

Impactos  
del uso  
de las nuevas  
tecnologías  
digitales en la  
libertad  
de expresión

El documento comienza con una definición de las tecnologías digitales y la inteligencia artificial para continuar con la delimitación del derecho a la libertad de opinión y de expresión, tal y como está recogido en el Sistema Universal de Derechos Humanos. Por último lleva a cabo un análisis del estado de la cuestión de algunos de los impactos y retos actuales que el uso de las tecnologías digitales, con especial atención a la inteligencia artificial, supone para la libertad de expresión.

Autoría: June Orenge

Fecha: Marzo 2022

Edición:

Institut de Drets Humans de Catalunya

Av. Meridiana 32, entr. 2a. Esc B

08018 Barcelona

[www.idhc.org](http://www.idhc.org)

Diseño y maquetación: [nadianmartin.com](http://nadianmartin.com)



Esta obra está bajo una licencia de Creative Commons Reconocimiento-No-Comercial 4.0 Internacional. Se puede copiar, distribuir, comunicar públicamente, traducir y modificar, siempre que sea para fines no comerciales y se reconozca su autoría.



El contenido de esta publicación es responsabilidad exclusiva del Institut de Drets Humans de Catalunya y no refleja necesariamente la opinión de la Open Society Initiative for Europe.

# Índice

---

I.	Introducción	4
II.	Tecnología digital e inteligencia artificial	7
	2.1. ¿Qué entendemos por tecnología digital?	8
	2.2. ¿Qué entendemos por inteligencia artificial?	9
III.	Delimitación del derecho a la libertad de opinión y de expresión en el Sistema Universal de Derechos Humanos	14
	3.1. Qué es y no es la libertad de expresión para el SUDH	16
	3.2. Caracterización del derecho de acceso a la información	17
	3.3. Límites y restricciones a la libertad de expresión	18
	3.4. Otros derechos vinculados a la libertad de expresión	19
IV.	Vínculos, problemas y retos en relación al uso de tecnologías digitales para la libertad de expresión	22
	4.1. El perfilado y la presentación del contenido	24
	4.2. Moderación del contenido: bloqueo y eliminación automática	26
	4.3. Desinformación y contaminación informativa	27
	4.4. Vigilancia masiva y selectiva	31
	4.5. Discursos de odio <i>online</i>	32
	4.6. El riesgo de no ser olvidada	34
V.	Conclusiones	36

cap. 1

# Introducción

Las tecnologías digitales moldean cada vez más la forma en que las personas ejercen la libertad de expresión, en cómo acceden a la información, en cómo interactúan de manera *online*, opinan o difunden sus opiniones u otras. Es indudable que, quien pueda acceder y utilizar las tecnologías digitales, podrá disfrutar de ciertas ventajas, pero también se enfrentará con determinados retos y desafíos. El gran impacto y alcance del uso de las tecnologías digitales en los derechos humanos en general y en la libertad de expresión en particular, demandan examinar detenidamente y de manera regular en el tiempo, las formas en que estas tecnologías se están aplicando ahora, la forma en que avanza esa aplicación, cómo esta va estructurando nuestra vida y, en la medida de lo posible, a quién o quiénes beneficia ese uso que finalmente se les da.

Este documento pretende llevar a cabo un análisis básico, un recorrido general sobre algunos de estos impactos y sus consecuencias para la libertad de expresión. Es un tema complejo porque abarca multiplicidad de factores y porque además está en constante evolución: la rapidez con que van surgiendo, mejorando o cambiando las tecnologías digitales dificulta una identificación y posterior análisis actualizado y completo de estos vínculos e impactos. Lo que se recoge aquí es, por tanto, una primera aproximación a través de los principales problemas o retos identificados y que están en debate a fecha de hoy (marzo de 2022). No pretende ser, en este sentido, un documento cerrado; está y estará abierto a nuevos análisis, contribuciones y propuestas, porque lo idóneo sería que estos impactos se analicen en cada momento y contexto determinado.

La dificultad de abarcar todas las tecnologías digitales, todos los impactos, todos los problemas o retos e incluso todos los factores de cada uno de esos retos genera que podamos centrarnos solo en algunos de ellos. Existen cuestiones importantes que no aborda específicamente este documento, como la gestión de grandes cantidades de datos, que en ocasiones son recolectados a través de dudosos métodos, o la afectación concreta en el pluralismo, la diversidad y libertad de los medios de comunicación. Son ambos temas importantes que impactan en la libertad de expresión y en otros derechos, pero que están siendo tratados y discutidos intensamente en otros espacios.

Otra dificultad añadida del análisis es que el límite o la frontera entre lo bueno y malo, correcto e incorrecto es impreciso cuando tratamos los vínculos entre las tecnologías digitales y la libertad de expresión. Veremos cómo concretamente algunas tecnologías, técnicas o herramientas pueden ser perjudiciales para este derecho en un momento dado y en otro, beneficiosas o provechosas para contrarrestar o mitigar algunos efectos adversos. El refrán “*dos lados de una misma moneda*” cobra todo el sentido cuando hablamos del uso de este tipo de tecnologías.

El texto a continuación se estructura, principalmente, como sigue:

Un primer apartado donde se delimitan conceptos, así como características generales de las tecnologías digitales y, dentro de ellas, de la inteligencia artificial, como protagonista, dado su impacto de gran alcance en la libertad de expresión.

Un segundo apartado centrado en la delimitación del derecho a la libertad de opinión y de expresión en el Sistema Universal de Derechos Humanos, que nos permita clarificar qué comprende el derecho y qué no desde este marco normativo. Se hará también una breve referencia a otros derechos muy relacionados con la libertad de expresión que pueden verse afectados por el uso de las tecnologías digitales, así como mención general a algunas de esas afectaciones.

Por último, un tercer apartado donde se lleva a cabo un análisis del estado de la cuestión de algunos de los vínculos y consecuentes retos en relación al uso de tecnologías digitales, con especial atención a la inteligencia artificial, para la libertad de expresión.

cap. **2**

Tecnología  
digital e  
inteligencia  
artificial

## 2.1. ¿Qué entendemos por tecnología digital?

La palabra “tecnología” proviene del griego (*téchnē*) y significa arte, oficio o destreza.<sup>1</sup> Por lo tanto, desde esta definición, la tecnología no sería tanto un objeto, sino un proceso, una capacidad de transformar o combinar algo ya existente para construir algo nuevo o bien darle, a algo ya conocido, otra función. La tecnológica, que comprende por tanto un gran número de cuestiones,<sup>2</sup> surge de las necesidades, de la curiosidad o inquietud por modificar algo de nuestro entorno para que genere un mayor bienestar o una mejora en nuestras condiciones de vida. Las tecnologías no pretenden describir o entender el mundo, como sí lo hace la ciencia, sino que su objetivo es modificar, moldear ese mundo o ciertas características de él para adaptarlo a las necesidades humanas.<sup>3</sup>

La palabra “digital”, por su parte, viene del latín y significa *digitus*<sup>4</sup> o dedo, en referencia a una de las herramientas más antiguas para contar. Actualmente se asocia básicamente con tecnología y hace referencia a que cuando la información se almacena, transmite o reenvía en formato digital, se convierte en números (en el nivel más básico de ceros y unos), medibles y alterables de forma fácil gracias al cálculo matemático. La tecnología asociada a “lo digital” pasaría de ser así algo mecánico o análogo, a algo computarizado. Podemos ver la diferencia en cómo ha evolucionado la forma de escuchar música por ejemplo: mientras que antes utilizábamos dispositivos que reproducían la música de forma mecánica como los radiocasetes o *walkmans*, ahora podemos almacenar en el móvil miles de canciones que la reproducen de forma digital.

8

1 Definición extraída de la Real Academia Española, Diccionario de la lengua española: [www.rae.es](http://www.rae.es), [consultado el 15 de marzo de 2022].

2 Para más detalle consultar: Pont A., Passera M. A. y Castilla K., *Impactos de las nuevas tecnologías en los derechos humanos*, Institut de Drets Humans de Catalunya, 2022, disponible en: <https://www.idhc.org/es/investigacion/publicaciones/otras-publicaciones/impactos-de-las-nuevas-tecnologias-en-los-derechos-humanos.php>

3 Kapur R., *Significance of digital technology*, *International Journal of Transformations in Business Management*, Vol. No 8, 2018, p. 1-3, disponible en: [https://www.researchgate.net/publication/323829721\\_Significance\\_of\\_Digital\\_Technology](https://www.researchgate.net/publication/323829721_Significance_of_Digital_Technology)

4 Véase nota *supra* 1.

La tecnología digital comprende una gran cantidad de herramientas, sistemas, dispositivos y recursos que generan, almacenan o procesan datos. Actualmente, es utilizada para funciones como el procesamiento de grandes volúmenes de datos, la automatización de procesos, las comunicaciones masivas, la creación de sistemas que imitan el funcionamiento de la mente humana, los cómputos sumamente potentes, entre muchas otras. Algunos ejemplos concretos y conocidos de tecnología digital serían las mismas webs o la propia internet; teléfonos inteligentes o *smart phones*; video *streaming*, que vemos por ejemplo en famosas aplicaciones que nos permiten ver series o películas o hacer videoconferencias; *e-books* o libros electrónicos; electrodomésticos inteligentes como cafeteras, neveras, aspiradores robotizados, cocinas inteligentes, etc.; *Blockchain* o cadenas de bloques, que es la tecnología utilizada para almacenar datos de forma segura, usando la criptografía.

