



# APM 3.9.

Publicación digital. - Asociación Profesional de la Magistratura

ALFONSO PERALTA GUTIÉRREZ

ANDALUCÍA ORIENTAL

## TECNOCENSURA DE LOS ALGORITMOS A LA LIBERTAD DE EXPRESIÓN

### Resumen/Abstract:

Las redes sociales puede decirse que no son realmente plataformas de libertad de expresión sino plataformas de publicidad y tienen políticas de moderación de contenidos y operan bajo reglas privadas de gobernanza. Esto significa que se está produciendo un desplazamiento de los derechos humanos y fundamentales de la libertad de expresión o de información desde el derecho público estatal al derecho privado de las reglas de moderación de contenidos. Así, se convierten en las leyes de las redes sociales, emergiendo las plataformas tecnológicas como "gobernantes de la libertad de expresión", "guardianes de la esfera pública", "tribunal digitales del discurso" o "administradores de la cultura pública". El problema es que estos derechos humanos están girando desde la ley pública hacia una censura tecnológica, la "tecnocensura".

Las políticas de las redes sociales deciden qué está permitido, y los algoritmos de recomendación promocionan lo que es visible. Esto nos lleva a la siguiente pregunta, ¿quién controla a los moderadores?

### Palabras clave/ Keywords:

plataformas digitales, redes sociales, censura, moderación de contenido, algoritmos y libertad de expresión.

Digital platforms, social networks, censorship, content moderation, algorithm, freedom of expression.

## **Tecnocensura de los algoritmos a la libertad de expresión**

Las redes sociales permiten una comunicación inmediata, mayor conexión entre los usuarios y gente de distintas partes del mundo, y pueden potencialmente ayudar a la prensa libre a ejercer su papel en la sociedad.

La inteligencia artificial puede ayudar al periodismo a examinar y entender vastas cantidades de datos no estructurados, traducir rápidamente textos relevantes y transcribir y analizar contenido de vídeo y audio.

El uso del contenido generado por la IA (principalmente contenido basado en datos o que sea mayoritariamente rutinario o de poco valor añadido) puede liberar a los periodistas tiempo para enfocarse en contenido más complejo, importante e impactante, el periodismo de investigación. También puede ayudar a la distribución de ciertas noticias a los destinatarios, especialmente aquellas que son muy sensibles a ser de "última hora", como por ejemplo las noticias financieras.

Además, la inteligencia artificial puede ayudar a los periodistas a identificar las fake news, la desinformación o los deep fake. Usando IA en búsquedas entrecruzadas e inteligentes de datos mediante por ejemplo relación de entidades nominales o líneas temporales se puede ayudar a hacer comprobación de hechos (fact-checking) de las declaraciones que han hecho los políticos y contribuir a la salud democrática, el control político y la rendición de cuentas.

Las plataformas digitales han superado las fronteras de los países, la distancia y las distintas lenguas. En sólo un minuto, ¿cuánta actividad digital se genera? Las estimaciones publicadas por el sitio web AllAccess<sup>1</sup> en 2021, basadas en la investigación de Lori Lewis, se calcula que en un solo minuto se compartirán alrededor de 695.000 stories en Instagram, los usuarios de WhatsApp y Facebook Messenger enviarán 69 millones de mensajes y se subirán 500 horas de contenido a YouTube. Aunque vertiginosas, estas cifras son sólo una pequeña muestra del ritmo frenético con

---

<sup>1</sup> LEWIS Lori. Infographic: What Happens In An Internet Minute 2021 13 de abril 2021. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.allaccess.com/merge/archive/32972/infographic-what-happens-in-an-internet-minute>

el que se generan e intercambian datos y actividades en la Red. De ese inmenso big data con el que alimentamos fruto de nuestros datos a las grandes plataformas digitales.

Aunque vertiginosas, estas cifras son sólo una pequeña muestra del ritmo frenético con el que se generan e intercambian datos y actividades en la red. De ese inmenso big data con el que alimentamos fruto de nuestros datos a las grandes plataformas digitales.

Actualmente, de las 10 mayores empresas por capitalización bursátil del mundo, 9 son empresas tecnológicas y además norteamericanas<sup>2</sup>. Y estas compañías tienen un valor de mercado superior al de muchos países. Así, Apple vale como Brasil, Amazon como Corea del Sur, o Microsoft como el PIB de España<sup>3</sup>. Esto es lo que ha venido llamándose un "capitalismo digital", "tecnocapitalismo" o "capitalismo de datos o de vigilancia".

A diferencia del periodismo, en las redes sociales no se verifican las fuentes de la información, no se contrastan los hechos antes de su publicación y no jerarquizan las opiniones tras un análisis sosegado, consultan a expertos u opiniones cualificados. Los periodistas tratan de hablar con un cierto conocimiento de los hechos o una previa investigación periodística. Las plataformas digitales son como una gran taberna con opiniones espontáneas, igualan a todo el mundo, situando en la misma posición quien tiene un conocimiento básico, opiniones ofensivas o totalmente falsas, que un premio Nobel o un experto que las afirmaciones de tu cuñado. Da igual tu experiencia profesional, tus estudios o tu carrera, todo gira en torno al número de cuántos likes retuits o seguidores algo o alguien consigue. Cuánta viralidad consigue.

Las redes sociales están borrando la frontera entre la verdad y la mentira, todo es relativo y opinable, NOAH<sup>4</sup> asegura que "hoy en día la verdad la determina la primera página de resultados del buscador Google". En palabras de

---

<sup>2</sup> DRINCOVICH Sol. Empresas que valen más que países: casos sorprendentes sobre quién es el que manda en la economía digital. IproUp. 02 de octubre 2020 [Consultado el 15 de mayo de 2022] Disponible en: <https://www.iproup.com/innovacion/16761-economia-digital-que-empresas-valen-mas-que-paises>

<sup>3</sup> WALLACH, Omri. The World's Tech Giants, Compared to the Size of Economies. Publicado el 7 de julio de 2021. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.visualcapitalist.com/the-tech-giants-worth-compared-economies-countries/>

<sup>4</sup> NOAH HARARI, Yuval .21 lecciones para el siglo XXI. Debate. Barcelona. 2018.

PÉREZ REVERTE<sup>5</sup>, "eso de que todas las ideas son respetables es una imbecilidad. Lo respetable es el derecho a expresarlas".

Asimismo, las plataformas digitales tienen sus propias reglas o políticas de moderación de contenidos. Son empresas privadas y tienen derecho de admisión, de tal manera que son capaces de decidir lo que es permisible, puesto que de lo contrario eliminan el contenido e incluso te suspenden o eliminan la cuenta (baneo), y lo que es visible, de tal manera que los algoritmos determinan qué contenido ves, te hace sugerencias para ver más por los sistemas de recomendación y personalización de contenido, microsegmentándolo con la finalidad de la mayor optimización de la publicidad. El conocido como microtargetting.

Esto supone por tanto, que las grandes redes sociales están ejerciendo la denominada "tecnocensura" (Alberto Escorcía<sup>6</sup>).

Ello significa que se está produciendo un balance de los derechos humanos y fundamentales de la libertad de expresión y de información desde una ley estatal y pública hacia leyes o normas privadas, reglas de moderación de contenidos, los términos del servicio que se ven convertidos en las "leyes de las redes sociales". Emergiendo por tanto las grandes plataformas online como "gobernadores del discurso online", "custodios de la esfera pública" o<sup>7</sup> "administradores de la cultura pública".

Si el código es la ley según LESSIG<sup>8</sup>, los nuevos legisladores son las plataformas digitales y en concreto sus programadores, elegidos por los usuarios con nuestro uso de esas aplicaciones, pero ¿responden ante alguien? ¿cómo legislan?. Lo que nos lleva a la pregunta, ¿quién modera a los moderadores? Pregunta, que no puede responderse de manera sencilla ni corta, pero que nos lleva

---

<sup>5</sup> PÉREZ-REVERTE Arturo. @perezreverte Eso de que todas las ideas son respetables es una imbecilidad. Lo respetable es el derecho a expresarlas. 8:52 p. m. · 14 ago. 2011 Twitter. [Consultado el 15 de mayo de 2022] Disponible en: <https://twitter.com/perezreverte/status/10281477453191168?lang=es>

<sup>6</sup> La tecnocensura, explicada por el activista mexicano Alberto Escorcía. 15 de septiembre de 2016. Youtube. [Consultado el 15 de mayo de 2022] Disponible en: <https://youtu.be/2J6Hz-Uz1fY>

<sup>7</sup> SANDER, Barrie, Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation (May 23, 2020). Fordham International Law Journal, Vol. 43, No. 4, 2020, [Consultado el 15 de mayo de 2022] Disponible en: <https://ssrn.com/abstract=3434972> <http://dx.doi.org/10.2139/ssrn.3434972>

<sup>8</sup> LESSIG, Lawrence. El código 2.0. Madrid. Proyecto editorial Traficantes de Sueños. 2009.

hacia la Estrategia Europea Digital, al concepto de la denominada "soberanía digital", al Reglamento Europeo de "Ley de Servicios Digitales (DSA Digital Services Act)", y al Proyecto de Reglamento Europeo de Libertad de los Medios (EMFA European Media Freedom Act)<sup>10</sup>.

Está claro, que resulta absolutamente imposible encomendar a los jueces la moderación de los contenidos en línea, y el registro en una red social implica la obligación de someterse a sus términos de uso, sin que en caso de incumplimiento -a día de hoy- pueda reclamarse nada, en cuanto que dicho acuerdo firmado entre las partes constituye un contrato que vincula a ambas y establece una serie de derechos y obligaciones para cada una de ellas, recogiendo también las consecuencias de esos incumplimientos, entre las que se encuentra la suspensión de la cuenta sin que ello suponga la privación de la libertad de expresión, en cuanto que se trata de una red privada. Ciertamente es, que los sistemas de moderación de contenido son necesarios pues el gigantesco volumen de datos impide que se pueda realizar únicamente por humanos, de manera manual o artesanal.

En este sentido se ha pronunciado alguna jurisprudencia pionera<sup>11</sup> y recientísima<sup>12</sup>.

Sin embargo, esto supone que los derechos fundamentales como libertad de expresión están balanceándose desde una perspectiva pública a un área privada, y desde la ley a una censura privada, la "tecnocensura" (ESCORCIA, Alberto<sup>13</sup>).

Las políticas de moderación de contenidos de las plataformas digitales deciden lo que es permisible, y sus

---

<sup>9</sup> Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Reglamento de Servicios Digitales). <https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-81573>

<sup>10</sup> Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establece un marco común para los servicios de medios de comunicación en el mercado interior (Ley Europea de Libertad de los Medios de Comunicación) y se modifica la Directiva 2010/13/UE Bruselas, 16.9.2022 COM(2022) 457 final 2022/0277 (COD)

<sup>11</sup> Sentencia SAP Palma de Mallorca 124/2020, Sección Tercera, de 26 de marzo (ponente Ana Calado) (ROJ: SAP IB 513/2020 - ECLI:ES:APIB:2020:513 ) Consultado 15/5/2022)

<sup>12</sup> La recientísima SAP Valencia sección 8 del 11 de mayo de 2022 (ROJ: SAP V 2910/2022 - ECLI:ES:APV:2022:2910; Sentencia: 213/2022; Recurso: 852/2021; Ponente: ANA VEGA PONS-FUSTER OLIVERA).

<sup>13</sup> La tecnocensura, explicada por el activista mexicano Alberto Escorcía. 15 de septiembre de 2016. Youtube. [Consultado el 15 de mayo de 2022] Disponible en: <https://youtu.be/2J6Hz-Uz1fY>

algoritmos de recomendación lo que es visible<sup>14</sup>, puesto que de lo contrario eliminan el contenido e incluso te suspenden o eliminan la cuenta (baneo), y lo que es visible, de tal manera que los algoritmos determinan qué contenido ves, te hace sugerencias para ver más por los sistemas de recomendación y personalización de contenido, microsegmentándolo con la finalidad de la mayor optimización de la publicidad. El conocido como microtargetting.

BRKAN categoriza las potenciales interferencias con la libertad de expresión y de información en cuatro grupos: personalización, bloqueo automático o eliminación de contenido ilegal, degradación o dilución del contenido dañino (desinformación) e implementación automatizada del derecho al olvido<sup>15</sup>.

En este sentido, las plataformas digitales establecen restricciones en un amplio espectro de categorías de contenido -comúnmente discurso del odio, contenido gráfico o verbal violenta, contenido sexual, acoso, spam, fake news, copyright, y actividades ilegales. Esto resulta ser la interpretación y ejecución de lo que puede llamarse como "las leyes de las redes sociales". Se denominan técnicas de moderación de contenido online. Están basadas en una combinación de denuncias automáticas y comunitarias para detectar potencialmente contenido no permitido, y posteriormente será revisado ex post mediante una apelación automática y/o mediante moderadores humanos<sup>16</sup>.

