

LLIBERTAT D'EXPRESSION I PLATAFORMES DIGITALS: REPTES DE LA REGULACIÓ DE CONTINGUTS EN UN ENTORN GLOBAL

Joan Barata Mir*

Resum

Els drets fonamentals a la llibertat d'expressió i a la llibertat d'informació es troben reconeguts, protegits i delimitats en la pràctica totalitat de les constitucions dels sistemes polítics de les democràcies liberals, entre elles la Constitució espanyola. Aquests drets formen igualment part del quadre fonamental de drets humans establert per les normes de dret internacional. Els intermediaris s'han convertit en els principals actors del procés de difusió i distribució de tot tipus de contingut. També tenen un paper molt destacat a l'hora d'assegurar la "visibilitat" del contingut produït pels mitjans tradicionals. Els interessos que modulen les decisions preses pels intermediaris no són únicament i necessàriament la protecció del dret a la llibertat d'expressió. Els intermediaris són entitats orientades comercialment i segueixen models de negoci basats en la venda de publicitat, entre altres aspectes. Els sistemes legals amb més influència en el món (particularment els de la Unió Europea i els Estats Units) incentiven la moderació del contingut per part de les plataformes, garantint que no són penalitzades com a conseqüència del contingut subministrat per tercers o a causa de l'adopció de determinades mesures orientades a eliminar contingut il·legal o contingut inadequat. Per aquests motius, els intermediaris tenen un paper molt important en la creació de les condicions per a l'exercici del dret a la llibertat d'expressió en línia.

Paraules clau: llibertat d'expressió; regulació de continguts; responsabilitat d'intermediaris; autoregulació de mitjans.

FREEDOM OF EXPRESSION AND DIGITAL PLATFORMS: CHALLENGES FOR CONTENT REGULATION IN A GLOBAL ENVIRONMENT

Abstract

The fundamental rights to freedom of expression and freedom of information are recognized, protected and delimited in practically all the constitutions of the political systems of liberal democracies, including the Spanish Constitution. These rights are also part of the human rights framework established by the rules of international law. Intermediaries have become the main actors in the process of disseminating and distributing all types of content. They also play a very important role in ensuring the "visibility" of content produced by traditional media. The interests that modulate the decisions taken by intermediaries are not solely and necessarily the protection of the right to freedom of expression. Intermediaries are commercially oriented entities and follow business models based on advertising, among other things. The most influential legal systems in the world (particularly those in the European Union and the United States) encourage the moderation of content by platforms, ensuring that they are not penalized as a result of content provided by third parties or due to the adoption of certain measures aimed at eliminating illegal or objectionable content. For these reasons, intermediaries play a very important role in creating the conditions for the exercise of the right to freedom of expression online.

Key words: freedom of expression; content regulation; liability of intermediaries; self-regulation of media.

* Joan Barata Mir, investigador associat al Cyber Policy Center de la Universitat de Stanford, ex-secretari general del Consell de l'Audiovisual de Catalunya i ex-assessor principal de la Representant de Llibertat de Mitjans a l'OSCE, Freeman Spogli Institute for International Studies. Encina Hall, 616 Jane Stanford Way, Stanford, CA 94305-6055 (Estats Units d'Amèrica). jbaratamir@gmail.com, [@JoanBarata](https://twitter.com/JoanBarata).

Article rebut el 13.09.2020. Avaluació cega: 25.09.2020 i 04.10.2020. Data d'acceptació de la versió final: 07.10.2020.

Citació recomanada: Barata Mir, Joan. (2020). Llibertat d'expressió i plataformes digitals: reptes de la regulació de continguts en un entorn global. *Revista Catalana de Dret Públic*, 61, 1-13. <https://doi.org/10.2436/rcdp.i61.2020.3519>.

Sumari

1 Introducció

2 La llibertat d'expressió com a dret que es protegeix tant en línia com fora de línia

3 El paper dels intermediaris i la moderació de continguts

4 Els intermediaris i els sistemes d'exempció de responsabilitat

5 La llibertat d'expressió dels usuaris de les plataformes

6 Conclusions

Referències bibliogràfiques

1 Introducció

La protecció de la llibertat d'expressió en els moments actuals suposa prendre consciència dels canvis tecnològics que afecten l'activitat consistent a cercar, impartir i rebre informació, idees i opinions de tot tipus. Més important que això, tanmateix, és entendre correctament quins són els diversos actors i elements normatius que en aquests moments es troben implicats en la definició del règim aplicable a la difusió de continguts en línia, especialment a través de les anomenades “plataformes o serveis d'intermediació”.

Aquest règim es basa en la concurrència dels principis proporcionats, d'una banda, per les normes i les institucions internacionals en matèria de drets humans, així com les normes jurídiques adoptades per les autoritats nacionals i, de l'altra, per les regles i estàndards establerts per les mateixes plataformes amb la intenció i finalitat d'establir un marc de civilitat pel que fa a les condicions de diàleg i convivència entre els usuaris.

Aquest article intenta analitzar les principals qüestions i problemes jurídics vinculats a les matèries esmentades.

2 La llibertat d'expressió com a dret que es protegeix tant en línia com fora de línia

Els drets fonamentals a la llibertat d'expressió i a la llibertat d'informació es troben reconeguts, protegits i delimitats en la pràctica totalitat de les constitucions dels sistemes polítics de les democràcies liberals, entre elles la Constitució espanyola (article 20). També formen part nominalment de les constitucions d'altres tipus de règims, de caràcter autoritari o semiautoritari, si bé en aquests casos l'abast d'aquest dret es troba sovint força limitat sobre la base de límits molt amplis o remissions completament obertes a la voluntat del legislador.

Els drets de referència formen igualment part del quadre fonamental de drets humans reconegut per les normes de dret internacional. Concretament, cal esmentar la Declaració Universal dels Drets Humans (article 19) i, de manera més desenvolupada, el mateix article 19 del Pacte Internacional de Drets Civils i Polítics, ambdós instruments adoptats en el marc de Nacions Unides. Val la pena recordar el contingut d'aquest darrer (en la versió oficial en espanyol):

1. Nadie podrá ser molestado a causa de sus opiniones.
2. Toda persona tiene derecho a la libertad de expresión; este derecho comprende la libertad de buscar, recibir y difundir informaciones e ideas de toda índole, sin consideración de fronteras, ya sea oralmente, por escrito o en forma impresa o artística, o por cualquier otro procedimiento de su elección.
3. El ejercicio del derecho previsto en el párrafo 2 de este artículo entraña deberes y responsabilidades especiales. Por consiguiente, puede estar sujeto a ciertas restricciones, que deberán, sin embargo, estar expresamente fijadas por la ley y ser necesarias para:
 - a) Asegurar el respeto a los derechos o a la reputación de los demás;
 - b) La protección de la seguridad nacional, el orden público o la salud o la moral públicas.