El uso de la tecnología digital conlleva innumerables ventajas<sup>5</sup> como el progreso del sector sanitario y el aumento en la esperanza y nivel de vida, nuevas oportunidades educativas y acceso a ellas por parte de un mayor número de personas (aunque sigue restringido a otras), la accesibilidad y el gran número de información disponible en tiempo real, la capacidad de comunicarnos con seres queridos en largas distancias antes inimaginables, en la diagnosis de la situación del medio ambiente, que nos permite predecir desastres o analizar el estado de la agricultura, e incluso en la defensa de los derechos humanos. Pero es igualmente innegable que la tecnología digital, o más bien el uso que se le da, puede suponer vulneraciones de derechos humanos, tal y como se analiza en más detalle en otra publicación<sup>6</sup> del Institut de Drets Humans de Catalunya (IDHC).

Dentro de las tecnologías digitales, la inteligencia artificial (IA) impacta de manera destacada en los derechos humanos y concretamente en la libertad de expresión, no solo porque la IA es la que permite a las personas buscar, acceder, recibir y difundir la información, sino porque se utiliza cada vez más en los procesos de toma de decisiones en relación a esos procesos (y otros), que antes eran responsabilidad exclusiva de los seres humanos. Es por esto que nos detendremos brevemente en la IA a continuación, pero sin dejar de lado otras tecnologías digitales que, sin la implicación de la IA, pueden ser igualmente determinantes para ejercer la libertad de expresión. Pensemos por ejemplo en el acceso a internet, o a una computadora, *tablet*, *smart phone*, entre otras que veremos a los largo de estas páginas.

9

## 2.2. ¿Qué entendemos por inteligencia artificial?

No existe una única definición de la IA porque, entre otras cosas, no hace referencia a una única cosa y también porque sufre cambios constantemente. Comprende una gran variedad de aplicaciones, técnicas, de diferente nivel y complejidad, autonomía y abstracción que permiten que las computa-

---

5 United Nations News, The impact of Digital Technologies, United Nations, 2020, disponible en: <https://www.un.org/en/un75/impact-digital-technologies>

6 Véase nota *supra* 2, p. 5-19.

doras complementen o reemplacen tareas específicas que, de otro modo, serían ejecutadas por seres humanos. Básicamente tomar decisiones y resolver problemas.

Fue John McCarthy, quien acuñó el concepto de IA en la Conferencia de Dartmouth de 1956, definiéndola como "la ciencia y la ingeniería de crear máquinas inteligentes, especialmente programas de computación inteligente. Está relacionada con la tarea similar de utilizar ordenadores para comprender la inteligencia humana, pero la IA no se limita a métodos que sean observables biológicamente".<sup>7</sup>

David Kaye, el que fue Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión (en adelante RE sobre libertad de expresión), de la Organización de las Naciones Unidas (ONU) por su parte, reconociendo la influencia creciente de la IA en el entorno de la información a nivel mundial, la definía en 2018 como "una constelación de procesos y tecnologías, en cuya base están los algoritmos, que permiten que las computadoras complementen o reemplacen tareas específicas que de otro modo serían ejecutadas por seres humanos, como tomar decisiones y resolver problemas".<sup>8</sup>

En el imaginario popular es común que la IA se relacione con esos sistemas que se aproximan mucho a la inteligencia humana o incluso la supera en muchos ámbitos, en forma de robots muchas veces, y que funciona según sus propios conceptos, criterios y reglas. Pero esta, llamada IA fuerte o general, no existe. La IA generalmente "optimiza la ejecución de tareas computarizadas y asignadas o dirigidas por seres humanos mediante repetición e intentos iterativos"<sup>9</sup>. Este es el tipo de IA que conocemos actualmente, denominada limitada o débil (Artificial Narrow Intelligence – ANI).<sup>10</sup>

10

Aunque nació hace décadas, el gran impacto de la IA comenzó sobre todo cuando potentes ordenadores capaces de experimentar con ella de forma global y cotidiana, se hicieron accesibles al gran público. Surgieron los agentes inteligentes, capaces de dar respuesta a partir de unos datos y unas reglas, o los *chatbots* que son capaces de mantener una conversación con personas. La gran diferencia de la IA respecto a un programa de ordenador, es que este último es solo una lista de órdenes que un computador debe hacer. La IA no recibe órdenes continuamente, sino que, sobre una base (reglas, criterios, etc.) que se le incorpora en su diseño, decide qué hace, qué respuesta dar ante determinados problemas. Puede incluso aprender del proceso para mejorar esas respuestas, como veremos a continuación.

La IA se agrega a productos ya existentes, sean *hardware* o *software*, y es importante algún subconjunto o técnica de este tipo de inteligencia que le permitirá obtener un valor agregado. Es así como

---

7 Computer Science Department, *What is Artificial Intelligence? Interview to John MacCarthy*, Stanford University, review revised in November 2004, p. 2, disponible en: <http://www-formal.stanford.edu/jmc/whatisai.pdf>

8 Informe A/73/348 del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión de la ONU, remitida por la Asamblea General en agosto de 2018, pág. 3, párr. 3.

9 En matemáticas se denomina iteración a la repetición de un proceso para generar una secuencia de resultados (posiblemente ilimitada). La secuencia se acercará a algún punto final o valor final. Cada repetición del proceso es una única iteración, y el resultado de cada iteración es el punto de partida de la siguiente iteración.

10 Véase nota *supra* 8, p. 4, párr. 3.

las automatizaciones, *bots*<sup>11</sup> y sistemas inteligentes se combinan junto a un gran volumen de datos para mejorar tecnologías de cualquier ámbito (hogar, ocio, trabajo, etc.) y tipo (dispositivos domésticos, ordenadores, móviles, etc.). La IA se aplica en prácticamente todos los sectores: industrial, económico, militar, académico, sanitario, legal, transportes y muchos más.<sup>12</sup>

En este punto es importante subrayar, y lo veremos a lo largo de todo el documento, que detrás de un mecanismo de IA está la mente y mano humana, que es quien diseña los programas, las reglas o los criterios sobre los que se basará ese producto final de IA. Y esto es clave, porque la persona o personas que están detrás de ese diseño y puesta en marcha de la IA, con sus emociones, conocimientos, estereotipos y prejuicios, trasladará sus sesgos, y por lo tanto ese producto o proceso final no será neutral, con todo lo que eso conlleva.

Lo mismo pasa con los algoritmos que están en la base de la IA y que son códigos informáticos, diseñados y escritos por seres humanos que son quienes dan las instrucciones y traducen los datos en conclusiones, información legible o productos concretos.<sup>13</sup> Dicho de forma sencilla: un algoritmo es un conjunto ordenado y finito de operaciones simples a través de las cuales podemos hallar la solución a un problema, simulando la inteligencia humana. Permiten ejecutar una acción o resolver un problema mediante una serie de instrucciones definidas, ordenadas y limitadas.

El gran desarrollo y despliegue de los algoritmos se ha visto incrementado por el gran volumen de datos que se manejan hoy en día; la IA es capaz de analizarlos rápidamente y tomar decisiones sobre ellos, facilitando mucho el trabajo al sector privado que gestiona los datos. Cuanta más disponibilidad de datos tengan, más precisos y fuertes pueden ser los algoritmos. Sin detenernos en este punto, señalar que esta “necesidad” de datos ha conllevado ya de entrada ciertos problemas relacionados con la privacidad de las personas, que afectan también a la libertad de expresión.<sup>14</sup>

En otra publicación<sup>15</sup> del IDHC, centrada en el análisis de la igualdad y no discriminación en la IA, se explican algunas características generales de los algoritmos, la importancia de los datos, así como la

11 *Bot* es la abreviatura de robot. Los Bots son programas que realizan tareas repetitivas en internet, explorando la web de forma recurrente. Tienen diferentes propósitos, como indexar páginas web a motores de búsqueda o cosechar direcciones de correo electrónico para los *spammers*. Ver en <https://www.arimetrics.com/glosario-digital/bot>

12 Ver breve guía en Wikipedia, *Applications of artificial intelligence*, disponible en: [https://en.wikipedia.org/wiki/Applications\\_of\\_artificial\\_intelligence](https://en.wikipedia.org/wiki/Applications_of_artificial_intelligence)

13 Larrondo M. E., Grandi N. M., *Inteligencia Artificial, algoritmos y libertad de expresión*, Universidad Nacional de La Plata, 2021, p.180, disponible en: <https://www.redalyc.org/journal/4761/476165932008/html/>

14 Ver por ejemplo Sanz M., *Microsoft se suma a la lista de empresas que espían nuestras conversaciones*, Computer Hoy, 2019, disponible en: <https://computerhoy.com/noticias/tecnologia/microsoft-grabaciones-%20escucha-conversaciones-474759> o Álvarez L., *Espías en casa: del dispositivo móvil al robot aspirador*, El Periódico 2019, disponible en: <https://www.elperiodico.com/es/activos/innovadores/20190716/tecnologia-hogar-espia-7554028> o Bohm D., *Google's 'field research' offered people \$5 to scan their faces for the Pixel 4*, 2019, disponible en: <https://www.theverge.com/2019/7/29/8934804/google-pixel-4-face-scanning-data-collection>

15 Castilla K., *Cuatro ángulos de análisis de la igualdad y la no discriminación en la inteligencia artificial*, Institut de Drets Humans de Catalunya (IDHC), 2021, disponible en: <https://www.idhc.org/es/investigacion/publicaciones/otras-publicaciones/cuatro-angulos-de-analisis-de-la-igualdad-y-la-no-discriminacion-en-la-inteligencia-artificial.php>

tipología general, por lo que no nos detendremos en este punto. Simplemente cabe recordar las dos principales técnicas de AI existentes: *Machine Learning* y *Deep Learning*.