Sin embargo, investigadores de la Universidad de Auburn (Alabama) y Adobe Research<sup>17</sup> han descubierto que existe un problema común en los sistemas de procesamiento del lenguaje natural pues son incapaces de identificar o

---

<sup>14</sup> SANDER, Barrie, Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation (May 23, 2020). Fordham International Law Journal, Vol. 43, No. 4, 2020, Disponible en SSRN: <https://ssrn.com/abstract=3434972> or <http://dx.doi.org/10.2139/ssrn.3434972>

<sup>15</sup> BRKAN, Maja, Freedom of Expression and Artificial Intelligence: On Personalisation, Disinformation and (Lack Of) Horizontal Effect of the Charter (March 17, 2019). Disponible en SSRN: <https://ssrn.com/abstract=3354180> or <http://dx.doi.org/10.2139/ssrn.3354180>

<sup>16</sup> Sander, Barrie, Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation (May 23, 2020). Fordham International Law Journal, Vol. 43, No. 4, 2020, Disponible en SSRN: <https://ssrn.com/abstract=3434972> or <http://dx.doi.org/10.2139/ssrn.3434972>

<sup>17</sup> PHAM, THANG & BUI, TRUNG & MAI, LONG & NGUYEN, ANH. (2021). Out of Order: How important is the sequential order of words in a sentence in Natural Language Understanding tasks?. 1145-1160. [Consultado el 15 de mayo de 2022] Disponible en: 10.18653/v1/2021.findings-acl.98. <https://arxiv.org/pdf/2012.15180.pdf>

entender las palabras cuando están desordenadas, el contexto. Los sistemas de moderación automáticos tampoco pueden leer entre líneas, comprender los giros idiomáticos o locales, la ironía, la parodia, los matices, o captar aspectos culturales. En el mismo sentido, investigadores de la Universidad del Sur de California, del Instituto Allen de IA y de la Universidad de Washington<sup>18</sup> han descubierto que los PNL (sistemas de procesamiento del lenguaje natural) pueden crear frases pero no pueden entenderlas porque no saben marcar las palabras clave, de tal manera que pueden crear frases ininteligibles. No tienen sentido común. Así, por ejemplo, si juntamos las palabras "pelota", "perro" y "lanzar", el sistema de inteligencia artificial puede crear como frase "el perro lanzó una pelota", en vez de una sencilla "yo lancé una pelota al perro".

Vemos que por ejemplo se ha convertido en noticia la censura de fotografías de esculturas desnudas<sup>19</sup>, cuadros de Rubens<sup>20</sup> o Delacroix<sup>21</sup>, o fotos con valor histórico-cultural<sup>22</sup>.

En España hace unos años ocurrió una curiosa "purga digital" basada en la frase "te mato". En 2018<sup>23</sup>, docenas de usuarios fueron baneados porque en algunos de sus tuits expresaban la frase "te mato", pero no como amenaza ni de una forma violenta, sino como broma, ironía o sarcasmo entre amigos o conocidos. Pero el algoritmo no puede entenderlo de una manera literal.

---

<sup>18</sup> LIN, BILL & ZHOU, WANGCHUNSHU & SHEN, MING & ZHOU, PEI & BHAGAVATULA, CHANDRA & YEJIN, CHOI & REN, XIANG. (2020). CommonGen: A Constrained Text Generation Challenge for Generative Commonsense Reasoning. 1823-1840. [Consultado el 15 de mayo de 2022] Disponible en: 10.18653/v1/2020.findings-emnlp.165. <https://arxiv.org/pdf/1911.03705.pdf>

<sup>19</sup> Facebook censura por «pornográfica» una escultura de hace 30.000 años. Cultura. ABC. 28 de febrero 2018. [Consultado el 15 de mayo de 2022] Disponible en: [https://www.abc.es/cultura/arte/abci-facebook-censura-pornografica-escultura-hace-30000-anos-201802280948\\_noticia.html](https://www.abc.es/cultura/arte/abci-facebook-censura-pornografica-escultura-hace-30000-anos-201802280948_noticia.html)

<sup>20</sup> ARAÚJO, SANTI .Facebook censura las pinturas de Rubens y varios museos lanzan una campaña que trollea esta situación. 25 Julio 2018. GENBETA. [Consultado el 15 de mayo de 2022] Disponible en:<https://www.genbeta.com/redes-sociales-y-comunidades/facebook-censura-pinturas-rubens-varios-museos-lanzan-campana-que-trollea-esta-situacion>

<sup>21</sup> Facebook censura 'La Libertad guiando al pueblo' por mostrar un pecho, EL PAÍS/AFP, 19 marzo 2018. [Consultado el 15 de mayo de 2022] Disponible en: [https://elpais.com/cultura/2018/03/19/actualidad/1521444972\\_981468.html](https://elpais.com/cultura/2018/03/19/actualidad/1521444972_981468.html)

<sup>22</sup> Facebook censuró una foto histórica y desata la polémica en Noruega. 9 de septiembre de 2016. LA NACION [Consultado el 15 de mayo de 2022] Disponible en: <https://www.lanacion.com.ar/tecnologia/facebook-censura-una-historica-fotografia-y-desata-la-polemica-en-noruega-nid1936172/>

<sup>23</sup> MIGUEL TRULA, ESTHER. Twitter tiene problemas con el lenguaje no literal y banea a quien dice "te mato" en sentido figurado. 25 Septiembre 2018. Xataka. [Accessed May 15th 2022]. Available at: <https://magnet.xataka.com/preguntas-no-tan-frecuentes/no-digas-te-mato-broma-podrias-ser-proxima-victima-censura-twitter>

Además, está convirtiéndose en una preocupación importante el surgimiento de sesgos en los algoritmos, especialmente aquellos que afectan a minorías étnicas (principalmente afroamericanos y latinos) así como mujeres<sup>24</sup> <sup>25</sup>. En muchas ocasiones, los desarrolladores son hombres blancos americanos, y los datos están alimentados procedentes del mismo segmento social, por lo que hay falta de diversidad en los equipos y en los datos de entrenamiento que alimentan los algoritmos.

De acuerdo al European Protection Board, "la segmentación de los usuarios de las redes sociales puede suponer que, directa o indirectamente, se produzcan efectos discriminatorios relacionados con la raza u origen étnico, estado de salud, orientación sexual u otras cualidades concernientes a la persona. Como por ejemplo, el uso de dichos criterios en el contexto de publicidad sobre ofertas de trabajo, vivienda o créditos (préstamos e hipotecas), que puede reducir la visibilidad de oportunidades a personas de ciertos grupos. El perfilado puede perpetuar la existencia de estereotipos y segregación racial. Y puede encerrar a una persona en una determinada categoría y restringirle las preferencias sugeridas. Esto puede perjudicar su libertad de elegir, por ejemplo, respecto a ciertos productos o servicios como libros, música o noticias. En algunos casos, el perfilado puede llevar a predicciones inexactas y en otros casos a denegar bienes o servicios con en una injustificada discriminación.<sup>26</sup> <sup>27</sup>.

Asimismo, existe en muchas ocasiones un black box o caja negra entre los datos de entrenamiento y las conclusiones obtenidas de tal manera que es imposible motivar por qué la IA ha llegado a cierta conclusión o aprender la razón de alcanzar un objetivo.

---

<sup>24</sup> TORRES LÓPEZ, Leopoldo Salvador y PERALTA GUTIÉRREZ, Alfonso. Uso policial de sistemas de inteligencia artificial en el ámbito comparado. CDF Cuadernos Digitales de Formación 3/2021 El Derecho y la Inteligencia Artificial.

<sup>25</sup> TORRES LÓPEZ, Leopoldo Salvador y PERALTA GUTIÉRREZ, Alfonso. Herramientas de inteligencia artificial en el ámbito jurídico comparado. CDF Cuadernos Digitales de Formación 7/2020 El Derecho y la Inteligencia Artificial.

<sup>26</sup> Guidelines 8/2020 on the targeting of social media users. Version 2.0. Adopted on 13 April 2021. European Data Protection Board. [https://edpb.europa.eu/system/files/2021-04/edpb\\_guidelines\\_082020\\_on\\_the\\_targeting\\_of\\_social\\_media\\_users\\_en.pdf](https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_082020_on_the_targeting_of_social_media_users_en.pdf)

<sup>27</sup> Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. Adopted on 3 October 2017. As last Revised and Adopted on 6 February 2018. ARTICLE 29 DATA PROTECTION WORKING PARTY. 17/EN WP251rev.01 [http://ec.europa.eu/justice/data-protection/index\\_en.htm](http://ec.europa.eu/justice/data-protection/index_en.htm)



El reto tecnológico es aprovechar ese ágora de mayor información, comunicación, libertad y espontaneidad, reduciendo la mentira, el odio, la falsedad, el acoso o la discriminación.

Debido a este tipo de errores del algoritmo se produce un doble problema en la moderación de contenidos, los denominados "falsos positivos" y los "falsos negativos".

Los "falsos positivos" son aquellos contenidos eliminados que no debían de haberlo sido. Esto puede afectar a la libertad de expresión y de información y el derecho a impartir ideas de toda clase sin restricciones. Recordemos el Art. 19 de la DUDH, "todo individuo tiene derecho a la libertad de opinión y de expresión; este derecho incluye el de no ser molestado a causa de sus opiniones, el de investigar y recibir informaciones y opiniones, y el de difundirlas, sin limitación de fronteras, por cualquier medio de expresión".

Como decíamos, supone una tecnocensura sobre un contenido que no sólo no es ilegal sino que además puede ser conforme con las políticas de moderación de contenidos de la propia plataforma.

En los últimos tiempos estamos asistiendo a "hackeos" masivos de las reglas de moderación de Twitter a través de grupos organizados con ideología de "ultraizquierda", a través de grupos de Telegram con un par de centenares de suscriptores, y muy posiblemente otro par de centenares de bots, que se dedican a reportar masivamente cuentas supuestamente de "ultraderecha", para "limpiar" la red social, consiguiendo el bloqueo de contenido e incluso la suspensión de la cuenta (baneo) tanto temporal como definitiva inclusive de parlamentarios y diputados<sup>28 29 30</sup>.

El procedimiento es bastante sencillo, mediante la herramienta de "búsqueda avanzada" de Twitter u otras

---

<sup>28</sup> Twitter bloquea la cuenta de un diputado de Vox por un tuit con la palabra 'machirulo'. EL INDEPENDIENTE . Sección España. 13 de enero de 2021. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.elindependiente.com/espana/2021/01/13/twitter-bloquea-la-cuenta-de-un-diputado-de-vox-por-un-tuit-con-la-palabra-machirulo/>

<sup>29</sup> Twitter bloquea a un diputado de Vox y Echenique aprovecha para mofarse. 25 abril 2022. Esdiario. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.esdiario.com/opinion/47923493/twitter-bloquea-diputado-vox-echenique-aprovecha-mofarse.html>

<sup>30</sup> GARCÍA NAVIA, Alejandra. Grupos prorrusos consiguen bloquear las cuentas de Twitter de varios líderes de opinión. 13 de abril 2022. The Objective. [Consultado el 15 de mayo de 2022] Disponible en: <https://theobjective.com/espana/2022-04-13/grupos-prorrusos-twitter-opinion/>

herramientas de análisis de redes sociales OSINT se busca un tuit que pueda tener alguna palabra sacada de contexto<sup>31</sup>, utilizada irónicamente, un chiste, una broma con alguna palabra "reportable", o incluso un mensaje con supuestos "datos personales" o con menciones sucesivas de tal manera que pueda atacarse como spam. Así, esos cientos de cuentas o de bots, reportan masivamente la cuenta y el tuit marcándolo como spam, datos personales o incitación al odio principalmente, de tal manera que "hacen saltar" el algoritmo. Reportadores o "baneadores" profesionales, que bien podrían llamarse reportadores o señalizadores negros, o más bien rojos (por su denominación y el color de su logotipo, pájaro rojo, en contraposición al pájaro azul de Twitter). De esta manera consiguen el silenciamiento del adversario, del oponente, hackeando el algoritmo utilizando un contenido legal y encima conforme a las reglas de moderación contractualmente previstas.

Y si tenemos en cuenta la posible adicción a las redes sociales, y el ansia desmesurado de notoriedad en las mismas, supone llevar a la persona a una especie de "muerte civil del siglo XXI" similar a la que se daba en la Edad Media.

Hay que añadir además, que en estos días, hemos visto, como el multimillonario Elon Musk reclama ahora "detalles que respalden el cálculo" y corroboren que las cuentas falsas y de spam representan menos del 5% de los usuarios activos y monetizables, pues podría ser un porcentaje superior<sup>32</sup>.

Y esta forma de silenciamiento puede constituir además una forma de intimidar, humillar o amenazar. O retiras la cuenta, o dejas de publicar cierto contenido ideológico, o vamos a conseguir que te quedes sin perfil, con todo lo que eso conlleva.

Todavía se puede encontrar como si fuera un trofeo la captura de pantalla del reporte masivo de un tuit de un

---

<sup>31</sup> MIGUEL TRULA, Esther. Twitter tiene problemas con el lenguaje no literal y banea a quien dice "te mato" en sentido figurado. 25 Septiembre 2018. Xataka [Consultado el 15 de mayo de 2022] Disponible en: <https://magnet.xataka.com/preguntas-no-tan-frecuentes/no-digas-te-mato-broma-podrias-ser-proxima-victima-censura-twitter>

<sup>32</sup> PREGO, Carlos. Twitter no sabe cuántos bots hay entre sus 61 millones de usuarios. Ahora Musk tiene un plan para averiguarlo. Xataka. 14 Mayo 2022. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.xataka.com/empresas-y-economia/twitter-no-sabe-cuantos-bots-hay-sus-61-millones-usuarios-ahora-musk-tiene-plan-para-averiguarlo>

magistrado jubilado del Tribunal Supremo, ahora dedicado a la abogacía, que era un prolífico tuitero y que fue objeto de una de estas campañas de reportes masivos.

Y es que en estos grupos de Telegram de reportadores "rojos" (por el color de su logotipo, pájaro rojo, RedBird, en contraposición al pájaro azul de Twitter), se congratulan y enorgullecen del número de mensajes que reportan y eliminan de la red, o las cuentas que consiguen suspender, temporal o definitivamente, colgando capturas de sus éxitos como si fuera un salón de trofeos de caza. En el otro extremo político existe otro "ejército" de reportadores, "Hunga Hunga Army" que se dedican literalmente a lo mismo pero en esta ocasión buscan aquellas cuentas de izquierdas.