A banda de l'article 19, el Pacte també inclou un article que fou objecte de gran controvèrsia en el procés d'elaboració d'aquest instrument legal, atès que una part dels estats de tall majoritàriament liberal entenien que l'article esmentat en primer lloc contenia ja totes les previsions necessàries per a manejar qualsevol possible abús o excés en l'exercici de la llibertat d'expressió. Finalment, però, les negociacions es varen decantar a favor dels qui proposaven la prohibició, en un article específic, de certes formes d'expressió (Kearney, 2007: 128, 131). Val a dir que la incorporació d'aquest article al dret internacional dels drets humans ha suscitat un gran nombre de debats interpretatius, especialment a l'entorn de l'abast de la noció de “propaganda per a la guerra” (paràgraf 1) o els termes en què el paràgraf 2 estableix els elements bàsics de la figura del que es coneix com a “discurs de l'odi”:¹

¹ Vegeu el *non-paper* de la representant en matèria de llibertat de mitjans de l'Organització per a la Seguretat i Cooperació a Europa (OSCE) intitulat “[Propaganda and Freedom of the Media](#)”.

1. Toda propaganda en favor de la guerra estará prohibida por la ley.
2. Toda apología del odio nacional, racial o religioso que constituya incitación a la discriminación, la hostilidad o la violencia estará prohibida por la ley.

És important destacar una sèrie d'elements fonamentals que es deriven d'aquest marc jurídic internacional.

El primer radica en el fet que, deixant de banda la dinàmica específica de l'article 20, l'article 19 preveu la possibilitat d'imposar certs límits al dret a la llibertat d'expressió. Ara bé, cal tenir en compte que allò que caracteritza l'esmentat precepte (el paràgraf 3 en particular) és el fet de delimitar amb precisió els requeriments sobre la base dels quals les autoritats nacionals podran, precisament, establir els esmentats límits o condicions. Dit d'altra manera, l'article 19 estableix, pel que fa a la llibertat d'expressió, els *límits dels límits*.

De manera particular, totes les restriccions, les regulacions i les limitacions que afectin el dret a la llibertat d'expressió i el dret a la llibertat d'informació han de ser establertes de manera clara i precisa per l'ordenament jurídic, estar orientades a la protecció dels drets i interessos esmentats al paràgraf 3 de l'article 19 i respectar els principis de necessitat i proporcionalitat. És a dir, les mesures en qüestió han de ser les estrictament necessàries i implicar la mínima restricció possible. Aquest "test tripartit"² és una obligació que es deriva del dret internacional i, per tant, és aplicable a totes les polítiques, legislacions i regulacions públiques adoptades en l'àmbit nacional. D'altra banda, aquest test és emprat també pels respectius tribunals regionals de drets humans, sobre la base de l'instrument corresponent (en el cas d'Europa, la jurisprudència del Tribunal Europeu de Drets Humans a partir de les previsions contingudes a l'article 10 del Conveni Europeu de Drets Humans) (Bychawska-Siniarska, 2017).

En segon lloc, la llibertat d'expressió es reconeix i protegeix "sense consideració de fronteres". Això té implicacions importants en l'era d'internet, la qual podria permetre que finalment aquesta dimensió potencialment global i il·limitada de la llibertat d'expressió es convertís en realitat. És evident, en aquest sentit, que internet permet situar la dinàmica de la distribució i disseminació d'idees, opinions i informacions en un àmbit completament diferent del que ha caracteritzat el règim jurídic (principalment nacional) aplicable als mitjans de comunicació.

En tercer lloc, l'article 19 protegeix l'exercici de la llibertat d'expressió per part de qualsevol persona i a través de qualsevol tipus de mitjà o tecnologia, i no només els que existeixen en el moment d'adopció del Pacte (l'any 1966). Com ha afirmat el Comitè de Drets Humans de Nacions Unides:³

En la función periodística participan una amplia variedad de personas, como analistas y reporteros profesionales y de dedicación exclusiva, autores de blogs y otros que publican por su propia cuenta en medios de prensa, en Internet o por otros medios.

En tot cas, és evident que ens trobem davant d'una concepció més àmplia dels conceptes de mitjans de comunicació i periodisme, la qual ha donat lloc a nous debats dins l'àmbit de la definició d'estàndards internacionals (Richter, 2016).

És necessari recordar, en aquest context, que el Consell de Drets Humans de Nacions Unides ha declarat en la Resolució 32/13, d'11 de juliol de 2016, que "[...] the same rights that people have offline must also be protected online, in particular freedom of expression, which is applicable regardless of frontiers and through any media of one's choice, in accordance with articles 19 of the UDHR and ICCPR", en línia igualment amb el que estableixen les resolucions 20/8, de 5 de juliol de 2012, i 26/13, de 26 de juny de 2014, sobre la promoció, la protecció i el gaudi dels drets humans a internet.

2 Comitè dels Drets Humans de Nacions Unides. Sessió número 102. Ginebra, 11-29 juliol 2011. [Observació general número 34](#).

3 Observació general número 34, ja esmentada.

3 El paper dels intermediaris i la moderació de continguts

El 2018, el Comitè de Ministres del Consell d'Europa (CoE) va adoptar una Recomanació “sobre el paper i les responsabilitats dels intermediaris d'Internet”,⁴ que descrivia aquests actors com un ampli ventall de jugadors que evoluciona ràpidament, i amb funcions importants:

facilitate interactions on the internet between natural and legal persons by offering and performing a variety of functions and services. Some connect users to the internet, enable the processing of information and data, or host web-based services, including for user-generated content. Others aggregate information and enable searches; they give access to, host and index content and services designed and/or operated by third parties. Some facilitate the sale of goods and services, including audio-visual services, and enable other commercial transactions, including payments.