El *Machine Learning* (ML) es una de técnica, dentro de la IA, que se utiliza para crear y mejorar el comportamiento de las computadoras para que se asemejen a la inteligencia humana. Una cosa es configurar un sistema para que realice una tarea determinada (programa de ordenador tradicional), y otra es configurar ese sistema para que aprenda a realizar por su cuenta esa misma tarea (IA) y además mejore esa ejecución (ML). Por poner un ejemplo sencillo, la IA sería la capacidad de jugar al ajedrez y el ML sería la técnica de entrenamiento para jugar al ajedrez.

El *Deep Learning* (DL), que se considera un algoritmo de aprendizaje profundo y automático, es la técnica de ML que utiliza redes neuronales y la gran diferencia con esta estriba en cómo aprende cada algoritmo: el DL automatiza grandes fragmentos del proceso de procesamiento de características, lo que elimina parte de la intervención humana manual requerida y permite el uso y procesamiento de conjuntos de datos más grandes.

El aprendizaje que se realiza a través del ML puede ser supervisado (*Supervised Machine Learning*) o no (*Unsupervised Machine Learning*). En el primer caso, está detrás el seguimiento de una persona. En el segundo no tanto porque es la propia IA la que debe descubrir qué hace bien y qué no y mejorar sobre eso, aunque bajo criterios o reglas diseñadas por personas. En el primero se pretende que el sistema logre hallar una solución a través de entrenamiento y en el segundo, además de esto, se añade el auto entrenamiento sin indicaciones externas.

Muchas de las tecnologías a las que actualmente se hace referencia como IA o se relaciona con ella son, estrictamente hablando, ML de aprendizaje supervisado o no. La acción humana es parte fundamental de la IA en todos los casos, como hemos comentado, y se da en todas las fases de creación y supervisión de la IA:<sup>16</sup>

- En la **etapa de diseño**, donde se entiende el problema a abordar, se adquieren sistemas de IA y se definen sus intenciones, los criterios a seguir, así como los objetivos que debe conseguir esa aplicación final.
- En la **etapa de despliegue**, donde finalmente se diseñan los algoritmos, se limpian y etiquetan los datos de entrada que utilizará la IA, se definen (en ocasiones) los productos del sistema de IA. También se define el modelo usado, es decir, cómo los algoritmos aprenden y se adaptan.
- En la **etapa de aplicación** se decide cómo aplicar y usar finalmente los productos de IA.

---

<sup>16</sup> Véase nota *supra* 8, p. 5, párr. 5.

Ejemplos de aplicaciones o herramientas actuales que incorporan la IA:

- Dispositivos del hogar, como termostatos, aspiradoras móviles, neveras inteligentes, etc.
- Filtros de *spam* que detectan y pueden archivar correos “sospechosos”.
- Videojuegos.
- *Speech recognition* o reconocimiento del habla, que permite recibir e interpretar la voz del usuario como un dato de entrada más en las aplicaciones web.
- Anuncios personalizados, que nos aparecen al abrir una página web, y que se basan en la información que buscadores o redes sociales tienen de nosotras.
- Los llamados “agentes expertos”, que son sistemas muy entrenados en una actividad intelectual específica, como jugar al ajedrez.
- *Chatbots*, que utilizan el NLP (*Natural Language Processing*) y aprenden y mejoran en cada experiencia permitiendo una comunicación bidireccional coherente con las personas, ya sea oral o escrita.
- Vehículos autónomos que son capaces de conducir solos.
- Asistentes virtuales, orientados principalmente a facilitar la realización de tareas a los usuarios de los mismos a través normalmente de la voz.

# cap. 3

## Delimitación del derecho a la libertad de opinión y de expresión en el Sistema Universal de Derechos Humanos<sup>17</sup>

---

<sup>17</sup> Apartado extraído, en su mayor parte, de Castilla K., Libertad de expresión y derecho de acceso a la información. Mecanismos y estándares internacionales, Institut de Drets Humans de Catalunya (IDHC), 2019, p. 13-20, disponible en: <https://www.idhc.org/es/investigacion/publicaciones/otras-publicaciones/libertad-de-expresion-y-derecho-de-acceso-a-la-informacion-mecanismos-y-estandares-internacionales.php>

**E**n el Sistema Universal de Derechos Humanos (SUDH), la libertad de opinión y de expresión está reconocida ampliamente en dos instrumentos: La Declaración Universal de Derechos Humanos (DUDH) y el Pacto Internacional de Derechos Civiles y Políticos (PIDCP), ambos en su artículo 19.

**Declaración Universal de Derechos Humanos (DUDH).** Proclamada por la Asamblea General de las Naciones Unidas el 10 de diciembre de 1948.

### Artículo 19

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.

15

**Pacto Internacional de Derechos Civiles y Políticos (PIDCP).** Adoptado el 16 de diciembre de 1966. En vigor desde el 23 de marzo de 1976.

### Artículo 19

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

Además de en la declaración y el tratado, el derecho a la libertad de expresión también está reconocido expresamente en la Convención Internacional sobre la Eliminación de todas las Formas de Discriminación Racial (artículo 5, d, viii); en la Convención sobre los Derechos del Niño (artículo 13); en la Convención Internacional sobre la Protección de los Derechos de todos los Trabajadores Migratorios y de sus familiares (artículo 13); y en la Convención sobre los Derechos de las Personas con Discapacidad (artículo 21).

## 3.1. Qué es y no es la libertad de expresión para el SUDH

- El derecho, reconocido a toda persona o todo individuo sin discriminación de ningún tipo, comprende:
  - La libertad de buscar informaciones e ideas de toda índole.
  - La libertad de recibir informaciones e ideas de toda índole.
  - La libertad de difundir informaciones e ideas de toda índole.
  - El derecho a no ser molestado a causa de sus opiniones.
  - El derecho de investigar informaciones y opiniones.
  - El derecho de difundir informaciones y opiniones.
  - El derecho a cambiar de opinión en el momento y por el motivo que la persona elija libremente.<sup>18</sup>
  - La libertad de no expresar las opiniones propias.<sup>19</sup>
  
- El derecho no comprende:
  - La propaganda en favor de la guerra.
  - La apología del odio nacional, racial o religioso que constituya incitación a la discriminación, la hostilidad o la violencia.
  
- El derecho protege todas las formas de opinión, como son: Políticas o pensamiento político, científicas, culturales, artísticas, históricas, morales, religiosas o pensamiento religioso, periodísticas, educativas o de enseñanza, comentarios sobre los asuntos propios o sobre asuntos públicos, campañas puerta a puerta, discusión sobre derechos humanos, publicidad comercial e incluso expresiones que pueden considerarse profundamente ofensivas.

16

<sup>18</sup> Comité de Derechos Humanos, *Observación General No. 34 sobre el Artículo 19, Libertad de opinión y libertad de expresión*, Comité de Derechos Humanos, 2011, p. 2, párr. 9, disponible en: <https://www.acnur.org/fileadmin/Documentos/BDL/2012/8507.pdf>

<sup>19</sup> *Ibidem*, p. 3, párr. 10.

Adicionalmente la Observación General N° 34 del Comité de Derechos Humanos de la ONU (encargado de interpretar y velar por la implementación del PIDCP), explícitamente reconocía en 2011 que el Artículo 19 del Pacto también aplica a todas las formas de expresión que se transmitan a través de internet y/o de aparatos electrónicos, además de las ya recogidas previamente.<sup>20</sup> Es decir, la libertad de expresión debe ser protegida de igual manera, sea ejercida de forma *online* o de forma *offline* o fuera de internet.

El que fue RE sobre libertad de expresión, Frank La Rue, en su informe de 2011 añadía que al ser la libertad de expresión un “factor coadyuvante de otros derechos”, internet, al funcionar como catalizador del derecho a la libertad de expresión, facilita igualmente otros derechos humanos “entre ellos los derechos económicos, sociales y culturales, como el derecho a la educación y el derecho a participar en la vida cultural y gozar de los beneficios del progreso científico y de sus aplicaciones, así como los derechos civiles y políticos, como los derechos de libertad de asociación y reunión”. En este sentido el Relator señalaba que el acceso a “la infraestructura y las tecnologías de la información y la comunicación, como cables, módems, computadoras y *software*, para acceder a Internet en primer lugar implican libertad de expresión”.<sup>21</sup>

## 3.2. Caracterización del derecho de acceso a la información

El derecho de acceso a la información ha derivado de la libertad de expresión y se ha configurado como un derecho independiente que en el SUDH tiene como caracterización mínima la siguiente:

- Es un derecho para acceder a la información en poder de los organismos públicos:
  - Poderes del Estado (ejecutivo, legislativo y judicial).
  - Autoridades públicas o de gobierno, cualquiera que sea su nivel (nacional, regional o local).
  - Entidades que ejerzan funciones públicas (entidades semi estatales o privadas).
- Esta información comprende los registros de que disponga el organismo público, independientemente de la forma en que esté almacenada la información, su fuente y la fecha de producción.
- El derecho comprende:
  - El derecho de los medios de comunicación a tener acceso a la información sobre los asuntos públicos.

20 Estas son: oralmente, por escrito, en forma impresa, de forma artística u objetos artísticos, por lenguas de señas o signos, braille, medios y formatos aumentativos y alternativos de comunicación y todos los demás modos, medios y formatos de comunicación accesibles que elijan las personas con discapacidad, imágenes, por cualquier procedimiento, por cualquier medio de expresión y mediante cualquier forma de comunicación, por ejemplo: libros, periódicos, folletos, internet, medios electrónicos, pancartas, etc.

