se trate, de acuerdo con el tiempo aproximado de vigencia de gran parte del contenido en línea.

Por supuesto, todo lo anterior debe entenderse sin perjuicio de las posibilidades de acceder a la justicia en el mismo sentido que ya prevén los arts. 12.3, 13.2, 14.3 y 18 de la Directiva e-commerce.

V. CONCLUSIONES

Cualquier usuario habitual de redes sociales habrá advertido en los últimos tiempos un notable aumento del control que ejercen las principales plataformas digitales (así, Instagram, Facebook, YouTube, Twitter y TikTok) sobre los contenidos que alojan creados por sus usuarios. Infracciones de propiedad intelectual, apología del terrorismo, infracción de las normas de uso, discurso del odio, vulneraciones al honor,

suelen ser los motivos que alegan para, o bien no permitir la publicación un contenido, o bien para eliminarlo con posterioridad a su publicación.

También es fácil de percibir la enorme cantidad de casos en los que estas medidas resultan totalmente desproporcionadas o claramente erróneas, sin que exista un mecanismo efectivo para cuestionarlas. En la creciente tendencia legislativa a exigir a las plataformas mayores esfuerzos para la erradicación de los ilícitos en línea (como el art. 17 de la Directiva CDSM) no parece que vayamos a encontrar mejoras.

Lograr un equilibrio entre los derechos de los afectados (así, entre otros, las víctimas de vulneraciones de derechos de la personalidad o los titulares de derechos de autor) y los usuarios activos de las plataformas debe pasar por regular un mecanismo de apelación de las decisiones que tenga estas características: que sea fácil y rápido de utilizar, que exija unos altos niveles de transparencia, que sancione el abuso, la mala fe y la temeridad de los denunciantes de contenido lícito y que tenga garantías de resolución imparcial y adecuada a derecho, ya sea en vía judicial o mediante el arbitraje de terceros ajenos a la empresa. En mi opinión, esta es la única forma de garantizar derechos y libertades fundamentales de los usuarios, como la libertad de expresión o artística, por ser en muchas ocasiones demasiado gravoso acudir a los tribunales en defensa de los mismos y para no generalizar una actitud de pasividad y resignación ante pequeñas (¡pero importantes!) vulneraciones de estos.

La transposición nacional del art. 17 de la Directiva CDSM es una buena oportunidad para regular con detalle estas medidas para que tengan la eficacia pretendida y no se conviertan en papel mojado.

VI. BIBLIOGRAFÍA

ANGELOPOULUS, Christina – SMET, Stijn. *Notice-and-Fair-Balance: How to Reach a Compromise between Fundamental Rights in European Intermediary Liability*. Journal of Media Law, Vol. 8, 2016, pp. 266-301.

ARROYO AMAYUELAS, Esther. “La responsabilidad de los intermediarios en internet. ¿Puertos seguros a prueba de futuro?”. CDT, 2020, Vol. 12, Nº 1, pp. 808-837.

BUSTO, José Manuel. “La responsabilidad civil de los prestadores de servicios de la sociedad de la información (ISPs)”. En REGLERO CAMPOS, Fernando [coord.] *Tratado de Responsabilidad Civil*. Aranzadi, 2003.

DRET D'INTERNET (UOC). *Contribución a la consulta pública previa sobre la incorporación al ordenamiento jurídico español de la Directiva (UE) 2019/790 de 17 de abril de 2019*, 2019. Disponible en: <http://edcp.blogs.uoc.edu/wp-content/uploads/2019/12/Contribuci%C3%B3n-a-la-consulta-p%C3%ABblica-Directiva-DADMUD.pdf>

FROSIO, Giancarlo. *Reconciling Copyright with Cumulative Creativity*. Edwar Elgar. 2018.

FROSIO, Giancarlo. "The Death of 'No Monitoring Obligations': A Story of Untameable Monsters" JIPITEC 8, 2017.

GARROTE FERNÁNDEZ-DÍEZ, Ignacio. "¿Puede crearse un nuevo límite en la Ley de Propiedad Intelectual española para dar cobertura a los contenidos generados por los usuarios?" En MARÍN LÓPEZ, J. J. – CASAS VALLES, Ramon (coords.). *Estudios sobre la ley de propiedad intelectual: últimas reformas y materias pendientes*. Dyckinson. Madrid, 2016.

GINSBURG, Jane C. "Online Service Providers liability in the EU and USA: Art. 17 Directive CDSM and Sec. 512 USCA." Seminario en la Universitat Oberta de Catalunya. Barcelona, 2019 (inédito).

HILTY, Reto – KÖKLÜ, Kaya. "Limitations and Exceptions to Copyright in the Digital Age. Four Cornerstones for a Future-Proof Legal Framework in the EU", en: STAMATOUDI, Irini [ed.] *New Developments in EU and International Copyright Law*. Wolters Kluwer, 2016.