Els intermediaris s'han convertit en els principals actors del procés de difusió i distribució de tot tipus de contingut. També tenen un paper molt destacat a l'hora d'assegurar la “visibilitat” del contingut produït pels mitjans tradicionals, com la premsa escrita o fins i tot la ràdio i la televisió. De fet, cada cop és més freqüent que aquest tipus de contingut es consumeixi com a conseqüència d'una recomanació o de la seva aparició a les xarxes socials i a tot tipus de plataformes, i no a causa de l'accés directe per part dels usuaris.⁵

Tot i això, també és evident que els interessos que modulen les decisions preses pels intermediaris (en particular els que estan relacionats amb el contingut) no són únicament i necessàriament la protecció del dret a la llibertat d'expressió. Els intermediaris són, sens dubte, entitats orientades comercialment i segueixen models de negoci basats en la venda de publicitat (que pot donar lloc, en alguns casos, a la promoció d'un discurs socialment indesitjable o perjudicial). En aquest sentit, Bietti (2020) adverteix que els intermediaris promouen contingut per tal de maximitzar la implicació, l'addicció, l'orientació conductual i la polarització dels usuaris. Les plataformes han intentat contrarestar aquest tipus d'arguments mostrant el seu compromís a proporcionar als usuaris contingut que els sigui “significatiu”, més enllà d'un sistema de classificació basat en allò que els agrada i en què fan clic.⁶

També són rellevants altres pressions (no sempre de naturalesa econòmica): cal mostrar responsabilitat a l'hora de tractar “continguts problemàtics” dins del marc de diferents societats, gestionar la pressió provinent de comissions parlamentàries o mitjans de comunicació i l'opinió pública, gestionar les “amenaces” de regulació dels governs i els legisladors, i ser capaç de respondre adequadament al comportament dels rivals empresarials (Keller, 2018b). En paral·lel, els òrgans competents del sistema internacional i la societat civil demanen a les plataformes que respectin el dret internacional dels drets humans a l'hora no només de moderar continguts, sinó també de prendre en consideració la legislació o els requeriments provinents de les respectives autoritats nacionals. Per exemple, en l'informe temàtic de 2018 davant del Consell de Drets Humans, el relator especial de les Nacions Unides sobre la promoció i protecció del dret a la llibertat d'opinió i expressió, David Kaye, va adreçar-se directament a les plataformes i els va demanar que reconeguessin que “the authoritative global standard for ensuring freedom of expression on their platforms is human rights law, not the varying laws of States or their own private interests, and they should re-evaluate their content standards accordingly”.⁷

Cal destacar que, en qualsevol cas, sens perjudici de la intensa activitat de bona part de les plataformes per tal d'evitar la presència de determinats continguts indesitjables (ja sigui per motius de violació de la legalitat o de les mateixes normes internes de comportament), aquestes plataformes cometen de manera regular errors importants a l'hora de classificar continguts suposadament nocius i prendre decisions al respecte (York i Gullo, 2018; Perrigo 2020a, 2020b). En aquest àmbit, l'ús de mecanismes d'intel·ligència artificial per tal de moderar els continguts esmentats abans sorgeix com una necessitat pràcticament ineludible en un context en què les plataformes es veuen obligades a aplicar normes jurídiques i estàndards interns *a escala*. És a

4 Recomanació CM/REC(2018)2, de 7 de març de 2018, del Comitè de Ministres als estats membres sobre el paper i les responsabilitats dels intermediaris d'Internet.

5 Vegeu per exemple l'informe del Pew Research Center de 2018, “[News Use Across Social Media Platforms](#)”.

6 Vegeu l'entrevista a Mark Zuckerberg per part d'Ezra Klein “[Mark Zuckerberg on Facebook's hardest year, and what comes next](#)”. *Vox*, 2 d'abril de 2018.

7 Disponible en línia a: www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ContentRegulation.aspx.

dir, sens perjudici que en alguns casos, especialment després d'obtenir inicialment un *positiu*, existeixi un cert nombre de continguts que són objecte d'una revisió més detallada per part de moderadors humans, les plataformes es veuen en la necessitat de detectar continguts inadequats de manera molt ràpida i en grans quantitats. Per tant, és materialment impossible d'aplicar, a tots els casos, sistemes d'anàlisi i avaluació equivalents als que empraria una autoritat administrativa o judicial.

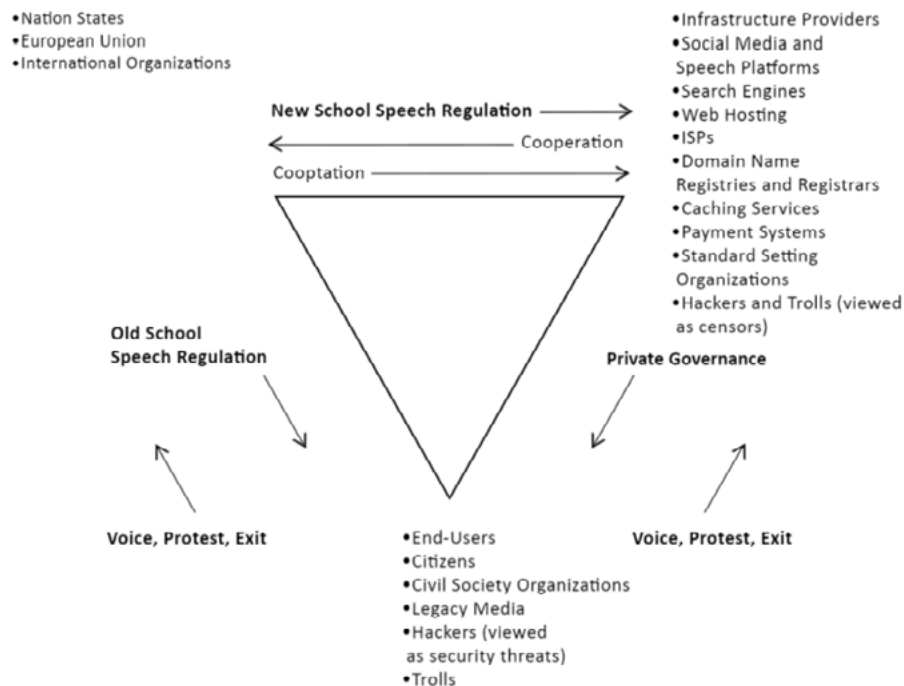
Un exemple especialment gràfic de l'anterior té lloc en relació amb els continguts del que podem anomenar extremisme violent o en casos de terrorisme. A la vista de situacions com ara l'aparició de vídeos de propaganda de l'organització terrorista ISIS en algunes plataformes o fins i tot la retransmissió en directe, pels mateixos terroristes, d'algun dels seus atemptats, les principals plataformes varen desenvolupar tècniques de concordança de contingut a partir d'una base de dades d'empremtes digitals o *hash*. Cada cop que alguna plataforma identifica i elimina un contingut per considerar-lo terrorista, hi afegeix l'esmentada empremta digital i el comparteix en aquesta base de dades comuna, a disposició de la resta dels socis de la iniciativa. En particular, el Fòrum Global d'Internet per Combatre el Terrorisme (GIFCT, en les seves sigles en anglès)⁸ va ser creat el 2017 per Facebook, Twitter, Microsoft i YouTube com una iniciativa del sector “per evitar que terroristes i extremistes violents explotin plataformes digitals”. Actualment, el GIFCT està en camí de convertir-se en un organisme gestionat independentment i no només per les mateixes plataformes. No obstant això, preocupacions relacionades amb la manca de transparència, la vulnerabilitat davant la pressió dels governs i l'impacte en les accions dels grups de la societat civil (degut a les conseqüències dels possibles errors o *falsos positius*) fan que sigui un instrument força controvertit. En tot cas, l'impacte d'aquesta base de dades en l'exercici de la llibertat d'expressió és de gran importància i, tal com han assenyalat alguns autors, es tracta de l'acrònim més important del qual mai no heu sentit parlar, el qual pot, no obstant això, canviar la manera de governar internet de manera radical (Radsch, 2020).