21 Informe A/HRC/17/27 del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, remitido por el Consejo de Derechos Humanos en mayo de 2011, p. 7 y 8, párr. 22, disponible en: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G11/132/04/PDF/G1113204.pdf?OpenElement>

- El derecho del público en general a que los medios de comunicación le proporcionen los resultados de su actividad.
- El derecho de verificar si hay datos personales almacenados en archivos automáticos de datos, de obtener información inteligible sobre cuáles son esos datos y con qué fin se han almacenado, así como que, en su caso, se rectifiquen esos datos.
- El derecho a consultar la historia clínica personal.

### 3.3. Límites y restricciones a la libertad de expresión

Las restricciones deben estar fijadas por la ley y deben cumplir pruebas estrictas de necesidad y proporcionalidad:

- a. Deben estar previstas por ley de manera clara y accesible para todos(as) (principios de previsibilidad y transparencia).
- b. Debe obedecer a uno de los fines establecidos en el párrafo 3 del artículo 19 del Pacto, que son: i) asegurar el respeto a los derechos o a la reputación de los demás, o ii) proteger la seguridad nacional, el orden público o la salud o la moral públicas (principio de legitimidad);
- c. Ser necesarias y proporcionales: el Estado debe demostrar que la restricción es necesaria para proteger un interés legítimo y que es el medio menos restrictivo para lograr el objetivo perseguido y deben ser proporcionales para:
  - Asegurar el respeto a los derechos humanos o a la reputación de los demás (a título individual o como integrantes de una comunidad).
  - La protección de la seguridad nacional, el orden público o la salud o la moral públicas.
  - Prevenir toda propaganda en favor de la guerra.
  - Prevenir toda apología del odio nacional, racial o religioso que constituya incitación a la discriminación, la hostilidad o la violencia.

Además, el artículo 20 (2) del PIDCP establece que cualquier odio racial o religioso que constituye una incitación a la discriminación, hostilidad o la violencia debe estar prohibido por la ley.

#### Artículo 20

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

En este sentido, tal y como señala el RE sobre libertad de expresión en el mismo informe de 2011 citado más arriba, “entre los tipos legítimos de información que pueden restringirse cabe mencionar la pornografía infantil (para proteger los derechos del niño y niña), la incitación verbal al odio (para proteger los derechos de las comunidades afectadas), la difamación (para proteger los derechos y la reputación de los demás contra ataques injustificados), la incitación directa y pública a cometer actos de genocidio (para proteger los derechos de los demás) y el fomento del odio nacional, racial o religioso que constituya incitación a la discriminación, hostilidad o violencia (para proteger los derechos de los demás, como el derecho a la vida).”<sup>22</sup>

## 3.4. Otros derechos vinculados a la libertad de expresión

Existen otros derechos vinculados o muy relacionados con el derecho a la libertad de expresión que pueden verse afectados por el uso que se dé a las tecnologías digitales, incluyendo la IA. No son objeto de este análisis pero los señalamos, junto a vínculos e impactos generales que luego detallaremos porque, por un lado, están referidos en apartados posteriores aunque sea de manera implícita, y, por otro, para visibilizarlos de cara a análisis específicos posteriores que puedan llevarse a cabo.

### Derecho a la privacidad

El derecho a la privacidad, recogido en el artículo 12 de la DUDH y el artículo 17 del PIDCP, establecen que “nadie será objeto de injerencias arbitrarias en su vida privada, su familia, su domicilio o su correspondencia, ni de ataques a su honra o a su reputación. Toda persona tiene derecho a la protección de la ley contra tales injerencias o ataques”.

En muchas ocasiones el derecho a la libertad de expresión está precedido por el derecho a la privacidad y los sistemas basados en IA que adoptan decisiones y moderan contenido, dependen del almacenaje y la explotación de datos personales, recopilan información de las usuarias para saber qué les gusta, qué hacen y qué piensan hacer en el futuro, por ejemplo, y lo hacen muchas veces sin su conocimiento sobre ese proceso de almacenaje y sobre todo del uso que se dará a esa información.<sup>23</sup> Las personas pierden, en definitiva, el control de sus propios datos.

Por otro lado, en un informe de 2015, el RE sobre libertad de expresión señalaba la necesidad de proteger la privacidad de las personas usuarias de internet como puerta de entrada a la libertad de opinión y de expresión. Esta privacidad pasa muchas veces, por “la necesidad de seguridad en línea para buscar, recibir y difundir informaciones sin riesgo de repercusiones, divulgación, vigilancia y otros usos indebidos de sus opiniones o expresiones”. Introducía el cifrado, el anonimato y el uso de pseudónimos

<sup>22</sup> Véase nota *supra* 21, p. 8, párr. 25.

<sup>23</sup> Article 19, *Blockchain and freedom of expression*. 2019, pág. 16 – 18, disponible en: <https://www.article19.org/wp-content/uploads/2019/07/Blockchain-and-FOE-v4.pdf>

como parte del derecho a la libertad de expresión y como herramientas facilitadoras de privacidad y de protección de opiniones, de expresión y creencias, aunque sin olvidar “el lado oscuro” de estas herramientas y el uso que personas acosadoras, terroristas y otras pueden hacer de ellas para perjudicar a otras personas.<sup>24</sup>

Como vemos, la relación entre ambos derechos es compleja porque, aun reconociendo que se refuerzan mutuamente, podrían entrar igualmente en conflicto, ya que puede darse el caso en que la publicación de ciertas informaciones vulnere el derecho a la privacidad de determinadas personas o grupos de personas.

### Libertad de pensamiento, de conciencia y de religión

El derecho a la libertad de pensamiento, de conciencia y de religión, que incluye la libertad de tener creencias, (párrafo 1 del artículo 18 de la DUDH) es profundo y de largo alcance; abarca la libertad de pensamiento sobre todas las cuestiones, las convicciones personales y el compromiso con la religión o las creencias, ya se manifiesten a título individual o en comunidad con otras personas. El artículo 18 además distingue entre la libertad de pensamiento, de conciencia, de religión o de creencias y la libertad de manifestar la propia religión o las propias creencias, muy vinculado con el artículo 19.

El Comité de Derechos Humanos señalaba<sup>25</sup> que “la libertad de tener o adoptar una religión o unas creencias comporta forzosamente la libertad de elegir la religión o las creencias, comprendido el derecho a cambiar las creencias actuales por otras” y esto pasa necesariamente por el derecho al acceso de la información como está concebido por el SUDH (ver apartado 3.2.).

20

Los impactos del uso de las tecnologías digitales en la libertad de expresión que veremos a continuación relacionados sobre todo con la gestión del contenido y la desinformación pueden, por tanto, afectar igualmente a la libertad de pensamiento, conciencia y religión.

### Libertad de reunión y asociación pacífica

El derecho a la libertad de expresión y el derecho de asociación y de reunión pacífica (artículo 20 de la DUDH) están muy interrelacionados: la propia acción de reunirse puede conllevar la misma expresión pública del tipo que sea; artística, de denuncia o protesta, entre otras.<sup>26</sup> La acción de protestar conlleva además el acceso a información de diferente índole y utilizando diversos medios, entre ellos los digitales. Los y las defensoras de derechos humanos y activistas deben poder ejercer todos estos derechos para llevar a cabo su labor.

---

24 Informe A/HRC/29/32 del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión (David Kaye), remitido por el Consejo de Derechos Humanos en mayo de 2015, p. 4, párr. 6 y 7, disponible en: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G15/095/88/PDF/G1509588.pdf?OpenElement>

25 Comité de Derechos Humanos de la ONU, *Observación General Nº 22, Libertad de pensamiento, de conciencia y de religión* (art.18), Comité de Derechos Humanos, 1993, párr. 5, disponible en: <https://wri-irg.org/en/story/1993/general-comment-22-article-18?language=es>

26 Para más detalle del derecho, consultar Rivera C. y Castilla K. *Libertad de reunión pacífica y de asociación: Mecanismos y estándares internacionales*, Institut de Drets Humans de Catalunya, 2019.

Según ha manifestado la ex Relatora Especial sobre el Derecho a la libertad de reunión pacífica y de asociación de la ONU, la libertad de reunión y asociación pacífica, da derecho a formar o a unirse a un grupo, y protege de la obligación de unirse a una asociación y entre estas asociaciones se incluyen sindicatos, clubes, asociaciones religiosas, partidos políticos y, cada vez más, grupos en Internet.<sup>27</sup> El uso de tecnologías digitales aporta ventajas a este derecho; por ejemplo, puede facilitar la organización de las protestas. Pero también tiene sus riesgos, muchos de los cuales coinciden con los de la libertad de expresión que veremos a continuación. Es el caso, por ejemplo, de la censura fruto de la moderación de contenido o la vigilancia masiva.

### No discriminación

El artículo 19 de la DUDH incorpora la no discriminación en el derecho a formarse opiniones, en la expresión de ideas e información y el acceso a ellas. El derecho a la libertad de expresión se incluye también, como ya hemos comentado anteriormente, en declaraciones relacionadas con grupos vulnerabilizados, tradicionalmente discriminados, como las mujeres, las niñas y niños, la población indígena, personas racializadas, entre otras. Es evidente que la no discriminación debe ser un principio innegable para ejercer el derecho a la libertad de opinión y expresión.