Pero no sólo un magistrado del Tribunal Supremo fue silenciado, sino que también se ha silenciado a diputados nacionales<sup>33</sup>, medios de comunicación<sup>34</sup> y periodistas<sup>35</sup>. Y hay que recordar que ambos tienen una especial protección del derecho a la libertad de expresión a través de la inviolabilidad parlamentaria o en el caso de los segundos en su función de perro-guardián de las instituciones frente a las asechanzas y los abusos del poder. Sin ir más lejos, todos recordamos el bloqueo de las cuentas en las redes sociales en Facebook y Twitter al expresidente Trump durante el asalto al Capitolio de los Estados Unidos, una censura que puede tener además muchas implicaciones.

Lo que pretenden estos grupos organizados no es más que un "chilling effect", un efecto disuasorio de la libertad de expresión de ciertas personas en las plataformas digitales. Así, en sus grupos puede comprobarse como los reportes de cuentas son sistemáticos hacia ciertas personas que consideran el enemigo, "fascistas" o de "ultraderecha" o "podemitas" o "ultraizquierda". Buscando que al final se acaben cansando de abrir cuentas con el

---

<sup>33</sup> Twitter bloquea la cuenta de un diputado de Vox por un tuit con la palabra 'machirulo'. EL INDEPENDIENTE. 13 de enero de 2021. [Consultado el 29 de octubre de 2022] Disponible en: <https://www.elindependiente.com/espana/2021/01/13/twitter-bloquea-la-cuenta-de-un-diputado-de-vox-por-un-tuit-con-la-palabra-machirulo/>

<sup>34</sup> Twitter expulsa a Estado de alarma TV, el medio de Javier Negre. 12 de agosto de 2022. Sport. [Consultado el 29 de octubre de 2022] Disponible en: <https://www.sport.es/es/noticias/television/twitter-expulsa-alarma-tv-medio-14269312>

<sup>35</sup> ALONSO Lucas, Aitana. Javier Negre recupera su cuenta de Twitter tras ser cerrada por difundir información privada. 3 Agosto 2022. Formula TV. [Consultado el 29 de octubre de 2022] Disponible en: <https://www.formulatv.com/noticias/javier-negre-recupera-twitter-informacion-privada-117554/>

trabajo que supone volver a recuperar los antiguos seguidores y además, vulnerando esta vez sí, las normas de uso, pues la expulsión de la red social no permite la sucesión con otros perfiles. No es descartable que esto mismo lo hagan también grupos de "ultraderecha" hacia personas de "ultraizquierda" o "prorrusos" frente a "proucranianos" y viceversa.. La periodista de El Mundo Lucía Méndez<sup>36</sup> reconoció públicamente que ella misma -y muchos periodistas- se ha autocensurado de escribir sobre el exvicepresidente del Gobierno Pablo Iglesias, puesto que ello suponía exponerse a la "ira de sus devotos en las redes".

Y todo esto en el marco del siglo XXI donde hemos de añadir a la propia tendencia humana a la vanidad, psicólogos y expertos del mundo de la tecnología vienen advirtiéndolo, desde hace tiempo, que el uso frecuente y continuado de las redes sociales puede resultar en una adicción, lo que, a su vez, podría derivar en ansiedad, dependencia emocional, pérdida de motivación, falta de autocontrol e irritabilidad. Un estudio realizado por la Chicago Booth School of Business (2012) señalaba que las redes sociales tienen una capacidad de adicción mayor incluso que la del tabaco porque, entre otros motivos, acceder a estas plataformas resulta sencillo y, además, gratuito.

Está demostrado que los procesos sociales en red aumentan la producción de dopamina un neurotransmisor que se produce en estados de placer y que abre los canales cognitivos, es así como se explica también las adicciones a la tecnología y la dependencia emocional de las redes sociales (Rodríguez, 2017).

Algunos autores consideran que "la necesidad de estar conectado o de mirar constantemente los likes puede convertirse en un pozo de vulnerabilidad del que será muy difícil salir porque nunca sacian a la persona, al contrario siempre se quiere más (Catalán, 2020)"<sup>37</sup>. Hay estudios que consideran que "una semana desconectado de las

---

<sup>36</sup> MÉNDEZ, Lucía. Examen de conciencia. El Mundo. Sección Opinión, Asuntos Internos. 7 de mayo de 2022. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.elmundo.es/opinion/columnistas/2022/05/07/62755e4b21efa071158b4594.html>

<sup>37</sup> .MARTÍN CRITIKIÁN, D. Y MEDINA NÚÑEZ, M. (2021). Redes sociales y la adicción al like de la generación z. Revista de Comunicación y Salud, 11, 55-76 [Consultado el 15 de mayo de 2022] Disponible en: <https://doi.org/10.35669/rcys.2021.11.e281>

redes sociales es suficiente para notar efectos positivos en la salud mental” con mejoras significativas en el bienestar, la depresión y la ansiedad<sup>38</sup>. Esto está produciendo en los jóvenes y no tan jóvenes una dependencia a las redes sociales y a la opinión de los demás que puede llegar a ser perjudicial, intentando mostrar solo el lado bueno de las cosas en lugar de mostrarse como son, así como la necesidad social imperiosa de aceptación a través de las redes sociales. Ya hay clínicas de desintoxicación que tratan la adicción a redes sociales para aquellos que se convierten en yonkis y esclavos de la notoriedad con un afán desmesurado de aceptación y reconocimiento social.

En este sentido, es archiconocida la frase de “si no pagas por el producto, es que el producto eres tú”. Y más aún a raíz de la inteligencia artificial cuyo funcionamiento requiere una gran base de datos -big data-, que se alimenta de textos, imágenes, vídeos, audios, datos de usuario, metadatos, datos del comportamiento, hábitos, datos biométricos y cualesquiera otros tipos de datos, como datos médicos. Los datos son el petróleo del siglo XXI.

El sistema es fácil de comprender, cada búsqueda en la plataforma digital o en el propio Internet es rastreado, los datos son almacenados y con esos datos se construye un perfil tuyo que después es utilizado para dirigir publicidad personalizada. Tracking-profiling-targeting<sup>39</sup>.

Kosinski y su equipo refinaron los modelos predictivos y en 2012 demostró que sobre la base de un promedio de 68 "me gusta" de Facebook de un usuario, era posible predecir su color de piel (con un 95% de precisión), su orientación sexual (88% de precisión) y su afiliación al partido demócrata o republicano (85 por ciento). Pero no se detuvo allí. Inteligencia, afiliación religiosa, así como el consumo de alcohol, cigarrillos y drogas, todo podría determinarse. A partir de los datos, incluso fue posible deducir si los padres de alguien estaban divorciados. En poco tiempo, pudo evaluar a una persona mejor que un colega de trabajo promedio, simplemente sobre la base de diez "me

---

<sup>38</sup> JIMÉNEZ DE LUI, ÁNGEL S. El secreto de la felicidad está en desconectarse de las redes sociales. 9 mayo 2022. El Mundo. Sección Tecnología. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.elmundo.es/tecnologia/2022/05/09/62797276fc6c83373f8b45a4.html>

<sup>39</sup> DANS, Enrique ¿Prohibir la publicidad hipersegmentada?, 20 de abril de 2022. Invertia. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.enriquedans.com/wp-content/uploads/2022/04/Prohibir-publicidad-Invertia.pdf>

gusta" de Facebook. Setenta "me gusta" eran suficientes para superar lo que sabían los amigos de una persona, 150 lo que sus padres sabían, y 300 "me gusta" lo que su pareja sabía. Más "Me gusta" podrían incluso superar lo que una persona creía saber sobre sí mismo. El día que Kosinski publicó estos hallazgos, recibió dos llamadas telefónicas: una amenaza de demanda y una oferta de trabajo. Ambas de Facebook<sup>40</sup> (Grassegger y Krogerus, 2017).

Es más, ingenieros de Facebook<sup>41</sup> han reconocido que aunque se anuncien como una plataforma cuya misión es conectar a la gente, no es verdad, porque sus algoritmos no están desarrollados para eso, sino para optimizar el tiempo de interacción de cada usuario por separado, y ese tipo de optimización favorece otro tipo de cosas. Pensamos que redes como Facebook son plataformas de comunicación cuando, en realidad, son plataformas publicitarias. El fin último de lo adictivo de las plataformas es una mayor monetización. No es casualidad que cuando uno inicie ciertas aplicaciones, no se pueda ver simultáneamente el reloj del teléfono o la tablet. Hay que salir de la app para ver la hora. Es un efecto similar al de los casinos o los centros comerciales, cerrados, con luz artificial, para no poder ver si es de día o de noche, si anochece o sale el sol y por lo tanto darse cuenta de cuánto tiempo lleva uno dentro.

En palabras de PARISER (2017) <sup>42</sup> "cuanta más información personal relevante seas capaz de ofrecer, más espacios publicitarios podrán vender, y en consecuencia, más probabilidades de que compremos los productos que nos están vendiendo".

En el caso de los "falsos negativos" - contenido que no es eliminado, que debía de haber sido eliminado- el problema es mucho mayor. Desinformación, bulos y propaganda están en las agendas de Seguridad Nacional de los distintos

---

<sup>40</sup> GRASSEGGER, H. & KROGERUS, M.:The Data That Turned the World Upside Down How Cambridge Analytica used your Facebook data to help the Donald Trump campaign in the 2016 election. Motherboard. (2017, 28 de Enero) [Consultado el 15 de mayo de 2022] Disponible en: [https://motherboard.vice.com/en\\_us/article/mg9vvn/how-our-likes-helped-trump-win](https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win)

<sup>41</sup> HOLMES Aaron. Un ejecutivo de Facebook asegura que lo crearon para ser "tan adictivo como el tabaco" y ahora teme que por ello EEUU pueda estar "al borde de una guerra civil"., Business Insider. 25 sep. 2020 [Consultado el 15 de mayo de 2022] Disponible en: <https://www.businessinsider.es/facebook-diseno-ser-adictivo-como-tabaco-ejecutivo-724243>

<sup>42</sup> PARISER, Eli. El Filtro Burbuja: Cómo la web Decide lo que Leemos y lo que Pensamos. Taurus, 2017. Pág 17.

países hace mucho tiempo. Se llaman "amenazas híbridas" y son uno de los desafíos más importantes en ciberseguridad. Intentar minar los valores fundamentales europeos como la dignidad humana, la libertad o la democracia. La posverdad y la propaganda puede afectar los procesos democráticos y elecciones, causa polarización y opiniones radicales en la sociedad o debido a agentes externos o extranjeros se intenta dañar, humillar o desestabilizar países<sup>43</sup>. En España, hemos visto investigaciones periodísticas de intentos de la Federación Rusa que intentaron desestabilizar el país colaborando con los movimientos separatistas catalanes extendiendo su propaganda y bulos utilizando más de 4800 bots<sup>44</sup>. En el caso de Birmania, las Naciones Unidas afirmó que Facebook tuvo "un papel determinante" en azotar la rabia contra la minoría rohingya<sup>45</sup>.

Siguiendo esta estela, debemos mencionar el concepto de "astroturfing" o estrategia de posverdad. Se trata de una técnica de comunicación usada por las organizaciones que consiste en crear falsos movimientos para manipular y difundir desinformación, intentando influir en la sociedad. Algunos métodos de astroturfing son crear contenido masivo (content seeding), usar bots a través de un troll center o crear un ecosistema falso. Esto constituye un negocio enorme y existen agencias nacionales e internacional de astroturfing<sup>46</sup>.

Muy probablemente una de las principales razones del crecimiento del populismo en muchos países es la potenciación en redes sociales de un clima de desconfianza y descrédito de los ciudadanos en las instituciones del estado e incluso en la democracia misma.

Una de las principales diferencias de las plataformas digitales y el periodismo es que los algoritmos amplifican

---

<sup>43</sup> PERALTA GUTIÉRREZ, Alfonso. La verdad no justifica la censura. El juicio comparado de derechos en materia de desinformación, seguridad nacional y libertad de expresión (1). Diario La Ley, ISSN 1989-6913, N° 9772, 2021. [Consultado el 15 de mayo de 2022] Disponible en: <https://diariolaley.laleynext.es/dli/2021/01/18/la-verdad-no-justifica-la-censura-el-juicio-comparado-de-derechos-en-materia-de-desinformacion-seguridad-nacional-y-libertad-de-expresion>

<sup>44</sup> GALÁN Javier, ABAD LIÑÁN José Manuel, ALAMEDA David. Los 4.800 bots que jalearon el 'procés'. El País. Sección Política. Madrid. 10 DIC 2017. [Consultado el 15 de mayo de 2022] Disponible en: [https://elpais.com/politica/2017/12/04/actualidad/1512389091\\_690459.html](https://elpais.com/politica/2017/12/04/actualidad/1512389091_690459.html)

<sup>45</sup> The country where Facebook posts whipped up hate. BBC Trending. 12 September 2018. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.bbc.com/news/blogs-trending-45449938>

<sup>46</sup> Confesiones de un bot ruso. Debate. 2022.

aquellas opiniones más extremas, polarizadas, extrañas, controvertidas o sesgadas porque atraen más la atención de los usuarios y por tanto, hacen que éstos pasen más tiempo conectados. Esta es la razón de por qué en las redes sociales se expanden todo tipo de teorías conspiratorias. Facebook ha confesado que su algoritmo explota la atracción del cerebro humano a la división, la polarización.<sup>47</sup> Según un informe interno realizado en 2016, el 64% de la gente que se unió a un grupo extremista lo habría hecho porque el algoritmo de la plataforma se lo habría recomendado, según investigaciones del Wall Street Journal<sup>48</sup>. Por lo tanto las fake news se hacen más virales, corren más rápido que las verdaderas, creando un estercolero de contenido de basura en las plataformas. De esta manera el algoritmo alimenta contenidos antiacadémicos, antiintelectuales que coinciden en el tiempo con ese crecimiento del populismo y de teorías rocambolescas. Las redes sociales alimentan más la reacción emocional que la reacción racional. Con poco margen para los grises y los matices, algo que sin embargo, es el "a, b y c" de un jurista. En cuanto a los algoritmos de recomendación o personalización de contenido, las plataformas digitales organizan, promueven y presentan el contenido para su visualización a los distintos usuarios. La inteligencia artificial es usada para evaluar el contenido de cara a marcarlo, monetizarlo, priorizarlo, eliminarlo o banear (eliminar) la cuenta. El sistema es fácil de entender, todo lo que buscamos en la plataforma (y en Internet) se rastrea (seguimiento o trazabilidad, tracking), y se recolecta un big data. Con esta vasta cantidad de datos se construye tu perfil (profiling), y después se fijan los objetivos para dirigir la publicidad personalizada o híper-segmentada (micro-targeting)<sup>49</sup>.