En aquest context general, Balkin (2018) ha creat la coneguda figura del “triangle de la lliure expressió” (figura 1), en què en un cantó hi ha estats i altres organismes públics i organitzacions supranacionals, en el segon cantó es poden trobar empreses d'internet (xarxes socials, motors de cerca, proveïdors d'internet, etc.), mentre que en el tercer costat trobem els usuaris, la societat civil i els mitjans de comunicació tradicionals. A través d'aquest triangle s'evidencia com la relació entre els poders públics i els usuaris i els mitjans tradicionals és la típica relació de poder en què es combinen els instruments legals que ja coneixem, així com la protesta o l'advocació ciutadana o empresarial. D'altra banda, la relació entre les empreses d'internet i els organismes estatals seria més complexa degut tant a la novetat de la matèria com als problemes per a garantir l'aplicació de les normes corresponents. Per aquest motiu, l'autor parla de regulació *de la nova escola*, també anomenada en alguns entorns *smart regulation*,⁹ basada en mecanismes de cooperació o corregulació, i en què els riscos de possible cooptació de les autoritats públiques serien més grans que dins l'àmbit de la regulació tradicional. Finalment, i com s'ha apuntat ja, la relació entre les empreses d'internet i els diversos tipus d'usuari encaixaria més pròpiament en la idea de governança o normació privada.

⁸ <https://www.gifct.org>.

⁹ Vegeu la proposta formulada per un dels vicepresidents de Google el febrer de 2019 “[Smart Regulation for combating illegal content](#)”.

Figure 1: The Pluralist Model of Speech Regulation¹⁰



Font: Balkin (2018).

Un exemple de la complexa interacció entre les plataformes i determinades instàncies amb poder polític i regulatori en relació amb els continguts en línia el trobem a la Unió Europea, en el marc de les iniciatives de foment de l'autoregulació. En concret, les empreses informàtiques Facebook, Microsoft, Twitter i YouTube (i aquí cal afegir-hi TikTok des del setembre de 2020) van signar un codi de conducta amb la Comissió Europea el maig del 2016 amb l'objectiu de contrarestar el contingut il·legal en línia.¹⁰ Aquesta iniciativa es troba connectada amb les comunicacions i les recomanacions prèviament formulades per part de la Comissió Europea en aquesta matèria.¹¹ Iniciatives semblants han tingut lloc igualment en relació amb el combat de la desinformació.¹²

En relació amb la legislació nacional, és obvi que les plataformes globals es veuen també obligades a tenir en compte i aplicar les lleis nacionals i les decisions d'autoritats nacionals (com ara autoritats de regulació o jutges), ja que els continguts publicats i consumits per part del usuari es poden associar normalment a un determinat territori. Per esmentar un exemple, l'1 de gener de 2018 va entrar en vigor a Alemanya la *Netzwerkdurchsetzungsgesetz* (Llei d'implementació del dret a la xarxa o NetzDG).¹³ Aquesta norma va ser objecte d'un gran nombre de debats i crítiques en el seu moment, particularment per part d'organitzacions que defensen els drets fonamentals, atès que la consideraven desproporcionada, inadequada i promotora de la censura privada a internet.¹⁴ De manera resumida, la NetzDG obliga a les plataformes (aquelles amb més de dos milions d'usuaris a Alemanya) a proporcionar mecanismes als usuaris per tal que aquests puguin informar d'un ampli ventall (22 específicament) de categories ja incloses a la legislació penal alemanya. Un cop rebuda la queixa corresponent, correspon a la plataforma determinar si el contingut és realment il·legal.

10 Disponible en línia a: https://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=54300.

11 Comunicació de la Comissió Europea al Parlament Europeu, el Consell, el Comitè Econòmic i Social i el Comitè de les Regions sobre la lluita contra el contingut il·legal en línia i cap a una major responsabilització de les plataformes en línia. COMM (2017) 555 final (28 de setembre de 2017) i Recomanació de la Comissió d'1 de març de 2018 sobre les mesures per lluitar de forma efectiva contra el contingut il·legal en línia. C (2018) 1177 final.

12 Vegeu [les diverses iniciatives en aquesta matèria](#) i en particular el codi de conducta signat amb les empreses tecnològiques.

13 Publicada a la *Gazeta Federal* l'1 de setembre de 2017, pàgines 3.352 i següents.

14 Vegeu l'anàlisi crítica elaborada per l'associació per a la defensa de la llibertat d'expressió en l'àmbit global, [article 19](#).

Segons la NetzDG, en casos de contingut “manifestament il·legal”, les plataformes l’han d’eliminar en un termini de 24 hores. En els casos en què aquesta determinació sigui més complexa, l’eliminació s’ha de dur a terme en un termini de 7 dies. L’incompliment d’aquestes obligacions està sancionat amb multes de fins a 50 milions d’euros. Així mateix, les plataformes estan obligades a publicar “informes de transparència” que reflecteixin les decisions adoptades sobre la moderació del contingut.