En otra publicación del IDHC<sup>28</sup> hablamos extensamente de la no discriminación en relación a la IA, con lo que no nos detendremos en esto ahora. Solo subrayar que, por un lado, en relación a las tecnologías digitales, no todo el mundo tiene acceso a las mismas y por tanto, esa falta de acceso puede agravar la discriminación de ciertos grupos y favorecer, por tanto, el incremento de la desigualdad en nuestras sociedades. Por otro, en relación a la IA y su uso en la moderación de contenidos y en la transmisión de noticias, pueden perpetuar y reforzar estereotipos, prejuicios y actitudes discriminatorias por el motivo que sea (género, orientación sexual, origen nacional, etc.). Por tanto, las herramientas de IA no solo incorporan los sesgos de las personas que están detrás de su diseño, sino que, dependiendo del uso que se les dé, pueden ser incluso un instrumento para potenciarlos.

21

---

<sup>27</sup> Ver en United Nations News, *Artículo 20: libertad de reunión y asociación*, United Nations, 2018, disponible en <https://news.un.org/es/story/2018/11/1447301>

<sup>28</sup> Véase nota *supra* 16.

# cap. 4

## Vínculos, problemas y retos en relación al uso de tecnologías digitales para la libertad de expresión

A lo largo del documento hemos podido ir identificando algunos vínculos o nexos entre el derecho a la libertad de expresión y el uso de las tecnologías digitales. Estos vínculos generan indudables ventajas pero también innegables zonas oscuras o retos para la protección y garantía de la libertad de expresión.

El objetivo de este apartado es llevar a cabo un análisis un poco más acotado en relación a algunos de estos vínculos, así como a problemas y desafíos derivados de ellos. Desde los sistemas internacionales de protección de derechos, academia, sociedad civil organizada y otros sectores especializados se han realizado investigaciones que tratan de analizar y recoger propuestas para contrarrestar estos problemas, parte de los cuales intentamos mostrar aquí.

No es este un análisis sencillo, son varias las dificultades que se enfrentan. Por un lado, los vínculos y los resultados de esos vínculos son tremendamente complejos y cada uno podría ser objeto de una investigación extensa e independiente.

Los problemas están muy relacionados entre sí, las fronteras entre ellos no siempre son claras y además esa relación puede ser muy variable: un vínculo o problema concreto podría o bien incentivar otro o bien estar dentro de las propuestas de solución de ese otro. Por ejemplo, el uso de algoritmos puede facilitar la difusión de un discurso de odio en línea, pero también podría ayudar a combatir ese mismo discurso.

Por último, ya venimos comentando que la constante evolución que sufre este tipo de tecnologías es vertiginosa, lo que dificulta aún más un análisis completo y actual, y es por esto que estos estudios deberían estar siempre abiertos a nuevas incorporaciones. Lo que aquí abordamos por tanto, son factores o rasgos generales de algunos de los impactos más destacados actualmente.

## 4.1. El perfilado y la presentación del contenido

A poca gente le es desconocida la palabra *cookie* a estas alturas. Una *cookie* no es más (ni menos) que un “trocito de código que se pega al navegador cuando navegamos por internet y nos identifica de manera única.”<sup>29</sup> Las cookies se utilizan para gestionar las preferencias de las personas usuarias y la personalización de los sitios web. Recaban información sobre qué webs visitamos, las compras o reservas que hacemos, las cosas que decimos que nos gustan, las que no nos gustan, las consultas del tipo que sean. Todos nuestros movimientos en internet, en definitiva. Con todos esos datos y también los clásicos que incorporamos en internet (nombre y apellidos, domicilio, estado civil, trabajo, etc.) se generan perfiles y, en función de esos perfiles, los algoritmos nos facilitan o hacen visible una información determinada y nos “esconden” o dificultan el acceso a otra. Esa información que tenemos más o menos visible puede ser de todo tipo: comercial, relacionada con noticias, sucesos de nuestro entorno o de otros más lejanos, entre otras muchas cosas.

El perfilado y la consecuente presentación del contenido suponen, como primera injerencia a la libertad de expresión, que las personas usuarias tengan una libertad relativa de elegir, acceder y buscar aquella información relevante para ellas. La información que, a priori, es objetiva e ilimitada en internet, se vuelve así bastante limitada y subjetiva. A esto se añade el hecho de que lo normal es que se consulten solo los primeros resultados que aparecen en un buscador y raramente los que aparecen en una segunda o tercera página.<sup>30</sup> Y estas búsquedas rápidas no son solo fruto de una motivación personal, sino que se ven incentivadas por técnicas facilitadas o promovidas por la IA, como el *engagement* (compromiso) sobre el que volveremos más adelante, o el *trigger*<sup>31</sup> (gatillo).

24

Otro impacto negativo muy relacionado con el perfilado es que contribuye a la desinformación y la polarización social. Gracias al trabajo del algoritmo, las usuarias no acceden a información diversa, plural o disonante o contraria a sus creencias, ya que las noticias que reciben tienden a confirmar y reforzar sus teorías. En las redes sociales, son los algoritmos quienes determinan la manera en que se selecciona y ordena la información: la que el algoritmo considera relevante, en función de la actividad de las personas en las redes, aparece en primer lugar, ocultando la no relevante. A lo que se suma que no solo aparece en determinado orden, sino limitada: muchas veces las personas reciben información que otras usuarias cuelgan y no necesariamente de la fuente primaria, de la web que la ha generado.

---

29 Peirano M., *El enemigo conoce el sistema: Manipulación de ideas, personas e influencias después de la economía de la atención*, Ed. Debate, 2019, p. 94.

30 Brkan M., *Freedom of expression and Artificial Intelligence: on personalisation, disinformation and (lack of) horizontal effect of the Charter*, Maastricht University, 2018, p. 2-3, disponible en <https://cris.maastrichtuniversity.nl/en/publications/freedom-of-expression-and-artificial-intelligence-on-personalisat>

31 En el entorno digital “*trigger*” (gatillo), hace referencia a un disparador, desencadenante, activador o señal que tiene el objetivo de implantar rutinas y generar necesidades para que las usuarias permanezcan en internet el mayor tiempo posible.

Es lo que Eli Pariser llama *filter-bubbles*<sup>32</sup> (filtros burbuja), término que describe los efectos de los algoritmos de búsqueda y organización en el tipo de información que las personas reciben. El autor define estos filtros como “una selección personalizada de la información que recibe cada individuo que le introduce en una burbuja adaptada a él para que se encuentre cómodo, pero que está aislada de las de los demás”. Para que se sienta cómodo y para producir lo que se ha denominado en inglés el *engagement*,<sup>33</sup> que hace referencia, entre otras cosas, a la generación de automatización del comportamiento de la persona usuaria para que no deje de consultar webs, para generar una rutina en ella, llegando incluso a la adicción en el peor de los casos.

Cabe señalar, de todas formas, que la existencia de estos filtros burbuja está siendo muy cuestionada por diversos estudios empíricos que básicamente argumentan que las personas usuarias de internet tienen acceso a mucha más información que las no usuarias y, por tanto, tienen la oportunidad al menos de conocer un mayor número (y más diverso) de opiniones. En todo caso, lo que sí parece evidente es que el riesgo de la existencia de estos filtros es mayor en aquellos países donde hay una menor diversidad de medios *online*<sup>34</sup> y es muy peligrosa en aquellos contextos donde la única fuente de información proviene de redes sociales.<sup>35</sup>

Algo parecido son los *echo-chambers*, que hacen referencia a “la tendencia de individuos con ideas afines a discutir entre sí para fortalecer sus ideas”.<sup>36</sup> Estos “espacios” favorecen solo a ciertos tipos de medios de comunicación, lo que aumenta los niveles de polarización en la sociedad, que puede devenir en pérdida de cohesión social y fragmentación de la esfera pública.

Las decisiones sobre este acceso, tradicionalmente tomadas por humanos, ahora son impulsadas por aplicaciones de IA con un funcionamiento muy poco claro. Estas aplicaciones tienen la capacidad de excluir o enfatizar selectivamente información crítica, y pueden servir para que los gobiernos las implementen sin rendir cuentas. El resultado es una falta de transparencia y gran opacidad en el acceso a la información.<sup>37</sup>

25

---

32 Ver por ejemplo la charla de Fundación telefónica *Eli Pariser: Fake news, los filtros burbuja y la batalla por la verdad*, disponible en <https://espacio.fundaciontelefonica.com/evento/eli-pariser-fake-news-los-filtros-burbuja-y-la-batalla-por-la-verdad/#:~:text=Esto%20es%20lo%20que%20Eli,de%20las%20de%20los%20dem%C3%A1s%E2%80%9D>

33 Además de esta connotación, el “*engagement*” hace también referencia al contrato legal donde el usuario (mediante un simple “clic”) renuncia a sus derechos en cuanto al uso de sus datos.

34 Véase nota *supra* 30, p. 3.

35 Reuters Institute, *Digital News Report 2021* (10th edition), University of Oxford, 2022 p. 52-56, disponible en <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021>

36 Véase nota *supra* 30, p. 3.

37 Véase nota *supra* 8, p. 12-14.

## 4.2. Moderación del contenido: bloqueo y eliminación automática

El primer gran bloqueo, y por tanto primera limitación a la libertad de expresión, viene de algunos estados que suspenden directamente el acceso a internet o a determinadas plataformas o páginas web, todas ellas tecnologías digitales. El objetivo último es impedir que las personas usuarias accedan a determinada información. Un ejemplo de esto son los datos que arroja un informe<sup>38</sup> que estudió cómo se da y dónde el bloqueo de medios sociales, que reporta que, desde 2015, uno de cada tres países han restringido estas plataformas y muchas de estas restricciones han ocurrido en eventos sociales o políticos como elecciones, referéndums y protestas.