Además, las plataformas crean el llamado "efecto burbuja" que impide un contraste de opiniones o de versiones distintas de los hechos desde ángulos distintos.

---

<sup>47</sup> STATT, Nick. Facebook reportedly ignored its own research showing algorithms divided users. Facebook fears harming engagement or enraging conservatives. The Verge. 26 de mayo 2020. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.theverge.com/2020/5/26/21270659/facebook-division-news-feed-algorithms>

<sup>48</sup> HORWITZ Jeff y SEETHARAMAN Deepa . Facebook Executives Shut Down Efforts to Make the Site Less Divisive. Wall Street Journal. Sección Tech. 26 mayo 2020. [Consultado el 15 de mayo de 2022] Disponible en: [https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499?mod=hp\\_lead\\_pos5](https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499?mod=hp_lead_pos5)

<sup>49</sup> DANS, Enrique. ¿Prohibir la publicidad hipersegmentada? 20 abril 2022. Invertia. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.enriquedans.com/wp-content/uploads/2022/04/Prohibir-publicidad-Invertia.pdf>



Todo se alimenta desde el mismo punto de vista, la sensación "más de lo mismo", ultrasesgada y con experiencias retroalimentadas (superfeedback)<sup>50</sup>. De acuerdo al Art. 3,1 del Código Civil debemos tener en cuenta en la aplicación de las normas "la realidad social", ¿nos permiten las redes apreciar verdaderamente la realidad social o sólo nuestra realidad social que nos recomienda el algoritmo?

Esto afecta al derecho humano de libertad de expresión que incluye, "el de investigar y recibir informaciones y opiniones, y el de difundirlas, sin limitación de fronteras, por cualquier medio de expresión" (Art. 19 DUDH y del PIDCP).

Como dice PARISER (2017), con frecuencia la creatividad se produce gracias a la colisión de ideas procedentes de otras disciplinas y culturas, encuentros casuales que aportan conocimientos y aprendizaje. Y por lo tanto, según este autor, un mundo en el que todo nos resulta familiar y que nos retroalimenta, es entonces un mundo en el que no hay nada que aprender. La personalización de contenidos puede impedir que entremos en lucha con nuestros prejuicios y cambiar nuestra forma de pensar. Un determinismo informativo donde lo clicado en el pasado, nos condicione el futuro a repetirlo una y otra vez. Esto impide a la sociedad abandonar la "tribu", convertirse en un ser individual de mente abierta, pensar por sí mismo, evolucionar de pensamientos o ideologías. Todo lo contrario, las plataformas incrementan el sentimiento de pertenencia tribal en grupos específicos. Acaban con la fragmentación social del individualismo y el pensamiento crítico y nos crean la necesidad de adherirnos a grupos ideológicos o -simplemente- emocionales.

BRONCANO y otros autores<sup>51</sup> lo han denominado también "cámaras de eco" o "burbujas epistémicas". Se tratan de sistemas sociales organizados que silencian cualquier opinión divergente, contraria u opuesta a nuestras

---

<sup>50</sup> MERINO RUS, Rafael. Los derechos humanos, la democracia y la igualdad en la era de los algoritmos y la inteligencia artificial El contenido de esta publicación corresponde a una ponencia presentada durante el IV Encuentro de Verano de Gernika-Urdaibai «Desdemocracia, los "ismos" y el abismo. Universales ante las políticas segregatorias», en el marco de los XXXVIII Cursos de Verano UPV/EHU, celebrado en Gernika-Lumo los días 19 y 20 de junio de 2019.

<sup>51</sup> BRONCANO, Fernando (2019). "Encerrados en la burbuja, El laberinto de la identidad, 10 de febrero de 19. [Consultado el 15 de mayo de 2022] Disponible: <<http://laberintodelaidentidad.blogspot.com/2019/02/encerrados-en-la-burbuja.html?m=1>>.

supuestas creencias, o por lo menos las creencias que el algoritmo supone que tenemos.

KOSINSKY y su equipo en 2012 probaron que con lo básico de una media de 68 "me gusta" de Facebook de un usuario, es posible predecir su color de piel (con un 98% de acierto), su orientación sexual (88% de acierto), su orientación política demócrata o republicana (85%). Pero no pararon ahí. El coeficiente intelectual, la religión que se profesa, si consume alcohol, drogas o tabaco también fue posible. De los datos era posible deducir cuando se habían divorciado los padres de alguien. 70 "likes" eran suficientes para superar el conocimiento que de una persona tiene un amigo, 150 para superar cómo te conocen tus padres, y 300 cómo te conoce tu pareja. Más "me gusta" puede suponer que la red social te conozca mejor que tú mismo<sup>52</sup>.

Estos datos además pueden combinarse con psicometrías o psicografías, enfocadas en medir los rasgos psicológicos como la personalidad. En 1980, dos equipos de psicólogos desarrollaron un modelo que buscaba clasificar a todos los seres humanos en 5 tipos psicológicos, conocido como los "Gran 5". Son: la receptividad o amplitud de miras (¿cuánto estás dispuesto a nuevas experiencias?), meticulosidad (¿cómo eres de perfeccionista?), extroversión (¿cómo eres de sociable?), complacencia (¿cómo de considerado, cooperativo y dispuesto a agradar a estás?) y neutoricismo (¿estás realmente enfadado?). Basado en estas dimensiones, que también se llaman OCEAN como acrónimo por sus iniciales en inglés -openness, conscientiousness, extroversion, agreeableness, neuroticism—podemos hacer con una bastante precisa eficacia una clasificación de la persona que tenemos frente a nosotros, lo que incluye los miedos y su posible comportamiento.

Los algoritmos de recomendación para dirigir contenido a los usuarios también pueden ser usados indebidamente para influir en los ciudadanos en materia de debate público y procesos electorales. Mezclando el big data con psicometrías, Cambridge Analytica utilizó millones de datos

---

<sup>52</sup>

GRASSEGGER Hannes & KROGERUS Mikael. The Data That Turned the World Upside Down. 28 de enero de 2017. Vice.com [Consultado el 15 de mayo de 2022]. Disponible en: <https://www.vice.com/en/article/mg9vvn/how-our-likes-helped-trump-win>

de usuarios de Facebook para microsegmentar los anuncios y la campaña electoral de Donald Trump.

Esto tiene dos consecuencias muy importantes:

- los usuarios no saben qué información u opiniones les está siendo excluida de su conocimiento o exhibición debido a su perfilado. Imagínate una carpeta de spam del correo electrónico a la que como sabes van en muchas ocasiones correos correctos marcados erróneamente por el asunto o por el remitente, pero a la que no pudieras acceder ni recuperar ese contenido. Eso puede ser una buena comparación con los algoritmos de recomendación de contenido. Y además uno no puede elegir entrar o no en la burbuja, simplemente te meten (PARISER 2017 53). En la investigación realizada por este autor, entrevistó a Jonathan McPhie, hombre clave en Google sobre la personalización, y éste sugirió que resultaba casi imposible adivinar cómo los algoritmos determinarían la experiencia de usuario. Simplemente serían demasiadas variables y datos que rastrear y determinar.

- los usuarios no pueden ejercer control sobre los datos almacenados que han sido utilizados para su perfilado. Como dice PARISER (2017), Google no te dice quién cree que eres o por qué te muestra los resultados que ves. No sabes si lo que se supone que sabe acerca de ti es correcto o incorrecto, y puede que ni siquiera seas consciente de que está haciendo conjeturas sobre ti. Tampoco puedes conocer si por ejemplo una búsqueda de un viaje a un territorio peligroso -simplemente aunque sea por curiosidad- puede suponer un aumento de la prima por parte de la compañía aseguradora, como sugiere el catedrático ZITTRAIN<sup>54</sup>.

Asimismo, los medios de comunicación tienen un rol social en contribuir a un clima favorable de debate público, contribuyendo a la diversidad y proveyendo una vía para canalizar ideas y opiniones y crear debate. Mientras que las plataformas digitales contribuyen a lo contrario, sólo se puede debatir sobre los mismos intereses con personas de similares puntos de vistas, y siendo limitado

---

<sup>53</sup> PARISER, Eli. El Filtro Burbuja: Cómo la web Decide lo que Leemos y lo que Pensamos. Taurus, 2017.

Pág 19.

<sup>54</sup> ZITTRAIN, Jonathan. The Future of Internet -And How to Stop it. New Heaven, Yale University. 2008.

sobre otras noticias o cuestiones importantes y reduciendo la exposición a contenido informativo diverso.

En España fue adoptada una reforma de la LOREG, Ley Orgánica de Régimen Electoral General en 2018 sobre la posibilidad de trazabilidad de opiniones políticas por partidos políticos para perfilar a ciudadanos basada en el "interés público y con suficientes garantías". Este precepto legal fue declarado inconstitucional por el Tribunal Constitucional<sup>55</sup> por su abstracción, vaguedad y falta de concreción.

Además, los continuos trending topics, crean una "presión sistémica y estructural" sobre lo que decimos, opinamos, comentamos e incluso debemos pensar. Debemos comentar todo aquello que el algoritmo nos sugiere que es trending topic, porque es la manera de aumentar seguidores, crear contenido y crecer en influencia, manteniendo la participación. Cada hora, o cada pocos minutos, se introduce un nuevo y urgente tema del cual debemos opinar. Según SOTO IVARS<sup>56</sup> esto es "la tiranía de las redes sociales debido a la presión del algoritmo". Y además esto produce una sobreinformación debido a que los individuos no pueden crearse una decisión formada e informada debido a la abundancia de información y de temas sobre los que hay que opinar -y en muchas ocasiones de manera poco fiable e ilustrada.

En este sentido cabe mencionar también lo que PARISER (2017) menciona como "dinámica de la irreversibilidad", descrita por Bob Metcalfe, inventor del protocolo Ethernet que conecta a los ordenadores. Como sabemos, el valor de una plataforma digital depende de su número de usuarios. Y cuando una tecnología o empresa es asumida por la generalidad de la población, no estar puede suponer una gran desventaja. Si todo el mundo está en Facebook, es más fácil que los usuarios se conecten más tiempo y más veces. Si todos los servicios que tienes y tu propio teléfono están asociados a Google o Apple, es más difícil que te pases a los mismos servicios pero de otras compañías.

---

<sup>55</sup> Sentencia del TC 76/2019, de 22 de mayo. Ref. BOE-A-2019-9548.

<sup>56</sup> SOTO IVARS, Juan. ¿Quién elige sobre qué asunto piensas hoy? J23 de marzo de 2022. El Periódico. Opinión. [Consultado el 15 de mayo de 2022]. Disponible en: <https://www.elperiodico.com/es/opinion/20220323/articulo-juan-soto-ivars-redes-sociales-libertad-agenda-13418809>

Los algoritmos de recomendación y personalización si bien pueden ser no sólo una herramienta más eficaz, dirigida verdaderamente a los clientes potenciales, y además un contenido más personalizado para satisfacer mejor nuestros intereses, no obstante, también hay autores que lo consideran como hacernos prisioneros de nuestros pensamientos, intereses y preferencias para siempre, sin posibilidad de crecer, evolucionar o progresar. BENKLER sostiene que para ser dueños y autores de nuestra vida, debemos ser conscientes de una amplia gama de opciones y estilos de vida<sup>57</sup>. Si el algoritmo sólo nos recomienda conocer a aquellas personas más afines a nuestros pensamientos e ideas, nos rodeamos únicamente de comunidades homogéneas, y como dice PARISER (2017) es mucho menos probable que entremos en contacto con personas con distintos puntos de vista.