La NetzDG es pot considerar com una llei realment única tant pel que fa al tipus d’obligacions que imposa a aquest tipus de proveïdors com a la seva extensió i intensitat i la gravetat de les sancions que preveu en casos d’incompliment de les seves provisions. Tot i que l’eficàcia de la NetzDG com a eina per combatre i eliminar certes formes de contingut il·legal en línia encara està pendent avui en dia de dades fiables, el cert és que la fama de la NetzDG va arribar a escala mundial i que s’ha utilitzat explícitament com a inspiració de règims autoritaris o semiautoritaris d’altres parts del món per justificar l’adopció de normes considerablement restrictives (Mchangama i Alkiviadou, 2020). Aquests règims inclouen els de Rússia, Veneçuela, Bielorússia, Hondures, Malàisia, Singapur o Filipines.

Una norma de les característiques de la NetzDG comporta un risc evident d’eliminació excessiva del contingut. Si una espasa de Dàmocles penja sobre les plataformes en forma de grans multes per a aquells casos en què no hagin considerat correctament un contingut determinat com a il·legal, és evident que aquestes empreses se sentiran empeses a eliminar qualsevol contingut que susciti la més mínima sospita de no respectar l’ordenament. L’impacte en l’exercici de la llibertat d’expressió és evident, sobretot si tenim en compte el fet que, en puritat, no ens trobem davant d’una decisió estrictament privada, sinó una limitació derivada originalment d’una norma legal, el compliment de la qual és verificat en darrera instància per les autoritats competents. Aquests casos impliquen ni més ni menys que l’exercici del poder de determinar l’abast legítim de la llibertat d’expressió i, el que és més important, actuar en conseqüència, especialment pel que fa a la retirada del contingut corresponent, acaba essent delegat a favor de corporacions privades. Aquest poder, fins ara, només havia estat exercit per part dels jutges o les autoritats de regulació. Si bé és cert que totes aquestes decisions poden sotmetre’s al control jurisdiccional, l’impacte de l’eliminació immediata d’un contingut determinat difícilment es pot compensar amb una decisió judicial favorable a les reclamacions del seu autor.

A banda d’això, hi ha una altra tendència emergent que aniria més enllà del marc normatiu acabat d’esmentar, i que representa l’adopció de normes jurídiques destinades a obligar les plataformes no només a complir restriccions legals generals en àmbits com el discurs d’odi, la difusió de continguts terroristes o la protecció dels menors, sinó també a adoptar normes i principis addicionals que suposarien, de manera paradoxal, la imposició per llei de l’obligació d’adoptar normes de restricció de la llibertat d’expressió que aniríem més enllà, precisament, dels mateixos límits legals.

La iniciativa orientada a regular els anomenats “danys en línia” o *online harms* al Regne Unit representa un bon exemple d’aquesta tendència. El Govern britànic va publicar, l’abril de 2019, un extens llibre blanc sobre aquesta matèria, la qual va ser objecte d’una extensa consulta pública.¹⁵ El febrer de 2020 es va publicar una resposta inicial sobre la base de les aportacions rebudes, que assenyalava alguns detalls més específics en algunes àrees.¹⁶ La idea principal darrere aquesta iniciativa és la imposició i verificació per part d’un regulador especialitzat d’un anomenat “deure de cura” o *duty of care*, sobre la base del qual les plataformes haurien d’adoptar una sèrie de mesures, a través dels seus codis i estàndards interns, orientades a eliminar o mitigar diverses modalitats de continguts no només il·legals, sinó també allò que el regulador acabi classificant com a “perjudicial”, perillós per a “l’estil de vida del Regne Unit” o “inacceptable”, és a dir continguts que, “encara que no siguin il·legals en totes les circumstàncies, també poden causar greus danys”. De manera especial, aquesta iniciativa mostra preocupació i voluntat d’atacar el problema de la desinformació en línia, especialment quan aquesta prové d’“actors hostils” que utilitzen la desinformació en línia per socavar “valors democràtics”.

En definitiva, es tracta d’un model de regulació de continguts en línia que, un cop més, incorpora importants “poders delegats” a les plataformes en relació amb l’establiment de límits al dret fonamental a la llibertat

15 Disponible en línia a: <https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper>.

16 Disponible en línia a: <https://www.gov.uk/government/consultations/online-harms-white-paper/public-feedback/online-harms-white-paper-initial-consultation-response>.

d'expressió. Aquests poders s'han d'exercir, concretament, en relació amb determinades categories considerades per la llei com a “problemàtiques”, però no enunciades ni definides en detall per aquesta, atès que és una tasca evolutiva que haurà d'exercir una instància independent de regulació en funció de les circumstàncies i els riscos que existeixin en cada moment en el marc de la societat. I si bé és cert que, en aquests casos, la decisió darrera queda en mans de la plataforma de què es tracti, serà responsabilitat del regulador verificar si aquesta ha pres mesures “estructurals” o de procediment adequades per a detectar i tractar contingut problemàtic, així com que aquestes siguin aplicades de manera transparent i coherent.

El cert en tot cas és que en aquest model existeix un alt grau d'incertesa sobre el que és o no és acceptable en el marc d'una determinada plataforma. Aquesta qüestió i el concret mecanisme de moderació de continguts que cada plataforma acabi adoptant es deixen de manera força discrecional a mans d'una autoritat administrativa de regulació. Aquesta incertesa no crea un marc favorable a l'exercici de la llibertat d'expressió, sinó més aviat el contrari (Smith 2020a, 2020b).

4 Els intermediaris i els sistemes d'exempció de responsabilitat

Els sistemes legals amb més influència en el món (particularment els de la Unió Europea i els Estats Units) incentiven la moderació del contingut per part de les plataformes, garantint que no són penalitzades com a conseqüència del contingut subministrat per tercers o a causa de l'adopció de determinades mesures orientades a eliminar contingut il·legal o contingut inadequat.