KOVACS, Andrea. *Die Haftung der Host-Provider für persönlichkeitsrechtsverletzende Internetäußerungen*. Nomos. 2018.

KUCZERAWY, Aleksandra. "From 'Notice and Take Down' to 'Notice and Stay Down': Risks and Safeguards for Freedom of Expression". En FROSIO, Giancarlo (ed). *The Oxford Handbook of Intermediary Liability Online* (próxima publicación, 2019).

KUCZERAWY, Aleksandra. *Intermediary Liability and Freedom of Expression in the EU: from Concepts to Safeguards*. Intersentia, KU Leuven Centre for IT & IP Law Series. 2019.

KUCZERAWY, Aleksandra. "The Power of Positive Thinking: Intermediary Liability and the Effective Enjoyment of the Right to Freedom of Expression", 8 (2017) JIPITEC. Nº 36.

KREUTZER, Till. "The EU copyright directive and its potencial impact on cultural diversity on the internet". Kluwer Copyright Blog. 2020. Disponible en:

<http://copyrightblog.kluweriplaw.com/2020/01/22/the-eu-copyright-directive-and-its-potential-impact-on-cultural-diversity-on-the-internet-part-i/>

PEGUERA, Miquel [coord.]. *Derecho y nuevas tecnologías*. Editorial UOC. Barcelona, 2005.

PEGUERA, Miquel [coord.]. *Principios de derecho de la sociedad de la información*. Cap. 6: Marco jurídico de los servicios de la Sociedad de la Información y del Comercio electrónico. Aranzadi. 2010.

PEGUERA, Miquel. *The DMCA Safe Harbors and Their European Counterparts: A Comparative Analysis of Some Common Problems*. Columbia Journal of Law & the Arts, Vol. 32, pp. 481. 2009. Disponible en: <https://ssrn.com/abstract=1468433>

QUINTAIS, João – FROSIO, Giancarlo – VAN GOMPEL, Stef – HUGENHOLZ, P. Bernt – HUSOVEC, Martin – JÜTTE, Bernd Justin – SENFTLEBEN, Martin. “Safeguarding User Freedoms in Implementing Article 17 of the Copyright in the Digital Single Market Directive: Recommendations From European Academics”. 2019. Disponible en SSRN: <https://ssrn.com/abstract=3484968>

RIIS, Thomas – PETERSEN, Clement. “Private enforcement of IP law by internet service providers: notice and action procedures”. En RIIS, Thomas. *User Generated Law: Re-Constructing Intellectual Property Law in a Knowledge Society*. Edwar Elgar. 2016.

RUBÍ, Antoni. “Autorregulació publicitària. Autocontrol”. Ponencia en: Curs de Dret de la Publicitat. ICAB, 30 de octubre de 2019 (inédito).

SERRANO, Eduardo. “Los prestadores de servicios para compartir en línea contenidos cargados por los usuarios” En SERRANO, Eduardo (dir.) *Medios de comunicación, contenidos digitales y derecho de autor*. Editorial Reus, 2019.

SPINDLER, Gerald. “Internet Intermediary Liability Reloaded. The New German Act on Responsibility of Social Networks and its (In-) Compatibility with European Law”. JIPITEC 8. 2017.

SPINDLER, Gerald. “¿Compartir contenidos en línea o no? Un análisis del artículo 17 de la Directiva sobre propiedad intelectual en el mercado único digital”, CDT, 2020, Vol. 12, Nº 1, pp. 962-1002.

SPOERRI, Thomas. “On Upload-Filters and other Competitive Advantages for Big Tech Companies under Article 17 of the Directive on Copyright in the Digital Single Market”. JIPITEC 10 (2) 2019.

SUZOR, Nicolas. *Lawless: The Secret Rules That Govern Our Digital Lives*. Cambridge University Press. Queensland, 2019.

TADDEO, Mariarosaria – FLORIDI, Luciano [eds.] *The Responsibilities of Online Service Providers*. Springer. 2017.

WOLLIN, Sören. *Störerhaftung im Immaterialgüter- und Persönlichkeitsrecht: Zustandshaftung analog § 1004 I BGB*. Nomos. Baden-Baden, 2018.

XALABARDER, Raquel. “Intermediarios tecnológicos y bases para la responsabilidad. Intervención voluntaria u obligatoria”. En *Difusión y gestión de las obras protegidas por el derecho de autor en internet*. ALAI. Cartagena, 2016.

XALABARDER, Raquel. “La responsabilidad de los prestadores de servicios en Internet (ISP) por infracciones de propiedad intelectual cometidas por sus usuarios”. *Revista IDP*, 2006.

YZQUIERDO, Mariano – ARIAS, Vicente. “Responsabilidad civil por daños a la propiedad intelectual”. En REGLERO CAMPOS (coord.) *Tratado de responsabilidad Civil*. Aranzadi, 2008.