Es más, según SCHULZ<sup>58</sup> los algoritmos de recomendación podrían estar afectando al sistema cognitivo el cual se basa en el conocimiento y fortalecimiento de las ideas existentes y conocidas y la adquisición de otras nuevas. Los filtros de personalización nos cercan con ideas ya conocidas, con las que estamos familiarizados y de acuerdo, induciéndonos a un exceso de confianza que podamos creer que estamos en lo conocido, cuando quizás sean ideas nuevas que aprendamos. Y además, eliminaría elementos de querer aprender. De esta manera, si nuestras ideas preconcebidas son por la creencia de una alta criminalidad, en la noticia de un asesinato, recordaremos la apariencia de la víctima, pero no por ejemplo la mención a que las tasas de delincuencia son bajas. Así, en experimentos desarrollados<sup>59</sup>, se comprobó suministrada una determinada información, los participantes tendían a confirmar lo que veían con esa información previamente suministrada, y la información que no cuadraba con lo que les habían dicho tendía a olvidarse a menudo. Doris Graber<sup>60</sup> descubrió incluso que los esquemas pueden ser tan poderosos que pueden llevar al cerebro a fabricar información, recordando

---

<sup>57</sup> BENKLER Yochai, "Siren songs and amish children: autonomy, information and Law". New York University. Law Review, abril de 2001.

<sup>58</sup> SCHULZ Kathryn. Being wrong. Adventures in the Margin of Error. Nueva York. HarperCollins 2010. (Traducción castellano "En Defensa del error". Madrid. Siruela. 2015.

<sup>59</sup> BRECKER, Steven James, OLSON James M. WIGGINS, Elizabeth Corinne. Social Osycology Alive. Belmont (California). Thomson Wadsworth 2006. p. 69.

<sup>60</sup> GRABER, Doris. Processing the news. p. 170.

los participantes del experimento detalles que ni siquiera habían existido pero que se asociaban a sus esquemas mentales.

Una vez que hemos adquirido ciertos esquemas, estamos predispuestos a reforzarlos y es lo que los psicólogos conocen como "sesgo de confirmación", tendencia a reforzar nuestros puntos de vista existentes. Así, el filtro burbuja tendería a amplificar ese sesgo. No es más fácil, más sencillo, e incluso más agradable adecuar la información que recibimos a nuestra visión del mundo que consumir aquella que nos invite a pensar de otra manera o cuestionar nuestra forma de pensar (PARISER 2017). "La abolición de la libertad y de la espontaneidad humana" (ARENDDT)<sup>61</sup>. Una predeterminación tecnológica que nos deniega la posibilidad de escapar y altera nuestra libertad y juicio crítico, una inhibición inducida de la libertad que propicia una mutación inconsciente que modifica la esencia de lo humano. Según DEWEY, "los obstáculos a la libertad y plenitud de comunicación levantan barreras que dividen a los seres humanos en [...] sectas y facciones antagónicas, y por lo tanto, socavan el estilo de vida democrático". De esta manera esa polarización y extremismos que se están dando en las redes sociales serían consecuencia de los sistemas de recomendación <sup>62</sup>. Además, los datos y su aprovechamiento pretenderían apartar al ser humano de ser quienes propiamente escriban su propia historia. Según LOEWENSTEIN<sup>63</sup>, "la curiosidad se despierta con cierta laguna de información", pero si no somos conscientes de qué información u opiniones nos está siendo excluida de su conocimiento o exhibición debido a su perfilado, no nos vemos tan obligados a ser curiosos. De manera similar opina VAIDHYANATHAN<sup>64</sup> "aprender es por definición un encuentro con lo que no sabemos, con lo que no hemos pensado, no podíamos concebir y que nunca entendimos o sopesamos como posible". Algo que por definición los algoritmos de recomendación impiden.

---

<sup>61</sup> ARENDT, Hannah. Los orígenes del totalitarismo. Taurus. Ciudad de México 2004.

<sup>62</sup> DEWEY, John. Essays, Reviews, and Miscellany, 1939-1941, The Later Works of John dewey 1925-1953, vol. 14, Carbondale, Southern Illinois University Press, 1998, p. 227.

<sup>63</sup> LOEWENSTEIN George. "The psychology of curiosity: a review and reinterpretation". Psychological Bulletin 116, nº 1 1994. P. 75-98. [Consultado el 2 de noviembre de 2022]. Disponible en: <https://www.cmu.edu/dietrich/sds/docs/loewenstein/PsychofCuriosity.pdf>

<sup>64</sup> VAIDHYANATHAN, Siva. La googlización de todo (y por qué deberíamos preocuparnos). Editorial Océano, 2012 - 238 páginas. p. 182.

A día de hoy, buscando un mismo término en Google por distintas personas no sólo dará distintos resultados, sino que también dará un número dispar de los mismos. Desde diciembre de 2009 los resultados son adaptados al usuario en particular, personalizados.

Pero es más, según PARISER (2017), en un mundo personalizado, cuestiones más complejas o poco agradables, pero que pueden ser esenciales para el debate público y la mejora democrática, es menos probable que nos sean recomendadas y nos llamen la atención, y por lo tanto los algoritmos no sólo nos limitarían la posibilidad de cambiar de opinión o de evolucionar, sino también de formarnos sobre cuestiones complejas y volvernos más inteligentes o con mayor espíritu crítico.

En este sentido, un ejemplo importante sería TikTok, plataforma china que fue en su caso objetivo de la administración americana de Donald Trump el cual mediante una acción ejecutiva en agosto de 2020 la obligó a ser participada por una empresa norteamericana de manera mayoritaria para poder operar en Estados Unidos. Y no fue sólo por los datos que pudiera recabar, el uso y la posible cesión de ellos al Gobierno chino. Sino también, por el contenido que dicha plataforma pudiera difundir. No sólo relativo a cuestiones de desinformación, fake news, bulos o propaganda de agentes extranjeros, como pueden ser teorías de conspiración, o que incluso podría intentar intervenir en las preferencias políticas de cara a las siguientes elecciones sino también, el riesgo tiene que ver con el simple hecho de que el algoritmo de TikTok es conocido por promover contenido viral generalmente inane y banal, que hace que los jóvenes pasen gran parte de su tiempo libre enganchados a los estímulos de dopamina del algoritmo, o buscando ellos mismos crear videos alineados a las tendencias, todas ellas usualmente triviales y sin ninguna trascendencia, a fin de obtener likes, creando así una especie de circuito de banalidad y entretenimiento vaciado de significado ulterior. Según HARA<sup>65</sup>, "el hecho de que esto sea algo negativo para la mente de los jóvenes es claramente entendido por China". Es revelador que TikTok,

---

<sup>65</sup> HARA, Luis Alberto. Tiktok es un peligro para la mente de los jóvenes y para la seguridad nacional: Estados Unidos investiga a la red social china. 09/09/2022. Pijamasurf. [Consultado el 2 de noviembre de 2022]. Disponible en: [https://pijamasurf.com/2022/09/tiktok\\_peligro\\_mente\\_jovenes\\_seguridad\\_nacional\\_estados\\_unidos\\_investiga\\_china/](https://pijamasurf.com/2022/09/tiktok_peligro_mente_jovenes_seguridad_nacional_estados_unidos_investiga_china/)

junto a otras redes sociales, está prohibido en China. En China, ByteDance tiene una versión llamada Douyin que está regulada por los censores chinos.

Así, ByteDance creó Douyin en 2016 para el mercado chino. El año siguiente, con la intención de expandir la red social al resto del mundo, lanzó su versión occidental TikTok y la fusionó con Musical.ly, llegando a tener un auge en 2020 con la llegada del distanciamiento social. TikTok cuenta con millones de usuarios; no obstante, Douyin y TikTok son apps completamente separadas, siendo la primera únicamente disponible en China y la segunda prohibida en China<sup>66</sup>. Así, respecto al contenido, En Douyin existe una mayor variedad de temas, muchos de ellos incluso relacionados con el aprendizaje y el uso de la plataforma como e-commerce, mientras que en TikTok lo que prima es el entretenimiento. En TikTok es más fácil encontrar videos divertidos, a menudo inspirados por los numerosos desafíos lanzados por los usuarios. El objetivo es principalmente conseguir viralidad y por tanto un elevado número de visualizaciones. En Douyin, en cambio, predominan los vídeos sobre animales, cocina y consejos sobre estilo de vida, belleza y moda. La propia versión occidental de TikTok, desde abril de 2020, está prohibida en la China continental por el mecanismo de censura oficial. Según WENTAO<sup>67</sup> "además del contenido de entretenimiento, también son populares (en Douyin) las historias, el entretenimiento (canto, baile y maquillaje), videos de divulgación científica y videos de habilidades". Incluso Douyin tiene entre sus modelos de negocio, la posibilidad de intercambio de conocimientos / formación en curso, compartiendo conocimientos profesionales y luego obtenga ganancias a través de la venta de cursos, capacitación y otros servicios, algo que no parece existir en Tik Tok.

WENTAO distingue que el contenido de Douyin se centra principalmente en cortometrajes elaborados, charlas tipo conferencias, fragmentos de series o películas, y cante y

---

<sup>66</sup> Así de diferente es Douyin, el TikTok exclusivo de China. El Comercio. Perú. 15 de agosto de 2022. Consultado el 2 de noviembre de 2022]. Disponible en: <https://elcomercio.pe/tecnologia/redes-sociales/asi-de-diferente-es-douyin-el-tiktok-exclusivo-de-china-redes-sociales-apps-espana-mexico-colombia-noticia/>

<sup>67</sup> WENTAO, Zhang. Diferencias entre China y Occidente en la creación de contenido en plataformas sociales de video. Trabajo Final De Grado. Universitat Politècnica De València. Escola Politècnica Superior De Gandia. Master en Comunicació Transmedia. Tutor/a: Jorge Ignacio Serrano Cobos. GANDIA, 2020 [Consultado el 2 de noviembre de 2022]. Disponible en: <https://riunet.upv.es/handle/10251/175091>



baile. Mientras que Tik tok su contenido es principalmente, influencers guapos y guapas cuyo estilo de vida principalmente es rico y refleja autos de lujo, compras de ropa o de bolsos, casas de lujo o fiestas, cante, baile, desafíos (challenges) y memes. Es decir, contenido viral poco formativo principalmente en Tik tok, a diferencia de Douyin.

En palabras de HARA, "Es como si China supiera que tiene una droga que crea daño cerebral y permite que sea consumida en otros países pero no en el suyo." Asimismo, la versión china de TikTok establece un límite de 40 minutos al día para los menores de 14 años. Este límite, implementado en China, no solo permite que utilicen la red social 40 minutos al día sino que también restringe su uso en horario nocturno, de 10 de la noche a 6 de la mañana<sup>68</sup>.

Como dice KLEIN<sup>69</sup> un artículo en el New York Times, dónde centramos nuestra atención interesa a los gobiernos, y mucho.

En palabras del responsable de Google, Eric Schmidt "La tecnología será tan buena que será muy complicado ver o consumir algo que no haya sido en algún sentido sugerido para ti"<sup>70</sup>. Según LASSALLE<sup>71</sup> la Historia humana ha sido el resultado de una gestión eficiente de datos que hacía el hombre orientando su vida a partir de la información obtenida, organizándola y procesándola. Pero a medida que se incrementa la confianza en el motor de búsqueda, disminuye nuestra capacidad para buscar respuestas por nosotros mismos. PARISER (2017) cita al periodista Clive Thompson quien considera "es una virtud cívica el estar abierto a aquello que puede estar fuera tus intereses".

Cómo hemos visto, los algoritmos pueden restringir las recomendaciones y sugerencias a una persona, su libertad de elección y de libre acceso a ideas y opiniones de cualquier clase, sin límites, y encerrarla en una determinada

---

<sup>68</sup> La versión china de TikTok establece un límite de 40 minutos al día para los menores de 14 años. 21/09/2021. Europa Press. [Consultado el 2 de noviembre de 2022]. Disponible en: <https://www.europapress.es/portaltic/socialmedia/noticia-tiktok-establece-limite-40-minutos-dia-menores-14-anos-china-20210920144342.html>

<sup>69</sup> Klein, Ezra. TikTok may be more dangerous than it looks. The New York Times, 8 de mayo de 2022. [Consultado el 2 de noviembre de 2022]. Disponible en: <https://www.nytimes.com/2022/05/08/opinion/tiktok-twitter-china-bytedance.html>

<sup>70</sup> <https://www.businessinsider.com/eric-schmidt-quotes-2013-11>

<sup>71</sup> LASSALLE, José María. Ciberleviatán. El colapso de la democracia liberal frente a la revolución digital. Arpa. 2019. Pág. 41 y 76.

categoría ya predeterminada sin posibilidad de cambio o evolución. En palabras de LASSALLE se nos neutraliza la capacidad decisoria jugando con la sutileza algorítmica que hace que las personas no actúen según su juicio sino que se optimice en su propio beneficio. Una neutralización de la capacidad decisoria que el exSecretario de Estado califica como "totalitaria".

En palabras de la Secretaria de Estado de Digitalización y de Inteligencia Artificial del Gobierno de España, Carme Artigas<sup>72</sup>, "si no tenemos oportunidad de ver un Porsche porque los algoritmos determinan que no podemos pagarlo, ¿dónde quedan nuestros sueños, nuestra capacidad de progreso, la búsqueda del éxito y las aspiraciones de hacer un sueño realidad? "

Y es que las plataformas digitales y en general el internet (que no sería más que ese reflejo monumental de las grandes tecnológicas) no mostrarían lo que realmente pasa en el mundo, sino lo que esas empresas y sus algoritmos creen que debes ver. En palabras de GOODMAN<sup>73</sup> "si buscas algo y no lo encuentras, ¿cómo puedes saber de su existencia". Este autor parafrasea una antigua pregunta filosófica: "si un árbol cae en Internet y ningún motor de búsqueda lo indexa, ¿emite algún ruido?". Cada vez más, si algo no existe no en Internet sino en Google, Facebook, Twitter o Instagram, parece que no existe.

Y a la vez debido a la desinformación, bulos y propaganda puede dar otro fenómeno curioso, que no se indexe aquello que ha ocurrido, pero sin embargo que sí aparezca aquello falso y que nunca ha ocurrido.