Por otro lado, las plataformas *online* o digitales<sup>39</sup> utilizan técnicas de *machine learning* para bloquear o eliminar contenido, lo que puede vulnerar o, cuanto menos, limitar el derecho a la libertad de expresión y de opinión. Según criterios humanos incorporados al algoritmo, este puede considerar que las opiniones vertidas son ofensivas, que incitan a la violencia, que son extremistas o que suponen o promueven discursos de odio.

El que fue RE, David Kaye, admitiendo la ventaja de que la IA permite un contenido más amplio y rápido de información a nivel mundial, porque maneja una gran cantidad de datos de manera rápida y ayuda a administrar una gran cantidad de contenidos al sector privado de la tecnología de la información y las comunicaciones, denunciaba su opacidad y falta de transparencia en cuanto a su funcionamiento. Según él, esto genera un alto riesgo de injerencia en la autodeterminación o acción individual y, por tanto, en la libertad de expresión y opinión y advierte de que los Estados corren el riesgo de socavar la libertad de expresión en Internet.<sup>40</sup> No hay que olvidar que un “elemento esencial del derecho a sostener un opinión es el derecho a formarse esa opinión y a desarrollarla mediante el razonamiento”<sup>41</sup> y existen serias dudas de que se utilice ese razonamiento en estos casos donde la información está tan elaborada y dirigida.

Existen varios problemas relacionados con este aspecto de la IA, pero me centraré brevemente en los dos siguientes:

Por un lado, los criterios sobre los que se basan los algoritmos para censurar estos contenidos están, como ya hemos dicho, diseñados por seres humanos y, en muchas ocasiones, reproducen sus sesgos y

38 Ver en Surfshark, *Social media censorship tracker, 2022*, disponible en <https://surfshark.com/social-media-blocking>

39 La Comisión Europea define plataforma digital como “plataformas de publicidad en línea, mercados, motores de búsqueda, medios sociales y puntos de venta de contenido creativo, plataformas de distribución de aplicaciones, servicios de comunicaciones, sistemas de pago y plataformas para la economía colaborativa”. Ver en Comisión Europea, *Las plataformas en línea y el mercado único digital. Retos y oportunidades para Europa* COM(2016) 288 final, 2017, p. 2-3, disponible en: <https://eur-lex.europa.eu/legal-content/es/TXT/?uri=CELEX:52016DC0288>

40 Véase nota *supra* 8, p. 1-2.

41 Informe A/HRC/38/35 del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, remitido por el Consejo de Derechos Humanos en 2018, disponible en <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/096/75/PDF/G1809675.pdf?OpenElement>

prejuicios. Un comentario que puede ser ofensivo para una persona, puede no serlo para otra o no ser ofensivo en absoluto para nadie. Incluso en el caso de que el contenido sea bloqueado por una persona, es decir, sin utilizar IA, es muy probable que la tarea se desarrolle de manera subjetiva, teniendo en cuenta las creencias y opiniones de la persona que la lleva a cabo. Pero si además son los algoritmos los que ejecutan esta actividad, no solo no será objetiva sino que podrá escapar mucho más a nuestro control, por su funcionamiento poco claro, ya que no sabemos cómo se produce exactamente en todas sus fases.

Por otro lado, el algoritmo no está diseñado o preparado para reconocer la ironía, el cinismo el sarcasmo o el humor, no reconoce las emociones que sí es capaz de detectar una persona. El resultado es una limitación a la libertad de expresión o la censura, ya que puede eliminar una gran cantidad de opiniones con un alto grado de error.

No existen instrumentos legales efectivos frente al bloqueo y la eliminación y censura de contenido. En el caso europeo, por ejemplo, tal y como muestra M. Brkan,<sup>42</sup> a pesar de que se ha desarrollado regulación para proteger el derecho a la libertad de expresión frente a estas actividades,<sup>43</sup> donde expresamente se requiere que el bloqueo o la eliminación de contenido debe tener en cuenta la libertad de opinión, no existe procedimiento alguno para restaurar el contenido que finalmente se considere legal. Anterior a esta directiva, una comunicación de la Comisión Europea (CE) relacionada con la lucha contra el contenido ilícito<sup>44</sup> hacía un llamado para abordar el contenido *online* ilegal y encomendaba a las plataformas *online* a que desarrollaran procesos efectivos para detectar y eliminar este tipo de contenidos, especialmente en aquellas tecnologías que utilizan algoritmos para ello, aunque no solo. Frente a esto, el resultado ha sido que las plataformas eliminan incluso de más por miedo a no eliminar lo suficiente y dado que luego no tienen que restaurar ese contenido que no era finalmente ilegal.

27

### 4.3. Desinformación y contaminación informativa

El uso de la desinformación para fines personales o lucrativos no es nuevo y no es exclusivo de la era de internet, ni siquiera actualmente. Hay autores que documentan su utilización ya en la antigüedad, situando su origen en la Grecia antigua. Otros incluso hablan de un origen previo.<sup>45</sup>

<sup>42</sup> Véase nota *supra* 30, p. 3-5.

<sup>43</sup> Ver por ejemplo Directiva 2000/31/CE del Parlamento y Consejo Europeo, relativa a determinados aspectos jurídicos de los servicios de la sociedad de la información, en particular el comercio electrónico en el mercado interior (Directiva sobre el comercio electrónico), 2000, disponible en: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2000-81295>

<sup>44</sup> Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, *Lucha contra el contenido ilícito en línea. Hacia una mayor responsabilización de las plataformas en línea*, COM(2017) 555 final, 2017, p. 2-3, disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52017DC0555&from=it>

<sup>45</sup> Ver por ejemplo Altares G., *La Larga historia de las noticias falsas*, El País, 2018, disponible en: [https://elpais.com/cultura/2018/06/08/actualidad/1528467298\\_389944.html](https://elpais.com/cultura/2018/06/08/actualidad/1528467298_389944.html)

Es habitual, en el imaginario popular, vincular la desinformación con la esfera política, pero no es la única donde se ha utilizado o se utiliza: también la podemos encontrar en el sector económico o empresarial, sobre todo a través de la publicidad engañosa, o en el sector sanitario, por ejemplo a través de campañas desinformativas en relación a la pandemia, o incluso en el sector medioambiental muchas veces en relación al cambio climático.<sup>46</sup>

En cualquier caso, lo nuevo no es, por tanto, su uso, sino las formas en que se utiliza y el grado de repercusión, velocidad, interconexión y alcance que ha conseguido gracias al desarrollo y amplio acceso a las tecnologías digitales, tales como computadoras, *smart phones*, internet, webs y redes sociales, entre otras. Los algoritmos, por su parte, están diseñados para mantener a las usuarias conectadas a las plataformas, como ya hemos visto, y para ello fomentan el acceso a noticias sensacionalistas y reducen la accesibilidad a otras fuentes más diversas y plurales, agitando las emociones y apelando a la indignación.<sup>47</sup> Los algoritmos participan activamente en la divulgación de la desinformación aunque no tanto en su creación, que sigue siendo competencia humana.<sup>48</sup> El acceso a ciertas tecnologías digitales, como las que acabamos de mencionar, facilita, por tanto, la creación y difusión de esa desinformación, que después la IA se encarga de potenciar, a través de la indexación y una mayor divulgación.

Uno de los problemas de inicio para abordar adecuadamente la desinformación, es la falta de consenso en cuanto a su definición. Según la RE sobre libertad de expresión, la desinformación “se presta a abusos por falta de definición y la magnitud y la naturaleza del problema son discutidas por falta de datos e investigaciones suficientes. Las respuestas de muchos Estados han sido problemáticas y excesivas y han afectado negativamente a los derechos humanos”.<sup>49</sup> Ella entiende por desinformación “la creación, difusión o amplificación de información falsa o manipulada por motivos políticos, ideológicos o comerciales”, mientras que el Consejo de Europa prefiere hablar de contaminación informativa (*information pollution*), término que parece muy adecuado, porque parte de la premisa de que el concepto *fake news* ha sido apropiado por políticos en todo el mundo que lo utilizan, muchas veces de manera interesada, para desacreditar a los rivales políticos o incluso a medios de noticias para, de esta manera, socavar la libertad de prensa.<sup>50</sup>

El Grupo de alto nivel sobre las noticias falsas y la desinformación en línea de la Comisión Europea (CE) define la desinformación como “información falsa, inexacta o engañosa diseñada, presentada y promovida para causar intencionalmente daño público o con fines de lucro”, pudiendo causar un perjuicio público. Distingue la desinformación (*disinformation*) de la información errónea (*misinformation*), que hace referencia a la información falsa o inexacta que se difunde, en teoría, de manera involuntaria y sin

46 Wadle C., Derakshan H., Informe del Consejo de Europa DGI(“017)09 *Information disorder: Toward an interdisciplinary framework for research and policy making*, Consejo de Europa, 2017, p. 10-11, disponible en: <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>

47 Ver por ejemplo Martel C., Pennycook G. D., Rand G., *Reliance on emotion promotes belief in fake news*, 2020, disponible en: <https://cognitiveresearchjournal.springeropen.com/articles/10.1186/s41235-020-00252-3>

48 Marsden C., Meyer T., *Regulating Disinformation with artificial intelligence*. European Parliament Research Service, 2019, p. 1-7, disponible en: <https://op.europa.eu/es/publication-detail/-/publication/b8722bec-81be-11e9-9f05-01aa75ed71a1>

49 Informe A/HRC/47/25 de la Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión (Irene Khan) *La desinformación y la libertad de opinión y expresión*, Comité de Derechos Humanos de la ONU, 2021, p. 2, párr.3, disponible en: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G21/085/67/PDF/G2108567.pdf?OpenElement>

50 Véase nota *supra* 45, p. 5 y 14-16.

el objetivo de hacer daño, por ejemplo, copiando y pegando en nuestro perfil de red social un post que hemos visto en otro perfil. Hay autores que señalan además una tercera distinción; la información errónea (*mal-information*), que se refiere a aquella información que es verídica pero que se difunde para producir un daño. Normalmente es información que se diseña para que sea privada pero que alguien la traslada a la esfera pública para perjudicar.<sup>51</sup>

Estas distinciones en los conceptos dan cuenta de otro problema a la hora de abordar la desinformación, que además preocupa de manera creciente entre las usuarias:<sup>52</sup> los límites entre realidad y mentira y entre querer o no causar daño. Porque una parodia puede utilizar información falsa o si pensamos en las creencias o en ciertos conocimientos, estos no son fácilmente catalogables entre verdad o mentira. De la misma manera que información verdadera puede ser expresada para hacer daño, y, por tanto, no sería responsabilidad, ese daño, solo de la información falsa. Ya hemos dicho anteriormente que los algoritmos no están preparados para distinguir las emociones, el sarcasmo o la ironía y que, aun sin utilizar algoritmos, si es una persona la que está detrás, lo está con sus sesgos, prejuicios y estereotipos.