Als Estats Units, la secció 230 de la Llei de comunicacions de 1934 (modificada per la Llei de telecomunicacions de 1996) eximeix de responsabilitat penal les plataformes en relació amb el contingut publicat per tercers. D'altra banda, proporciona als intermediaris immunitat pel que fa a qualsevol decisió de moderació de continguts, en la mesura en què: a) s'ha adoptat “de bona fe”, i b) fa referència a continguts o, de manera més àmplia, a informació que la plataforma considera com a “objectable”. La secció 230 ha jugat un paper fonamental en el desenvolupament d'internet tal com la coneixem. Sota les proteccions establertes per la legislació nord-americana, les plataformes han tingut l'incentiu d'operar i expandir els seus negocis sota un règim jurídic previsible i han comptat amb la capacitat per a intervenir-hi en relació amb la informació que distribueixen, especialment quan es tracta de continguts suposadament il·legals o susceptibles de ser considerats com a nocius d'acord amb els estàndards interns establerts per part de cada plataforma (Goldman, 2019).

Els articles 14 i 15 de l'anomenada Directiva sobre comerç electrònic¹⁷ contenen disposicions que afecten igualment la moderació del contingut per part dels intermediaris. Sobre la base d'aquestes disposicions, les plataformes no serien responsables en relació amb els continguts subministrats per tercers sobre la base d'una sèrie de requisits i condicions. Per tal de conservar la immunitat, les plataformes no han de tenir coneixement real d'activitat o informació il·legal i/o no han de ser conscients de fets o circumstàncies a partir dels quals es manifesta l'activitat o informació il·legal (segons reclamacions de responsabilitat civil). A més, la legislació de la Unió Europea (article 15 de la Directiva esmentada) també prohibeix la imposició d'obligacions generals de control (monitoratge) de continguts o de recerca activa de fets o circumstàncies que evidencin la presència d'activitats il·legals. En qualsevol cas, els intermediaris poden adoptar voluntàriament les seves pròpies (i ja esmentades) regles de moderació i supervisió del contingut i fer-les complir sense perdre necessàriament l'esmentada exempció (Barata Mir, 2020a).

Així, les disposicions legals vigents als Estats Units i a la Unió Europea (així com en altres parts del món, com l'Índia i diversos països d'Amèrica Llatina) protegeixen en diferent mesura la capacitat de les plataformes per decidir com organitzar, prioritzar, degradar o simplement eliminar el contingut. Aquesta protecció es proporciona mitjançant exempcions de responsabilitat legal.

Fruit d'aquest marc legal, les grans plataformes disposen, efectivament, d'un complet ventall de normes internes que van molt més enllà dels límits legals generalment establerts dins l'àmbit de la llibertat d'expressió.

¹⁷ Directiva 2000/31/EC del Parlament Europeu i el Consell, de 8 de juny de 2000, sobre determinats aspectes legals dels serveis de la societat de la informació, en particular el comerç electrònic, en el mercat interior.

Exemples d'això serien els estàndards de la comunitat de Facebook,¹⁸ les normes i polítiques de Twitter¹⁹ o la guia per a la comunitat de YouTube.²⁰

Això forma part d'un debat molt rellevant, sobretot des que la Comissió Europea s'ha compromès a presentar una proposta d'un paquet legislatiu de la Llei de serveis digital (DSA),²¹ que, entre altres coses, substituirà les disposicions de la Directiva sobre comerç electrònic. Aquest debat se centra, entre altres aspectes, en si l'increment del control "editorial" dels intermediaris sobre el contingut mitjançant pràctiques de moderació també podria justificar la introducció de modulacions legals que afectin el grau de responsabilitat respecte al contingut de tercers que faciliten. Dit d'una altra manera, si seria possible, en termes de responsabilitat, aprofundir en la distinció ja apuntada (si bé de manera una mica imprecisa) pel Tribunal de Justícia de la Unió Europea entre intermediaris que allotgen i/o faciliten continguts de manera completament passiva i aquells que intervenen d'alguna manera en la informació subministrada, en tot cas, per tercers.²²

Prèviament, la Unió Europea ha adoptat ja algunes directives que afecten directament el règim de responsabilitat d'intermediaris i han generat igualment importants debats en relació amb el seu impacte sobre la llibertat d'expressió.

La Directiva 2019/79 del Parlament Europeu i del Consell, de 17 d'abril de 2019, sobre els drets d'autor i drets afins en el mercat únic digital, i per la qual es modifiquen les directives 96/9/CE i 2001/29/CE, conté, particularment en l'article 17, una sèrie de previsions que obligarien les plataformes a jugar un paper actiu per tal de detectar continguts il·licits en el moment de la seva pujada per part dels usuaris, donant doncs lloc a un deure de monitoratge amb un possible impacte negatiu sobre la llibertat d'expressió (Quintais et al., 2019).

D'altra banda, la Directiva 2018/1808 del Parlament Europeu i del Consell, de 14 de novembre de 2018, per la qual es modifica la Directiva 2010/13/UE, sobre la coordinació de determinades disposicions legals, reglamentàries i administratives dels estats membres relatives a la prestació de serveis de comunicació audiovisual, inclou una sèrie de deures aplicables als continguts audiovisuals subministrats per les anomenades plataformes de compartició de vídeos de tercers (com ara YouTube). Aquests deures afecten àrees com ara la prevenció i la moderació de continguts que constitueixin discurs d'odi i pornografia infantil, afectin el desenvolupament físic i mental dels infants, incompleixin les obligacions en matèria de comunicacions comercials o es puguin considerar de naturalesa terrorista. Les autoritats nacionals (principalment organismes independents de regulació) tenen ara la responsabilitat de verificar que les plataformes han adoptat les "mesures adequades" per tractar adequadament els tipus de continguts acabats d'esmentar. Això inclou la garantia que les plataformes revisen i apliquen correctament les seves normes internes, disposen de mecanismes adequats de denúncia de continguts, implementen sistemes de verificació o classificació i control d'edat dels usuaris, disposen de procediments transparents, fàcils d'utilitzar i eficaços per a resoldre les queixes dels usuaris i proporcionen eines d'alfabetització mediàtica. La Directiva atorga igualment a les autoritats competents la possibilitat d'imposar a les plataformes el deure de restablir un contingut erròniament eliminat sobre la base d'una incorrecta aplicació dels criteris legals corresponents. Aquestes previsions plantegen en tot cas un nombre considerable de dubtes i incerteses des del punt de vista de la seva aplicació efectiva (Barata Mir, 2020b).

5 La llibertat d'expressió dels usuaris de les plataformes

Com s'ha assenyalat, tant el Consell d'Europa com altres instàncies internacionals posen un èmfasi especial en el fet que els intermediaris tenen un paper molt important en la creació de les condicions per a l'exercici del dret a la llibertat d'expressió en línia.

18 <https://www.facebook.com/communitystandards>.