En resumen, hoy en día, las opiniones son más libres y más espontáneas pero también más sesgadas, subjetivas y sectarias. La libertad de expresión es más amplia que nunca, pero a la vez más estrecha, menos global, cosmopolita, y abierta de mente, convirtiéndose en todo lo contrario, nacionalista, populista y extremista.

La controversia sobre las declaraciones de Elon Musk sobre las políticas de moderación de contenido de Twitter y

---

<sup>72</sup> Ética, IA y Derechos Digitales, 20 de abril 2022. Granada. Law Faculty. [Consultado el 15 de mayo de 2022]. Disponible en: <https://youtu.be/aWzueCq546>

<sup>73</sup> GOODMAN Marc, Los delitos del futuro: Todo está conectado, todos somos vulnerables, ¿qué podemos hacer al respecto? (Ariel). 5 noviembre 2015 Gemma Deza Guil (Traductor).

los cambios que pretende imponer reflejan el problema de que las leyes necesitan interpretación, por los jueces y el "balancing" o ponderación de los derechos de libertad de expresión está lejos de ser simple, pues plantea cuestiones complejas relativas a cómo trasladar los estándares de derechos humanos a reglas particulares.

Siempre habrá asuntos que serán desagradables u ofensivos para algunos usuarios o incluso para las compañías que pagan la publicidad, pero desde una perspectiva de derechos humanos y en palabras del TEDH la libertad de expresión debe desarrollarse sin estrechez ni miedo y las únicas limitaciones a la libertad de expresión deben estar bien concretadas bajo los principios de legalidad, proporcionalidad, necesidad y legitimación. En palabras del Tribunal Constitucional STC 177/2015 (LA LEY 104946/2015) (8), la libertad de expresión debe «desenvolverse sin angostura, sin timidez y sin temor», también en redes sociales. Como subraya el Tribunal Europeo de Derechos Humanos, la libertad de expresión vale no solo para la difusión de ideas u opiniones «acogidas con favor o consideradas inofensivas o indiferentes, sino también para aquellas que contrarían, chocan o inquietan al Estado o a una parte cualquiera de la población». Y justamente ese «inquietar al Estado», viene a ser una definición muy similar al concepto de «amenaza híbrida» utilizado en Seguridad Nacional. Es decir, lo que por una parte socava los valores democráticos, por otra no es sino expresión de la vigorosidad de una sociedad abierta y la democracia en sí misma.

Como ha dicho el Tribunal Constitucional en su sentencia 6/88 (LA LEY 53529-JF/0000) (10): «Las afirmaciones erróneas son inevitables en un debate libre, de tal forma que, de imponerse «la verdad» como condición para el reconocimiento del derecho, la única garantía de la seguridad jurídica sería el silencio.» No existe un bien común que justifique la censura, en cuanto que el único fin superior es la máxima posibilidad de información y es el pleno ejercicio del derecho a la expresión.

Como ya ha declarado el Tribunal Supremo en su doctrina sobre el discurso del odio, «la libertad de opinión ampara incluso a los que ataquen al propio sistema democrático, puesto que, la Constitución, en última instancia protege también a los que la niegan, ni, se puede

prohibir menos aún, la expresión de opiniones subjetivas sobre acontecimientos históricos o de actualidad».

No todo mensaje inaceptable o que ocasiona el normal rechazo de la inmensa mayoría de la ciudadanía ha de ser tratado como delictivo por el hecho de no hallar cobertura bajo la libertad de expresión (Roj: STS 31/2017 (LA LEY 85/2017) –ECLI:ES:TS:2017:31; LA LEY 85/2017, N.o de Recurso: 1619/2016; N.o de Resolución: 4/2017; Fecha de Resolución: 18/01/2017).

Como ha expresado Carolina Botero Marino –exRelatora Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos de la Organización de Estados Americanos– «un sistema de control al derecho de expresión en nombre de una supuesta garantía de la corrección y veracidad de la información puede ser fuente de grandes abusos y, en el fondo, viola el derecho a la información que tiene esa misma sociedad. Resulta en principio contradictorio invocar una restricción a la libertad de expresión como un medio para garantizarla».

Por consiguiente, ningún sistema que se establezca puede contemplar un sistema de censura de la libertad de expresión con la excusa de un control de la calidad, una veracidad informativa, un orden público o una salubridad pues supondría una restricción desproporcionada y altamente peligrosa de la libertad de expresión que pudiera dar lugar a abusos autoritarios y a un blindaje del Gobierno contra las críticas hacia su gestión política.

El riesgo es, que de las declaraciones de miembros del Gobierno antes mencionadas, parecen pretender querer arrogarse aquello justo que no puede hacer, la facultad de establecer qué informaciones u opiniones son «veraces» o «correctas» para ser conocidas por el público como si existiera una única verdad incuestionable lo que llevaría al control gubernativo de los medios y a impedir informaciones investigativas, críticas y debilitar el debate público e intercambio de ideas.

Quien fue Procurador General de los Estados Archibald Cox, afirmó: «Ningún hombre, ningún comité y ciertamente ningún gobierno, cuenta con la infinita sabiduría y el desinterés para separar con exactitud y sin egoísmo lo que

es verdadero de lo que es debatible, y estos dos de lo que es falso»

Es incompatible con la democracia cualquier forma de censura «purificadora» por la cual el poder caprichoso del Gobierno sea el guardián de la verdad y del alcance de la libertad de expresión. Este derecho fundamental tiene una especial dimensión institucional que garantiza una opinión pública libre y es pilar esencial de una sociedad libre y democrática. En palabras del Tribunal Constitucional STC 177/2015 (LA LEY 104946/2015) (8), la libertad de expresión debe «desenvolverse sin angostura, sin timidez y sin temor», también en redes sociales.

Los bulos o noticias falsas no son supuestos internacionalmente admitidos para la limitación de la libertad de expresión e Internet juega un rol primordial en garantizar el acceso público a las noticias y diseminar la información (STEDH Times Newspapers Ltd v. the United Kingdom (nos. 1 and 2), nos. 3002/03 and 23676/03, § 27, 10 March 2009).

Como subraya el Tribunal Europeo de Derechos Humanos, la libertad de expresión vale no solo para la difusión de ideas u opiniones «acogidas con favor o consideradas inofensivas o indiferentes, sino también para aquellas que contrarían, chocan o inquietan al Estado o a una parte cualquiera de la población». Y justamente ese «inquietar al Estado», viene a ser una definición muy similar al concepto de «amenaza híbrida» utilizado en Seguridad Nacional. Es decir, lo que por una parte socava los valores democráticos, por otra no es sino expresión de la vigorosidad de una sociedad abierta y la democracia en sí misma.

No parece haber una bala de plata al respecto.

Algunas herramientas útiles podrían ser unas políticas de moderación más claras sobre qué constituye ofensa o conducta que pueda suponer un bloqueo, enfatizar la importancia del contexto en algunas expresiones como la desnudez, temas históricos, culturales o educacionales, reforzar los matices lingüísticos y culturales en cada mercado. Y para ello una de las soluciones sería incrementar la revisión humana de los algoritmos, con una

apelación humana después de la decisión automatizada de moderación.

Para evitar falsos positivos podría ser bueno el dar la oportunidad de editar el contenido, poder explicar mejor el contexto o simplemente marcar el contenido reportado ("reportado, contenido bajo investigación"), antes de proceder a su eliminación. Así por ejemplo, en justicia, los efectos de la mayoría de casos no se producen hasta la sentencia firme. Y se podrían contratar más moderadores humanos de contenido en distintas lenguas, nacionalidades, culturas y contextos. Con más transparencia, políticas de respeto a los derechos humanos desde el diseño y compliance del resto de principios ético de la IA.

Una de las principales preguntas es ¿quién modera a los moderadores, guardianes de la información y opiniones en compañías privadas?

También, para evitar bots y campañas automáticas de acoso, desinformación, propaganda o manipulación, podría ser una buena opción, no el requerir la identidad real, pues permitiría desenmascarar seudónimos de creadores, bloggers o activistas y ponerlos en peligro, exponiéndolos a un riesgo grave y real de amenaza física, pero sí que podría suponer un sistema más seguro, una doble autenticación o sistema de identificación para comprobar que la cuenta está manejada por un humano y limitar el número de cuentas con los mismos datos o perfiles. En este sentido se dirige la intención de Musk de crear un Twitter de pago autenticado, Twitter Blue. Sin embargo, hemos visto que no todo es tan sencillo, y que promoviendo un sistema de suscripción de pago, se inundó la plataforma de cuentas fake que sin embargo habían pagado y estaban supuestamente verificadas<sup>74</sup>.

También, la colaboración de las plataformas digitales con las autoridades judiciales, debería ser más rápida y sencilla.

Finalmente, existe un último dilema: ¿tenemos derecho a saber que estamos hablando con un robot? En España, ¿sí!

---

<sup>74</sup> Twitter suspende el pago por verificación tras el caos con cuentas 'fake'. Grupo GDA: 13/11/2022. [Consultado el 15 de noviembre de 2022]. Disponible en: <https://elcomercio.pe/tecnologia/redes-sociales/twitter-suspende-el-pago-por-verificacion-tras-el-caos-con-cuentas-fake-elon-musk-apps-aplicaciones-redes-sociales-espana-mexico-colombia-noticia/>

La Carta de Derechos digitales española establece el derecho a conocer con transparencia si algunas opiniones son moderadas por humanas, automáticamente o con fines publicitarios. Y por lo tanto si estamos debatiendo con un humano o una máquina. Apartado XV:

1. Se promoverá la adopción, por parte de los prestadores de servicios intermediarios a través de los cuales directamente se ejerzan libertades de expresión e información, de protocolos adecuados que garanticen a los usuarios su derecho a recibir información veraz.

2. Tales protocolos asegurarán la transparencia de la información y opiniones al dar a conocer:

a) Si han sido elaboradas sin intervención humana mediante procesos automatizados.

b) Si han sido clasificadas o priorizadas por el proveedor mediante técnicas de perfilado o equivalentes.

c) Si tienen carácter publicitario o hubieran sido patrocinadas por un tercero.

En California han aprobado una reforma<sup>75</sup> que declara ilegal el intento de ocultar la identidad de una persona con un bot con fines comerciales o electorales.

Y sobre el bot, ¿tendría derecho a la libertad de expresión? Ya han sido creadas noticias generadas totalmente de manera automatizada, noticias autogeneradas por inteligencia artificial, o lo que se denomina periodismo-robot. (Govik, 2018)<sup>76</sup>. The Guardian publicó el primer artículo escrito por un bot, y en China hay un presentador de telediario generado por IA.

---

<sup>75</sup> Senate Bill No. 1001 CHAPTER 892. An act to add Chapter 6 (commencing with Section 17940) to Part 3 of Division 7 of the Business and Professions Code, relating to bots. [Approved by Governor September 28, 2018. Filed with Secretary of State September 28, 2018.]

<sup>76</sup> Prof. Dr. Natali Helberger; Sarah Eskens; Max van Drunen; Dr. Mariella Bastian; Dr. Judith Moeller IMPLICATIONS OF AI-DRIVEN TOOLS IN THE MEDIA FOR FREEDOM OF EXPRESSION. Artificial Intelligence – Intelligent Politics Challenges and opportunities for media and democracy Background Paper, Ministerial Conference, Cyprus 2020

¿Sería posible que Siri o Alexa puedan tener libertad de expresión y sean capaces de pensar y responder por sí mismas fuera de ciertos parámetros?

Por último me gustaría acabar con una cita de la Corte Suprema de Justicia de Estados Unidos en el caso *United States v. Alvarez* (2012), «el remedio para el discurso que es falso es el discurso que es verdadero. Este es el curso ordinario en una sociedad democrática. La respuesta para lo irracional es lo racional, para lo desinformado es lo ilustrado, para la mentira flagrante, la simple verdad».

La democracia no es eterna ni está consagrada, hay que protegerla y cuidarla de manera activa. Por ello, la tarea de identificar información falsa y fake news corresponde a todos los ciudadanos con un pensamiento crítico y una educación digital.

#### Bibliografía:

ACCESS NOW Human Rights in the Age of Artificial Intelligence. November 2018.

ALKIVIADOU, Natalie. Automated Content Moderation, Hate Speech and Human Rights. 09 November 2021 [Consultado el 15 de mayo de 2022]. Disponible en: <https://www.humanrightshere.com/post/automated-content-moderation-hate-speech-and-human-rights>

ARTICLE 29-PRIVACY INTERNATIONAL DATA PROTECTION WORKING PARTY. Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. Adopted on 3 October 2017. As last Revised and Adopted on 6 February 2018. 17/EN WP251rev.01 [http://ec.europa.eu/justice/data-protection/index\\_en.htm](http://ec.europa.eu/justice/data-protection/index_en.htm) -ARTICLE 19. Privacy and Freedom of Expression In the Age of Artificial Intelligence. April 2018.

ASWAD, Evelyn, Losing the Freedom to Be Human (February 29, 2020). *Columbia Human Rights Law Review*, Vol. 52, 2020, Available at SSRN: <https://ssrn.com/abstract=3635701>

BALAGUER CALLEJÓN, Francisco. La constitución del algoritmo. El difícil encaje de la constitución analógica en el mundo digital. In: GOMES, Ana Cláudia Nascimento; ALBERGARIA, Bruno; CANOTILHO, Mariana Rodrigues (Coord.). *Direito Constitucional: diálogos em homenagem ao 80º aniversário de J. J. Gomes Canotilho*. Belo Horizonte: Fórum, 2021. ISBN 978-65-5518-191-3.