Otro gran desafío que apunta el Parlamento Europeo<sup>53</sup> a la hora de abordar la desinformación es la zona gris en donde se encuentran las noticias falsas dentro de la expresión política, porque abarca, como ya hemos visto, la desinformación y la información errónea. Estas zonas oscuras son aprovechadas por gobiernos y en general en el ámbito de la lucha política, tal y como denuncian diferentes organismos y entidades, para deslegitimar aquellos argumentos divergentes, contrarios o que no interesan. Paralelamente, los intermediarios de internet<sup>54</sup> son presionados, bajo el argumento de luchar contra los discursos de odio, para moderar el contenido y las cuentas (de redes sociales) privadas, lo que puede derivar en censura.

29

Ante esta complejidad de factores, lo más adecuado sería analizar cada caso o campaña de desinformación de manera específica, fijándonos en ciertos aspectos clave que algunas autoras proponen y que apelan al comportamiento en línea, más que al contenido en sí. Se refieren a:<sup>55</sup>

1. Los agentes: conocer la motivación (política, económica, ideológica, etc.) de quienes crean, o producen (cuando se convierte en un producto) o difunden la información.
2. El contenido del mensaje o de la información: el formato, el tipo de información, las características que tiene, entre otros.
3. El sujeto que recibe la información y la interpreta: cómo la entiende y qué acciones lleva a cabo, si es que lleva alguna.

---

51 Véase nota *supra* 45, p. 20-22.

52 Ver por ejemplo Cellan-Jones R., *Fake news worries 'are growing' suggests BBC poll*, 2017, disponible en: <https://www.bbc.com/news/technology-41319683>

53 Véase nota *supra* 47, p. 8-15.

54 Según la Organización Article 19 intermediarios de internet es “un término amplio que se refiere a las entidades que permiten a las personas conectarse a Internet y transmitir contenido. Existen diferentes tipos de intermediarios, tales como los proveedores de acceso a Internet, los proveedores de servicio de alojamiento en la web, plataformas de redes sociales así como los motores de búsqueda. Los intermediarios se distinguen de los productores de contenido, siendo estos últimos aquellas personas u organizaciones que se encargan de la producción de información y su publicación en línea.” Ver en: <https://www.article19.org/es/resources/internet-intermediaries-dilemma-liability-q/>

55 Véase nota *supra* 45, p. 22-24.

Centrarse en los actores, en sus objetivos y fines, puede ser más interesante que centrarse en el contenido del mensaje, para la búsqueda de soluciones. En este sentido, es necesario destacar la necesidad de analizar lo que diversas autoras llaman “el negocio de la desinformación”. Los casos son numerosos,<sup>56</sup> algunos muy conocidos y vinculados además con la lucha política, incluso directamente con diversas campañas electorales. Un informe<sup>57</sup> de 2020 destacaba que durante ese año, 81 países (11 más que el año anterior) habían utilizado las redes sociales para difundir propaganda y desinformación sobre política. A su vez encontraron 48 casos de empresas privadas que desplegaron propaganda en nombre de un actor político, y desde 2009 el gasto en este tipo de contrataciones había sido de 60 millones de dólares a nivel global. La lucha contra la desinformación debe pasar necesariamente por el rastreo de este tipo de negocios.

Los efectos de la desinformación son múltiples y pueden ser demoledores para la libertad de expresión. Generan una clara limitación al pensamiento crítico y a la diversidad de opiniones; el uso de algoritmos para combatirla puede derivar en censura o autocensura; el riesgo de radicalización y polarización social es elevado; la utilización política para censurar y desacreditar argumentos disonantes o contrarios, es decir, lo que no interesa para fines políticos, es un hecho y la influencia de ello en la política, en las elecciones y finalmente en los gobiernos en el poder es evidente.

Por último, un apunte en relación a otro tipo de desinformación muy dañina; la ideológica. Subrayar que puede ser devastadora para personas vulnerabilizadas y también para los defensores y las defensoras de derechos humanos, así como para la labor que llevan a cabo. Gran parte de esta desinformación fomenta la discriminación y el odio hacia minorías, migrantes, mujeres, y otros grupos tradicionalmente excluidos, generando innumerables tensiones y problemas que pueden llegar incluso a situaciones tan graves como el genocidio de Ruanda (previo a internet, pero que fue facilitado por otra tecnología como es la radio) o el más reciente genocidio de la población Rohingya en Myanmar, donde Facebook tuvo, además, un papel determinante<sup>58</sup> divulgando contaminación informativa y facilitando los discursos de odio. En relación a las y los defensores de derechos humanos, la RE sobre libertad de expresión, Irene Kahn, ha denunciado cómo son diariamente violentados y violentadas a través de campañas de desinformación en línea y las terribles consecuencias que esto tiene en su libertad de expresión y en su vida.<sup>59</sup>

30

---

56 Ver Fisher M., *El oscuro negocio de la desinformación por encargo*, The New York Times, 2021, disponible en: <https://www.nytimes.com/es/2021/07/27/espanol/desinformacion-mexico-venezuela.html>

57 Bradshaw S., Bailey, H., Howard P. N. *Industrialized Disinformation 2020 Global Inventory of Organized Social Media Manipulation*, Oxford Internet Institute, University of Oxford, 2020, disponible en: <https://demotech.oii.ox.ac.uk/research/posts/industrialized-disinformation/>

58 Ver por ejemplo Galarraga N., *Facebook fue clave en la limpieza étnica del siglo XXI en Myanmar*, El País, 2018, disponible en: [https://elpais.com/internacional/2018/04/12/actualidad/1523553344\\_423934.html](https://elpais.com/internacional/2018/04/12/actualidad/1523553344_423934.html) o E. Akinwotu *Facebook's role in Myanmar and Ethiopia under new scrutiny*, The Guardian, 2021, disponible en <https://www.theguardian.com/technology/2021/oct/07/facebooks-role-in-myanmar-and-ethiopia-under-new-scrutiny>

59 Véase nota *supra* 48, p. 5-7.

## 4.4. Vigilancia masiva y selectiva

Como dice la periodista Marta Peirano: “toda tecnología desarrollada para luchar contra el terrorismo y por la libertad de otros países, acaba formando parte del aparato de vigilancia doméstico”.<sup>60</sup>

Desde que E. Snowden, a través de *The Guardian*<sup>61</sup> y *The Washington Post*, lo pusiera en evidencia en 2013, es conocido que bajo el pretexto de la seguridad, muchos gobiernos en todo el mundo ponen en marcha tecnologías digitales, concretamente instrumentos de vigilancia masiva, contra su población. Hablamos de tecnologías tales como aplicaciones informáticas de reconocimiento facial (vigilancia biométrica), cámaras de red que se usan para el seguimiento a distancia, sistemas de reconocimiento de emociones que analizan automáticamente las expresiones faciales con la intención de predecir comportamientos, entre otras.<sup>62</sup>

Esta vigilancia puede estar legitimada en algunos casos, pero en otros incumplen claramente uno o todos los supuestos de restricción de la libertad de expresión establecidos por el SUDH que hemos visto anteriormente (ver apartado 3.3.) y también del derecho a la privacidad.

En relación al funcionamiento de la IA, su opacidad y falta de transparencia, unido a la capacidad que tiene para identificar y rastrear comportamientos, puede tener un efecto disuasorio importante sobre la libertad de expresión y afectar cada vez más al espacio (tanto físico como *online*) de la sociedad civil. Las técnicas de *machine learning* o *deep learning* pueden utilizarse para identificar y rastrear a las personas, en diferentes dispositivos y en diferentes lugares: en las calles, en los hogares, en el espacio laboral. El uso de herramientas antes mencionadas, de tecnología digital, como video vigilancia, el reconocimiento facial y gestual, el análisis de comportamiento, etc., por parte de autoridades públicas y empresas privadas obstaculizan gravemente la libertad de expresión y también vulneran el derecho a la privacidad. Incluso en el caso de datos anonimizados, la IA se puede emplear para eliminar el anonimato, “desafiando la distinción entre datos personales y no personales, en el que se basa la normativa vigente en materia de protección de datos”.<sup>63</sup>

Las técnicas de reconocimiento facial pueden rastrear e identificar a personas y eliminar el anonimato en el espacio público. Un informe señalaba<sup>64</sup> en 2018 que los sistemas de aprendizaje automático incluso han podido identificar alrededor del 69% de manifestantes con gorras y bufandas para cubrirse la cara. El reconocimiento facial, dice este mismo informe, puede permitir que la policía identifique a las personas sin causa probable, sospecha razonable o cualquier otro estándar legal.