19 <https://help.twitter.com/en/rules-and-policies>.

20 <https://www.youtube.com/howyoutubeworks/policies/community-guidelines>.

21 <https://ec.europa.eu/digital-single-market/en/digital-services-act-package>.

22 Vegeu, principalment, els casos L'Oreal (Sentència de 12 de juliol de 2011, cas C-324/09) i Google France (Sentència de 23 de març de 2010, casos C-236/08 a C-238/08).

Aquesta perspectiva basada en la idea que els intermediaris faciliten i permeten l'exercici del dret a la llibertat d'expressió per part dels seus usuaris no es pot trobar, almenys amb els mateixos significat i implicacions, en el sistema jurídic dels Estats Units. Concretament, existeix una jurisprudència consolidada que determina que els ciutadans o usuaris no poden exercir el seu dret a la llibertat d'expressió contingut a la Primera esmena davant decisions (d'eliminació d'una determinada publicació o fins i tot la desactivació d'un compte) preses per part d'un intermediari. Ans al contrari, formaria part del dret a la llibertat d'expressió, del qual serien titulars, precisament, les plataformes, la llibertat de decidir quins continguts s'allotgen i faciliten a tercers, i quins no (Keller, 2018a; Goldman, 2020a). Un cas recent i particularment exemplificatiu en aquest sentit és la sentència del cas *Prager University v. Google LLC*,²³ en què s'analitza la decisió de YouTube de restringir l'accés a una sèrie de vídeos publicats per l'entitat Prager University i desmonetitzar-los. Concretament, el tribunal destaca aquí el següent:

Despite YouTube's ubiquity and its role as a public-facing platform, it remains a private forum, not a public forum subject to judicial scrutiny under the First Amendment [...] PragerU runs headfirst into two insurmountable barriers—the First Amendment and Supreme Court precedent [...] .The Internet does not alter this state action requirement of the First Amendment.

Com veiem, el tribunal rebutja un possible ús del terme “fòrum públic” equivalent a la idea d'un espai públic en què cal donar cabuda i protegir l'expressió de qualsevol participant. Aquest concepte s'extreu d'una decisió del Tribunal Suprem que va confirmar la sentència del Tribunal Suprem de Califòrnia,²⁴ el qual havia articulat aquest concepte pel que fa a la possibilitat d'expressar-se lliurement en un espai d'accés públic com són uns grans magatzems. Cal tenir present, en tot cas, que aquesta idea s'extrauria no ja de la Primera esmena, sinó de l'articulat de la Constitució de l'estat de Califòrnia.

Dit d'una altra manera, i segons la doctrina establerta a *PragerU*, el fet que els intermediaris proporcionin un espai obertament accessible per parlar ni priva aquest espai de la seva naturalesa intrínsecament privada ni crea l'obligació de protegir la llibertat d'expressió dels usuaris.

El debat és més complex més enllà dels Estats Units, i principalment a Europa.

A Alemanya, el Tribunal Regional Superior de Munic va dictaminar que un comentari publicat per un polític conservador que infringia, segons Facebook, les seves normes de contingut intern constituïa un exercici de llibertat d'expressió protegit per la Constitució alemanya. El Tribunal va ordenar en conseqüència a Facebook tornar a publicar el comentari i restablir el compte de l'usuari, que havia estat eliminat durant 30 dies. No està clar si aquesta decisió serà confirmada per tribunals superiors, però és rellevant en el sentit que en primera instància s'ha establert que el dret a la llibertat d'expressió dels usuaris de Facebook s'ha de protegir en termes essencialment iguals que en relació amb les institucions de l'Estat, i que quan apliquen les seves pròpies normes de moderació de contingut intern els intermediaris han de tenir en compte, com a mínim, aquestes normes constitucionals.²⁵

Hi ha hagut casos similars en alguns països europeus i altres llocs del món, tot i que encara no es pot considerar un enfocament jurisprudencial consolidat confirmat pels tribunals superiors en general. L'exemple més recent i rellevant en redactar aquest text és la decisió d'un jutjat civil de Roma d'obligar Facebook a reactivar el compte del partit d'extrema dreta CasaPound i pagar-li 800 € per dia en concepte d'indemnització per danys durant el període en què el compte havia estat suspès. De manera similar al cas alemany, el compte suposadament havia infringit les polítiques internes en matèria de discurs de l'odi de Facebook. No obstant això, el Tribunal va considerar que la decisió de la plataforma va crear una inacceptable exclusió o limitació de la veu de CasaPound en el debat polític italià.²⁶

Així, doncs, a Europa podem trobar casos en què els òrgans jurisdiccionals han decidit, en relació amb determinades plataformes, l'adopció de mesures positives per protegir determinats continguts, principalment sobre la base del raonament que els intermediaris en qüestió juguen un paper fonamental en la facilitació de

23 *Prager University v. Google LLC*, cas núm. 18-15712 (9è Circuit, 26 de febrer de 2020).

24 *PruneYard v. Robins*, 447 U.S. 74 (1980).

25 Vegeu una descripció i anàlisi d'aquest cas a la [Columbia University Global Freedom of Expression Database](#).

26 Vegeu una descripció i anàlisi d'aquest cas a la [Columbia University Global Freedom of Expression Database](#).

l'exercici efectiu dels drets fonamentals (en particular la llibertat d'expressió o el dret a la no-discriminació), o incideixen en principis constitucionals bàsics nacionals i europeus (com ara el pluralisme) (Kettemann i Tiedeke, 2020). Aquesta és, per descomptat, una qüestió encara molt oberta que mereixerà un desenvolupament per part dels tribunals constitucionals nacionals i del Tribunal Europeu de Drets Humans.

En conseqüència, el debat a Europa sobre l'existència de possibles obligacions “positives” dels estats per protegir la llibertat d'expressió davant decisions restrictives de moderació de continguts per part dels intermediaris encara es troba en una etapa molt primerenca, tot i que alguns estudiosos ja han establert igualment connexions interessants entre el principi de pluralisme i diversitat propi del règim tradicional de mitjans com ara la ràdio i la televisió, i la introducció de possibles restriccions i obligacions a la manera com les plataformes gestionen el contingut. Aquestes restriccions i obligacions no implicarien principalment l'eliminació de contingut il·legal, sinó la promoció de la lliure difusió i l'accés a determinats tipus de continguts (Helberger et al., 2019).