BALAGUER CALLEJÓN, Francisco. Inteligencia artificial, democracia y derechos. Publicado en: Willis Santiago Guerra Filho,



Lucia Santaella, Dora Kaufmn, Paola. Cantarini, Direito e Inteligência Artificial: Fundamentos - Volume 1, Lumen Juris, Rio de Janeiro, 2021.

BALAGUER CALLEJÓN, Francisco. The Impact of the New Mediators of the Digital Age on Freedom of Speech; YSEC Yearbook of Socio-Economic Constitutions 2021, 2022, Volume 2021; ISBN : 978-3-031-08513-0

BALAGUER CALLEJÓN, Francisco. (2022). The Crisis of the Representative Democracy in the Face of Digital Democracy. In: Blanco de Morais, C., Ferreira Mendes, G., Vesting, T. (eds) The Rule of Law in Cyberspace. Law, Governance and Technology Series, vol 49. Springer, Cham. [https://doi.org/10.1007/978-3-031-07377-9\\_4](https://doi.org/10.1007/978-3-031-07377-9_4)

BALTUSSEN, H. y Davis, P. J. (eds.). (2015). The Art of Veiled Speech. Self-censorship from Aristophanes to Hobbes. Filadelfia: University of Pennsylvania Press.

#BigData: Discrimination in data-supported decision making. HELPING TO MAKE FUNDAMENTAL RIGHTS A REALITY FOR EVERYONE IN THE EUROPEAN UNION. FRA Focus

BOIX PALOP, Andrés. Los algoritmos son reglamentos. La necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones. Revista de Derecho Público: teoría y método, ISSN-e 2695-7191, N°.1 1, 2020, págs. 223-269

BOTERO MARINO, Carolina . La regulación estatal de las llamadas "noticias falsas" desde la perspectiva del derecho a la libertad de expresión. Comisión Interamericana de Derechos Humanos (CIDH), Libertad de expresión: A 30 años de la Opinión Consultiva sobre la colegiación obligatoria de periodistas. Estudios sobre el derecho a la libertad de expresión en la doctrina del Sistema Interamericano de Derechos Humanos, 2 Noviembre 2017, [Consultado el 15 de mayo de 2022] Disponible en: <https://www.refworld.org/es/docid/5a67c0324.html>

BRINK, Hanno. AI, freedom of expression and the press. Synthesis Staff. Synthesis Machine Learning Engineer. 30 de abril 2021. [Consultado el 15 de mayo de 2022]. Disponible en: <https://www.synthesis.co.za/ai-freedom-of-expression-and-the-press/>

BRKAN, Maja, Freedom of Expression and Artificial Intelligence: On Personalisation, Disinformation and (Lack Of) Horizontal Effect of the Charter (March 17, 2019). Available at SSRN: <https://ssrn.com/abstract=33354180> or <http://dx.doi.org/10.2139/ssrn.33354180>

BRONCANO, Fernando (2019). "Encerrados en la burbuja, El laberinto de la identidad, 10 de febrero de 19. [Consultado el 15 de mayo de 2022] Disponible: <<http://laberintodelaidentidad.blogspot.com/2019/02/encerrados-en-la-burbuja.html?m=1>>.

CEBALLOS Noel. Internet Safari. Blackie Books. Enero 2015.

COMMITTEE OF MINISTERS Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries (Adopted by the Committee of Ministers on 7 March 2018 at the 1309th meeting of the Ministers' Deputies)

Confesiones de un bot ruso. Debate. 2022.

Contra el rebaño digital. Un manifiesto, Barcelona, Debate, 2011.

Directores DE LA QUADRA-SALCEDO, Tomás; FERNÁNDEZ DEL CASTILLO Y PIÑAR MAÑAS, José Luis. Coordinadores BARRIO ANDRÉS, Moisés y TORREGROSA VÁZQUEZ, José. Sociedad Digital Y Derecho. Ministerio De Industria, Comercio Y Turismo Red.Es. Boletín Oficial Del Estado. Madrid, 2018

EDWARDS Douglas. I'm Feeling Lucky : Falling On My Feet in Silicon Valley. 28 Junio 2012. Editorial Penguin Books Ltd

EUROPEAN DATA PROTECTION BOARD. Guidelines 8/2020 on the targeting of social media users. Version 2.0. Adopted on 13 April 2021. [https://edpb.europa.eu/system/files/2021-04/edpb\\_guidelines\\_082020\\_on\\_the\\_targeting\\_of\\_social\\_media\\_users\\_en.pdf](https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_082020_on_the_targeting_of_social_media_users_en.pdf)

GALÁN, Carlos. Amenazas híbridas: nuevas herramientas para viejas aspiraciones, Documento de trabajo 20/2018 - 13 de diciembre de 2018. Real Instituto Elcano.

GILLIES Lorna & MANGAN David. The Legal Challenges of Social Media. Leicester University's School of Law in December 2013

GILLESPIE Tarleton. Custodians of the Internet. Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. Paperback. 9780300261431. Tuesday, 24 Aug 2021

GOODMAN Marc, Los delitos del futuro: Todo está conectado, todos somos vulnerables, ¿qué podemos hacer al respecto? (Ariel). 5 noviembre 2015 Gemma Deza Guil (Traductor).

GRASSEGGER, H. & KROGERUS, M: The Data That Turned the World Upside Down How Cambridge Analytica used your Facebook data to help the Donald Trump campaign in the 2016 election. Motherboard. (2017, 28 de Enero) [Consultado el 15 de mayo de 2022] Disponible en: [https://motherboard.vice.com/en\\_us/article/mg9vvn/how-our-likes-helped-trump-win](https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win)

Guía para garantizar la libertad de expresión frente a la desinformación deliberada en contextos electorales / [Elaborada por la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos con los aportes del Departamento de Cooperación y Observación Electoral y el Departamento de Derecho Internacional de la Secretaría General de la Organización de los Estados Americanos]. v.;cm. (OAS. Documentos oficiales ;

OEA/Ser.D/XV.22) [OAS. Documentos oficiales ; OEA/Ser.G CP/CAJP/INF.652/19] ISBN 978-0-8270-6981-7

GRUPO ÁTICO 24. La nueva Ley de Servicios Digitales (DSA) de la Unión Europea. [Consultado el 15 de mayo de 2022] Disponible en: <https://protecciondatos-lopd.com/empresas/ley-de-servicios-digitales-dsa/#:~:text=El%20proyecto%20de%20Ley%20de,servicio%20digital%2C%20por%20peque%C3%B1o%20que>

HAAS, Julia. Global Conference for Media Freedom. November 16 2020. Freedom of the Media and Artificial Intelligence. Office of the OSCE. Representative on Freedom of the Media Freedom of the Media and Artificial Intelligence. <https://www.osce.org/files/f/documents/4/5/472488.pdf>

Prof. Dr. HELBERGER Natali; ESKENS, Sarah; DRUNEN Max van; DR. BASTIAN, Mariella; MOELLER, Dr. Judith. Implications of ai-driven tools in the media for freedom of expression. Artificial Intelligence - Intelligent Politics Challenges and opportunities for media and democracy Background Paper, Ministerial Conference, Cyprus 2020.

HOLMES Aaron. Un ejecutivo de Facebook asegura que lo crearon para ser "tan adictivo como el tabaco" y ahora teme que por ello EEUU pueda estar "al borde de una guerra civil"., Business Insider. 25 sep. 2020 [Consultado el 15 de mayo de 2022] Disponible en: <https://www.businessinsider.es/facebook-diseno-ser-adictivo-como-tabaco-executivo-724243>

HUGHES, C. (2019). It's time to break up Facebook. New York Times Sunday Review, 9-5-2019. Disponible en: <https://nyti.ms/2NUfDsE>.

HUMAN RIGHTS COUNCIL, A/HRC/17/3121 March 2011 Human Rights Council Seventeenth session. Promotion and protection of all human rights, civil, political, economic, social and cultural rights, including the right to development. Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises, John Ruggie

HUMAN RIGHTS COUNCIL. A/HRC/8/5 7 April 2008 HUMAN RIGHTS COUNCIL. Eighth session. Promotion and Protection of all Human Rights, Civil, Political, Economic, Social and Cultural Rights, including the Right to Development Protect, Respect and Remedy: a Framework for Business and Human Rights. Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises, John Ruggie.

KEEN Andrew. Internet no es la respuesta. Enciclopèdia Catalana; N.º 1 edición (18 mayo 2016)

KELLY, Kevin, What Technology Wants, Nueva York, Viking, 2010.

La inteligencia artificial es un gran reto para los derechos humanos en el entorno digital: relator especial de la ONU. 19 de octubre de 2018. <https://r3d.mx/2018/10/19/la-inteligencia-artificial->

es-un-gran-reto-para-los-derechos-humanos-en-el-entorno-digital-relator-especial-de-la-onu/

La tecnocensura, explicada por el activista mexicano Alberto Escorcía. 15 de septiembre de 2016. Youtube. [Consultado el 15 de mayo de 2022] Disponible en: <https://youtu.be/2J6Hz-Uz1fY>

LANIER, Jaron, *You Are Not a Gadget: A Manifesto*, Nueva York, Alfred A. Knopf, 2010. [Hay trad. cast.:

La Limón @velcrotime. El submundo del tuitel: Historias de una red social.1 diciembre 2021. Independently published

LARRONDO, Manuel Ernesto; GRANDI, Nicolas Mario. *Inteligencia Artificial, algoritmos y libertad de expresión*. Universidad Nacional de La Plata. Localización: Universitas-XXI: Revista de Ciencias Sociales y Humanas, ISSN-e 1390-8634, ISSN 1390-3837, N°. 34 (March 2021- August 2021), 2021 (Ejemplar dedicado a: Fake News, Communication and Politics), págs. 177-194. [Consultado el 15 de mayo de 2022] Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=8090679>

LASSALLE, José María. *Ciberleviatán. El colapso de la democracia liberal frente a la revolución digital*. Arpa. 2019.

LESSIG, Lawrence, *Code: And Other Laws of Cyberspace, Version 2.0*, Nueva York, Basic Books, 2006. [Hay trad. cast.: *El código y otras leyes del ciberespacio*, Barcelona, Taurus, 2001].

LESLIE, D., BURR, C., AITKEN, M., COWLS, J., KATELL, M., AND BRIGGS, M. (2021). *Artificial intelligence, human rights, democracy, and the rule of law: a primer*. The Council of Europe. FINCK, Michèle. *Artificial intelligence and online hate speech*. Enero de 2019. © 2019, Centre on Regulation in Europe (CERRE)

LIN, BILL & ZHOU, WANGCHUNSHU & SHEN, MING & ZHOU, PEI & BHAGAVATULA, CHANDRA & YEJIN, CHOI & REN, XIANG. (2020). *CommonGen: A Constrained Text Generation Challenge for Generative Commonsense Reasoning*. 1823-1840. [Consultado el 15 de mayo de 2022] Disponible en: [10.18653/v1/2020.findings-emnlp.165](https://arxiv.org/pdf/1911.03705.pdf)  
<https://arxiv.org/pdf/1911.03705.pdf>

LOSSE Katherine. *The boy kings: A Journey Into the Heart of the Social Network*. 18 octubre 2014. Free Press; 256 páginas. ISBN-10 : 1451668260

MARTÍN CRITIKIÁN, D. Y MEDINA NÚÑEZ, M. (2021). *Redes sociales y la adicción al like de la generación z*. *Revista de Comunicación y Salud*, 11, 55-76 [Consultado el 15 de mayo de 2022] Disponible en: <https://doi.org/10.35669/rcys.2021.11.e281>

MARTÍNEZ MARTÍNEZ, Ricard. (2019). *Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo*. *Revista Catalana de Dret Públic*, (58). 64-81.  
<https://doi.org/10.2436/rcdp.i58.2019.3317>

MERJÍA, A. F. (2020). La libertad de expresión en jaque, el panóptico del Siglo XXI. Big Data como amenaza para la democracia: a propósito del caso Cambridge Analytica. UNIVERSITAS. Revista De Filosofía, Derecho Y Política, (32), 79-105. <https://doi.org/10.20318/universitas.2020.5512>

MERINO RUS, Rafael. Los derechos humanos, la democracia y la igualdad en la era de los algoritmos y la inteligencia artificial El contenido de esta publicación corresponde a una ponencia presentada durante el IV Encuentro de Verano de Gernika-Urdaibai «Desdemocracia, los "ismos" y el abismo. Universales ante las políticas segregatorias», en el marco de los XXXVIII Cursos de Verano UPV/EHU, celebrado en Gernika-Lumo los días 19 y 20 de junio de 2019.

MIJATOVIĆ, Dunja. HUMAN RIGHTS COMMENT-Safeguarding human rights in the era of artificial intelligence. STRASBOURG 03 de julio de 2018. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.coe.int/en/web/commissioner/-/safeguarding-human-rights-in-the-era-of-artificial-intelligence>

NOAH HARARI, Yuval .21 lecciones para el siglo XXI. Debate. Barcelona. 2018.

NOTT, Lata. Does the first amendment protect speech made by artificial intelligence? Posted on 19 de septiembre 2019. Freedom Forum Institute. [Consultado el 15 de mayo de 2022]. Disponible en: <https://www.freedomforuminstitute.org/2019/09/19/does-the-first-amendment-protect-speech-made-by-artificial-intelligence/>

O'NEIL, C. (2018). Armas de destrucción matemática. Madrid: Capitán Swing.

PARISER, Eli. El Filtro Burbuja: Cómo la web Decide lo que Leemos y lo que Pensamos. Taurus, 2017.

PHAM, THANG & BUI, TRUNG & MAI, LONG & NGUYEN, ANH. (2021). Out of Order: How important is the sequential order of words in a sentence in Natural Language Understanding tasks?. 1145-1160. [Consultado el 15 de mayo de 2022] Disponible en: 10.18653/v1/2021.findings-acl.98. <https://arxiv.org/pdf/2012.15180.pdf>

PEIRANO, Marta y PÉREZ COLOMÉ, Jordi. La nueva censura digital. del algoritmo a la inteligencia artificial. Conversación- José María Lassalle • Revista MINERVA, REVISTA DEL CÍRCULO DE BELLAS ARTES [Consultado el 15 de mayo de 2022] Disponible en: <https://cbamadrid.es/revistaminerva/articulo.php?id=888>

PEIRANO, Marta. El enemigo conoce el sistema: manipulación de ideas, personas e influencias después de la economía de la atención. Debate, 2019. ISBN 978-84-17636-39-5

PERALTA GUTIÉRREZ, ALFONSO. Comunicación y principios éticos para Jueces y Magistrados en redes sociales (1) Diario La Ley, N° 9651, Sección Tribuna, 11 de Junio de 2020, Wolters Kluwer. LA LEY 5739/2020

PERALTA GUTIÉRREZ, Alfonso. La verdad no justifica la censura. El juicio comparado de derechos en materia de desinformación, seguridad nacional y libertad de expresión (1). Diario La Ley, ISSN 1989-6913, N° 9772, 2021. [Consultado el 15 de mayo de 2022] Disponible en: <https://diariolaley.laleynext.es/d11/2021/01/18/la-verdad-no-justifica-la-censura-el-juicio-comparado-de-derechos-en-materia-de-desinformacion-seguridad-nacional-y-libertad-de-expresion>

POLICY PAPER on Artificial Intelligence's (AI) Impact on Freedom of Expression in Political Campaign and Elections April 2021. Armin Rabitsch\*, Rania Wazir#, and Thomas Treml# \*Election-Watch.EU #data4good Initiative of the Vienna Data Science Group (VDSG)

Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. A/HRC/38/35. 6 de abril de 2018.

REPORTERS WITHOUT BORDERS. Online Harassment Of Journalists. Attack of the trolls.

RODRÍGUEZ MEDINA, José Luis Y ARISTA BUSTAMANTE, Luis Leoncio. Los algoritmos y la transgresión del derecho a la libertad de expresión y libre albedrío. Rev.SSIAS. Vol. 14 / No2 (2021). Fecha de recepción: 05 de octubre de 2021. Fecha de aprobación: 17 de noviembre de 2021. DOI: 10.26495/rcs.v14i2.1949

RONSON Jon. Humillación en las redes. B DE BOOKS, Noviembre 2015

RUDAS MURGA, C. (2021). Redes sociales: Inteligencia artificial en el derecho al honor desde una perspectiva peruana. Lucerna Iuris Et Investigatio, (1), 99-110. <https://doi.org/10.15381/lucerna.v0i1.20137>

SÁNCHEZ BARRILAO, Juan Francisco "El derecho constitucional ante la era de ultrón: la informática y la inteligencia artificial como objeto constitucional", Estudios de Deusto, Vol. 64/2, Julio-Diciembre 2016, p. 239.

SANDER, Barrie, Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation (May 23, 2020). Fordham International Law Journal, Vol. 43, No. 4, 2020, [Consultado el 15 de mayo de 2022] Disponible en: <https://ssrn.com/abstract=3434972> or <http://dx.doi.org/10.2139/ssrn.3434972>

SHIELDS, David, Reality Hunger: A Manifesto, Nueva York, Alfred A. Knopf, 2010. [Hay trad cast.: Hambre de realidad. Un manifiesto, Madrid, Círculo de tiza, 2015].

SOTO IVARS, J. (2017). Arden las redes. Barcelona: Debate.

SOTO IVARS, J. (2018). Manual para inquisidores novatos. El Món d'ahir, 8, 178-183. Disponible en: <https://botiga.som.cat/el-mon-d-ahir-8-968.1001>

STOA/ESMH Options Brief. EPRS | European Parliamentary Research Service. Scientific Foresight Unit (STOA) PE 624.279 - March 2019 EN. Regulating disinformation with artificial intelligence

SUNSTEIN, C. (2007). Republic.com 2.0. Princeton: Princeton University Press.

SUNSTEIN, C. (2017). #Republic: Divided Democracy in the Age of Social Media. Nueva York: Princeton University Press.

The Filter Bubble By The Daily Dish. 10 de octubre 2010. The Atlantic. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.theatlantic.com/daily-dish/archive/2010/10/the-filter-bubble/181427/>

THOMAS Justice: Judges aren't 'mass media icons,' should uphold the law. By Bill Rankin, The Atlanta Journal-Constitution. [Consultado el 15 de mayo de 2022] Disponible en: <https://www.ajc.com/news/local/justice-thomas-judges-aren-mass-media-icons-should-uphold-the-law/s5PyqmJWd7IOQK7XR2PsWI/>

TURNER, Fred, From Counterculture to Cyberculture. Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism, Chicago, The University of Chicago Press, 2006.

UNESCO. Recommendation on the ethics of artificial intelligence. November 2021.

UNESCO and The future Society and National Judicial College MOOC. Upholding the Rule of Law in the Age of Artificial Intelligence. [https://www.judges.org/ai\\_and\\_law/english/](https://www.judges.org/ai_and_law/english/)

UNESCO WORLD PRESS FREEDOM DAY- JOURNALISM UNDER DIGITAL SIEGE. Global Conference. 2 de mayo de 2022. Social Media 4 Peace: Content Moderation Practices to Advance Social Cohesion by UNESCO with the support of the European Union • Frances Haugen, Advocate for accountability and transparency in social media • João Brandt, Expert, Internet Policies, Instituto Cultura e Democracia • Quinn McKew, Executive Director, ARTICLE 19 • Gabriel Parra, Public policy Manager Asociación Latinoamericana de Internet (ALAI) MODERATORS: Adeline Hulin, Project Coordinator Social Media 4 Peace; and John Okande Otieno, UNESCO Project Officer in Nairobi, Social Media 4 Peace [Consultado el 15 de mayo de 2022] Disponible en: <https://vimeo.com/706253668/2154e0c159>

UNESCO WORLD PRESS FREEDOM DAY- JOURNALISM UNDER DIGITAL SIEGE. Global Conference. 2 de mayo de 2022. I've got my AI On You: Spotlight on Artificial Intelligence and Freedom of Expression by the Organization for Security and Cooperation in Europe (OSCE) • Teresa Ribeiro, Representative on Freedom of the Media, Organization for Cooperation and Security in Europe (OSCE) • Vladan Joler, Academic, Researcher, Founder of SHARE Foundation • Susie Alegre, Human Rights Lawyer, Writer • Arzu Geybullayeva, Journalist • Andrei Soldatov, Investigative Journalist, Agentura.ru MODERATOR: Deniz Wagner, Adviser, Office of the OSCE Representative on Freedom of the Media

[Consultado el 15 de mayo de 2022] Disponible en:  
<https://vimeo.com/706251784/d733d80718>

UNESCO WORLD PRESS FREEDOM DAY- JOURNALISM UNDER DIGITAL SIEGE. Global Conference. 2 de mayo de 2022. SILENCED - Democracy, Journalism and Censorship in the Digital Age by IBAHRI and the High Level Panel of Legal Experts on Media Freedom • Judge Chile Eboe-Osuji, Former President, International Criminal Court, Member of the High Level Panel of Legal Experts on Media Freedom • Dario Milo, Partner at Webber Wentzel, Attorneys and Member of the High Level Panel of Legal Experts on Media Freedom • Catherine Amirfar, International Lawyer, Co-Chair of International Dispute Resolution Group and the Public International Law Group at Debevoise & Plimpton; President of the American Society of International Law and Deputy Chair of the High Level Panel of Legal Experts on Media Freedom MODERATOR: Pedro Vaca Villareal, Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights [Consultado el 15 de mayo de 2022] Disponible en: <https://vimeo.com/708658967/c3fe10fe85>

UNESCO WORLD PRESS FREEDOM DAY- JOURNALISM UNDER DIGITAL SIEGE. Global Conference. 2 de mayo de 2022. The Role of International Community Media in Protecting the Public from Disinformation and Surveillance by PEN America • Nataliya Gumenyuk, author, documentary filmmaker, and journalist, Public Interest JournalismLab, Ukraine • Mukhtar Ibrahim, Executive Director and Editor of Sahan Journal • Dina Meza, Founder of PEN Honduras, Human Rights Defender, Journalist MODERATOR: Kavitha Rajagopalan, Community Engagement Manager at CUNY Newmark J-School's Center for Community Media. [Consultado el 15 de mayo de 2022] Disponible en: <https://vimeo.com/707468584/18b810e34b>

UNESCO WORLD PRESS FREEDOM DAY- JOURNALISM UNDER DIGITAL SIEGE. Global Conference. 2 de mayo de 2022. Addressing the Digital Challenges to Freedom of Expression through the UNESCO Recommendation on the Ethics of AI by UNESCO • SETTING THE STAGE ON THE AI ETHICS RECOMMENDATION: Gabriela Ramos, Assistant Director-General for the Social and Human Sciences of UNESCO • Rumman Chowdhury, Director for the Machine Learning Ethics, Transparency and Accountability, Twitter • Diego García-Sayán, Special Rapporteur on the Independence of Judges and Lawyers, OHCHR • Alfonso Peralta Gutiérrez, Judge of the First Instance and Criminal Investigation Court n. 1 of Roquetas de Mar/Spain, Specialised Court in Criminal International Cooperation • Fabrizio Scrollini, Executive Director, Open Data Latin American Initiative (ILDA) • Rodica Ciochina, Programme Officer, Media and Internet Governance Unit of the Information Society Department of the Directorate General for Human Rights and Rule of Law, Council of Europe. MODERATOR: Natalia Zuazo, Technology Journalist. [Consultado el 15 de mayo de 2022] Disponible en: <https://vimeo.com/706890647/582a017c8e>

UNESCO WORLD PRESS FREEDOM DAY- JOURNALISM UNDER DIGITAL SIEGE. Global Conference. 2 de mayo de 2022. Transparency and Beyond: Role of Think Tanks in Advancing Implementation of the Responsibilities of Intermediaries by the Center for Law and Democracy • Grace Githaiga, Convenor at the multi-stakeholder ICT Policy and Regulation advocacy



think tank, KICTANet, Kenya • Akriti Gaur, Resident Fellow at the Yale Information Society Project, India • Andrew Puddephatt, Director, Global Partners & Associates MODERATOR: Toby Mendel, Executive Director, Centre for Law and Democracy, Canada [Consultado el 15 de mayo de 2022] Disponible en: <https://vimeo.com/707792016/ab89ad0efb>

UNESCO WORLD PRESS FREEDOM DAY- JOURNALISM UNDER DIGITAL SIEGE. Global Conference. 2 de mayo de 2022. Latest Jurisprudence of the Inter-American Court of Human Rights on Freedom of Expression, Freedom of the Press and Safety of Journalists by the Columbia Global Freedom of Expression, the UNESCO Chair on Freedom of Expression at the Los Andes University, OBSERVACOM and UNESCO • Ernesto Fernandez Polcuch, Director of UNESCO's Regional Bureau for Sciences in Latin America and the Caribbean • Edison Lanza, Former Special Rapporteur for Freedom of Expression of the Interamerican Court for Human Rights • Jorge Jiménez Martín, Secretary General of the Ibero-American Network of Judicial Schools (RIA EJ) • Jonathan Bock, Executive Director, Fundación para la Libertad de Prensa - (FLIP) • Amy B. Van Zyl-Chavarró, Suffolk University Law School • Carlos Ayala, Vice President and Member of the International Commission of Jurists • Zuliana Lainez, Vice President, International Federation of Journalists (IFJ). [Consultado el 15 de mayo de 2022] Disponible en: <https://vimeo.com/707824377/5790bac7e0>

VAIDHYANATHAN, Siva. La googlización de todo (y por qué deberíamos preocuparnos). Editorial Océano, 2012 - 238 páginas. p. 182.

WATTS, Duncan J., Six Degrees: The Science of a Connected Age, Nueva York, W.W. Norton & Company, 2003. [Hay trad. cast.: Seis grados de separación. La ciencia de las redes en la era del acceso, Barcelona, Paidós, 2006].

WOLFGANG HOFFMANN-RIEM: Big Data. Desafíos también para el derecho (trad. Eduardo Knörr Argote). Cizur Menor (Navarra), Cuadernos Civitas, 2018, 179 págs.

WOLFGANG HOFFMANN-RIEM: el derecho ante el reto de la transformación digital. Aranzadi ISBN: 978-84-1390-703-1 Páginas: 384. 2022.

WU, Tim, The Master Switch. The Rise and Fall of Information Empires, Nueva York, Alfred A. Knopf, 2010.

ZITTRAIN, Jonathan, The Future of the Internet-And How to Stop It, New Haven, Yale University Press, 2008.

ZUAZO, Natalia. Conferencia TAI Granada 21 de octubre de 2022. [Consultado el 29 de octubre de 2022] Disponible en: <https://www.ideal.es/tatgranada/natalia-zuazo-20221021114702-nt.html>

ZUBOFF, Shoshana. "The Coup We Are Not Talking About", The New York Times, 29/1/2021, p. 4: [Consultado el 15 de mayo de 2022] Disponible en: <https://www.nytimes.com/2021/01/29/opinion/sunday/facebook-surveillance-societytechnology>