Autors com ara Leerssen (2015) admeten, en tot cas, que, en termes generals, la identificació de l'acció directa de l'Estat actualment continua essent essencial per exercir o reivindicar el dret a la llibertat d'expressió davant prestadors de serveis en línia, particularment quan això es duu a terme davant de tribunals o organismes de regulació. No obstant això, també convida instàncies com el Tribunal de Justícia de la Unió Europea a explorar la dimensió horitzontal de l'exercici del dret a la llibertat d'expressió en línia i a apartar-se d'una estricta distinció entre responsabilitats públiques i privades en la consideració de possibles accions legals. No hi ha dubte que es tracta d'una qüestió que pot canviar alguns dels fonaments i enfocaments constitucionals tradicionals sobre la noció i els mètodes per protegir la llibertat d'expressió a les nostres societats, la qual cosa ens introdueix al que s'ha anomenat “constitucionalisme digital” (De Gregorio, 2020).

6 Conclusions

Com hem vist, són actualment moltes les tensions que existeixen entre els diferents grups o conjunts normatius acabats d'esmentar. Els estats sovint regulen els continguts en línia d'acord amb les seves prioritats i interessos, obviant les directrius i els límits establerts per les normes internacionals de protecció dels drets humans. Això també significa la creació de “compartiments” normatius que segmenten i alteren la lliure circulació d'informació i opinions en el marc de plataformes d'abast global, cosa que dona lloc al que comunament s'anomena la *balcanització* d'internet. D'altra banda, podem observar igualment com, a banda d'establir els esmentats marcs normatius, existeix una creixent tendència per part dels estats a influir en la manera com les mateixes plataformes d'intermediació estableixen els principis i estàndards que regiran el comportament dels usuaris i els límits a l'expressió dins del marc concret de la respectiva plataforma. Aquests límits es mouran generalment en àmbits més amplis que els que defineix la legalitat vigent.

Finalment, i lligat amb l'anterior, assistim en l'actualitat a un debat incipient en relació amb la possibilitat que les autoritats públiques (nacionals o internacionals) puguin determinar i aplicar mecanismes de tutela dels drets fonamentals dels usuaris de les plataformes, en relació amb decisions i mesures adoptades per aquestes.

Referències bibliogràfiques

- Balkin, Jack. (2018). Free Speech is a Triangle. *Yale Law School, Public Law Research Paper*, 640.
- Barata Mir, Joan. (2020a). [Positive Intent Protections: Incorporating a Good Samaritan principle in the EU Digital Services Act](#). *Center for Democracy and Technology*.
- Barata Mir, Joan. (25 de febrer 2020b). [Regulating content moderation in Europe beyond the AVMSD](#). [Apunt de blog]. *LSE Blog*.
- Bietti, Elettra. (1 de juny 2020). [Free Speech is Circular](#). [Apunt de blog]. *Medium*.
- Bychawska-Siniarska, Dominika. (2017). *Protecting the right to freedom of expression under the European Convention of Human Rights. A handbook for legal practitioners*. Estrasburg: Consell d'Europa.

- Goldman, Eric. (2019). Why Section 230 is better than the First Amendment, *Notre Dame Law Review Reflection*, 95(1), 33-46.
- Goldman, Eric. (2 de març 2020). [YouTube Isn't a State Actor \(DUH\) – PragerU v. Google](#). [Apunt de blog]. *Technology & Marketing Law Blog*.
- De Gregorio, Giovanni. (2020). [The Rise of Digital Constitutionalism in the European Union](#). *International Journal of Constitutional Law*. Pendent de publicació.
- Helberger, Natali, Leerssen, Paddy, i Van Drunen, Max. (29 de maig 2019). [Germany proposes Europe's first diversity rules for social media platforms](#). [Apunt de blog]. *LSE Blog*.
- Kearney, Michael G. (2007). *The Prohibition of Propaganda for War in International Law*. Oxford: Oxford University Press.
- Keller, Daphne. (2018a). [Who do you sue?](#) *Hoover Institution. Aegis Series Papers, 1902*.
- Keller, Daphne. (2018b). [Internet Platforms. Observations on Speech, Danger and Money](#). *Hoover Institution Aegis Series Paper, 1807*.
- Kettemann, Matthias C., i Tiedeke, Anne Sophie. (2019). [Back up: Can Users Sue Platforms to Reinstate Deleted Content? A Comparative Study of US and German Jurisprudence on "Must Carry"](#). *Internet Policy Review*, 9(2).
- Leerssen, Patrick. (2015). Cut Out by The Middle Man. The Free Speech Implications of Social Network Blocking and Banning in the EU. *JIPITEC 1*, 99-119.
- Mchangama, Jacob, i Alkiviadou, Natalie. (2020). [The Digital Berlin Wall: How Germany \(accidentally\) created a prototype for global online censorship – Part Two](#). Copenhagen: Justitia.
- Perrigo, Billy. (2020a). [These Tech Companies Managed to Eradicate ISIS Content. But They're Also Erasing Crucial Evidence of War Crimes](#). *Time*.
- Perrigo, Billy. (4 de maig 2020b). [When Content Moderation Hurts](#). [Apunt de blog]. *Mozilla Foundation*.
- Quintais, João Pedro, Frosio, Giancarlo, van Gompel, Stef, Hugenholtz, P. Bernt, Husovec, Martin, Jütte, Bernd Justin, i Senftleben, Martin. (2019). [Safeguarding User Freedoms in Implementing Article 17 of the Copyright in the Digital Single Market Directive: Recommendations from European Academics 10 \(2020\)](#). *JIPITEC*, 277.
- Radsch, Courtney. (2020). [GIFCT: Possibly the Most Important Acronym You've Never Heard Of. Hint: It Involves the Internet and Terrorism](#). *Just Security*.
- Richter, Andrei. (2016). Defining media freedom in international policy debates. *Global Media and Communication*, 12(2), 127-142.
- Smith, Graham. (24 de juny 2020a). [Online Harms Revisited](#). [Apunt de blog]. *Cyberleague*.
- Smith, Graham. (16 de febrer 2020b). [Online Harms Deconstructed – the Initial Consultation Response](#). [Apunt de blog]. *Cyberleague*.
- York, Jillian C., i Gullo, Karen. (6 de març 2018). [Offline/Online Project Highlights How the Oppression Marginalized Communities Face in the Real World Follows Them Online](#). [Apunt de blog]. *Electronic Frontier Foundation*.