<sup>60</sup> Véase nota *supra* 29, p. 93.

<sup>61</sup> Ver Greenwald G., MacAskill E., *NSA Prism program taps in to user data of Apple, Google and others*, *The Guardian*, 2013, disponible en: <https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data>

<sup>62</sup> Ver por ejemplo Amnistía Internacional, *Empresas de la UE venden herramientas de vigilancia a responsables de abusos contra los derechos humanos en China*, Amnistía Internacional, 2020, disponible en <https://www.amnesty.org/es/latest/news/2020/09/eu-surveillance-sales-china-human-rights-abusers/>

<sup>63</sup> Article 19, *Privacy and freedom of expression in the Age of Artificial Intelligence*, 2018, p. 18. (original en inglés, traducción de la autora), disponible en: <https://www.article19.org/wp-content/uploads/2018/04/Privacy-and-Freedom-of-Expression-In-the-Age-of-Artificial-Intelligence-1.pdf>

<sup>64</sup> *Ibidem* p. 18.

La vigilancia masiva es una interferencia desproporcionada con la privacidad y la libertad de expresión, mientras que la vigilancia selectiva<sup>65</sup> solo puede justificarse cuando es prescrita por la ley, necesaria para lograr un objetivo legítimo y proporcionada al objetivo perseguido. Ambas pueden tener efectos inmediatos en la libertad de expresión: la autocensura, la alteración del comportamiento y actividades expresivas en los espacios públicos y cambio de actitudes en las comunicaciones privadas. Además de que, como señala la RE, “la vigilancia ha conducido a veces a la detención arbitraria, la tortura y ejecuciones extrajudiciales”.<sup>66</sup>

Frank La Rue ha advertido de que “las prácticas de los Estados habían puesto de manifiesto la ausencia de leyes nacionales adecuadas o la falta de rigor en su aplicación, la escasez de garantías procesales y la ineficacia de la capacidad de supervisión” en relación a la vigilancia digital que vulneraban, en muchas ocasiones, el derecho a la intimidad y a la libertad de opinión y expresión. Según él, la vigilancia digital ya no está reservada a los Estados que dispongan de recursos para llevar a cabo una vigilancia masiva o selectiva, ha entrado en escena el sector privado, “sin supervisión y con una forma de actuar muy cercana a la impunidad”.<sup>67</sup>

Frente a este panorama, las y los activistas, defensores y entidades de derechos humanos buscan y proponen estrategias para salvaguardar su seguridad, sus derechos y su trabajo.<sup>68</sup> Y esto pasa por proteger la propia conexión, las contraseñas, las comunicaciones, los archivos y la forma de compartirlos, entre otros. En este sentido, el anonimato del que hablábamos en relación al perfilado, cobra especial sentido y para ello existen redes de anonimato, siendo Tor (*The Onion Router*) el software más conocido, y técnicas específicas para resistir al bloqueo o evitar la censura.

32

## 4.5. Discursos de odio *online*

No existe una única definición del discurso de odio pero comúnmente hace referencia a “las expresiones que incitan directamente a la comisión de actos de discriminación o violencia por motivos de odio racial, xenófobo, orientación sexual, u otras formas de intolerancia”.<sup>69</sup> También pueden ser aquellos que fomentan el prejuicio<sup>70</sup>.

---

65 La diferencia entre la vigilancia masiva y la vigilancia selectiva es que la primera se hace al conjunto de la población y de grupos poblacionales, mientras que la segunda se dirige a personas concretas, como periodistas, defensoras y defensores de derechos humanos, activistas, políticos, etc.

66 Informe A/HRC/41/35 del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, *La vigilancia y los derechos humanos*, remitido por el Consejo de Derechos Humanos en mayo de 2019, p. 1, disponible en: <https://hchr.org.mx/puntal/wp/wp-content/uploads/2020/06/G1914879.pdf>

67 Informe A/HRC/23/40 del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión remitido por el Consejo de Derechos Humanos en abril de 2013, párr.3., disponible en <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G13/133/06/PDF/G1313306.pdf?OpenElement>

68 Consultar, por ejemplo, las propuestas por la Organización Front Line Defenders en <https://securityinabox.org/es/>

69 Cabo Isasi A., García Juanatey A., *Contrólale en las redes! El discurso de odio en las redes sociales: un estado de la cuestión*, Dirección de Servicios de Derechos de Ciudadanía y Diversidad, Ajuntament de Barcelona, 2017, p. 6, disponible en: <https://bcnroc.ajuntament.barcelona.cat/jspui/handle/11703/114545>

70 Bazzaco Edoardo, García J. Ana, Lejardi Jon, Palacios Anna y Tarragona Laia, *¿Es odio? Manual práctico para reconocer y actuar frente a discursos y delitos de odio*, Institut de Drets Humans de Catalunya, 2017, p. 7, disponible en: [https://www.idhc.org/arxius/recerca/1517393506-ES\\_ODIO\\_Manual\\_practico\\_vf.pdf](https://www.idhc.org/arxius/recerca/1517393506-ES_ODIO_Manual_practico_vf.pdf)

Las tecnologías digitales (internet, webs, plataformas), tienen ciertas características que pueden facilitar e incluso agravar los discursos de odio y que complejizan su erradicación. Algunas de ellas son:

- La “democratización” o descentralización de la comunicación: cualquier persona con acceso a tecnología digital, que no son todas, puede enviar un mensaje que sea leído por una gran cantidad de personas e incluso generar un efecto multiplicador y hasta, gracias al trabajo de la IA, conseguir viralizar el mensaje.
- La permanencia de los contenidos e itinerancia entre diferentes plataformas, que pueden potenciar y agravar el daño de un discurso de odio.
- El uso de pseudónimos, anonimato y transnacionalidad que favorecen los discursos y dificultan la persecución de quienes los crean y difunden.
- La falsa creencia de que la “virtualidad” no es real y lo que ocurre en la red no tiene repercusiones fuera de ella.

Los efectos o consecuencias del discurso de odio (sea *online* o no) son diversos y van desde el daño emocional o psicológico, la erosión de la dignidad de las personas y su reputación, la división y polarización de la sociedad, hasta la deshumanización de ciertos colectivos que puede derivar en actos discriminatorios, abuso y violencia. Casos extremos conocemos varios: el ya mencionado en Myanmar o el caso yugoslavo, que, según el fiscal del caso, incluso persiste<sup>71</sup> 25 años después del genocidio de Srebrenica.

33

El RE, en relación a los discursos de odio, la libertad de expresión y el vínculo con la IA, aceptando el peligro real que tienen ciertos discursos en línea para colectivos especialmente vulnerabilizados y en concreto para limitar sus derechos de expresión y opinión, alerta sobre la utilización que hacen algunos estados del concepto de manera errónea, para censurar y limitar la capacidad de expresión y opinión de ciertos sectores de la sociedad. Ya hemos visto que el uso de la IA es habitual para la gestión y moderación del contenido, siendo uno de sus objetivos el de eliminar ciertos discursos discriminatorios, abusivos o de odio. Utilizan los “clasificadores” de IA para identificar contenidos prohibidos sobre la base de palabras y análisis específicos, con éxito parcial.

Hemos visto (apartado 3.1.) que las expresiones de odio (nacional, racial, religioso) no están comprendidas en el derecho a la libertad de expresión pero el conflicto entre esta y la represión del discurso de odio es evidente: ¿cómo podemos garantizar el derecho a la libertad de expresión protegiendo a las personas de los discursos de odio? Un claro ejemplo de este brete es el uso de pseudónimos en internet que, por un lado, puede ser requisito previo para ejercer la libertad de expresión en línea y, por otro, facilita la proliferación de los discursos de odio bajo el anonimato.

Pero además existen otros desafíos a la hora de enfrentar los discursos de odio en línea: desde dificultad a la hora de atribuir responsabilidades a las redes sociales que juegan un papel determinante en la

---

71 Ferrer I., *El fiscal para Yugoslavia alerta contra el discurso del odio que persiste en los Balcanes*, El País, 2020, disponible en: <https://elpais.com/internacional/2020-06-25/el-fiscal-para-yugoslavia-alerta-contr-a-el-discurso-del-odio-que-persiste-en-los-balcanes.html>

difusión de estos discursos, hasta la falta de formación en instituciones implicadas, falta de concienciación en la sociedad en general, pasando por dificultades para acreditar la motivación de odio.<sup>72</sup>

## 4.6. El riesgo de no ser olvidada

El uso de internet ha generado preocupación en los últimos años en torno a lo que se viene denominando “el derecho al olvido”, entendido como aquel que permite a las personas exigir a los motores de búsqueda<sup>73</sup> la eliminación de la información que se dispone sobre ellas.<sup>74</sup> Lo que subyace en el fondo es el argumento de que cierta información puede perder importancia con el tiempo o incluso podría perjudicar a quien la haya creado, lo que confronta con la idea predominante de que la información debe recopilarse y permanecer accesible para la preservación del registro histórico y para poder reparar abusos a los derechos humanos si se da el caso.<sup>75</sup>

Cabe señalar que este no es un derecho reconocido ni en el ámbito internacional, ni en los nacionales, pero que salió a la luz a raíz de la decisión del Tribunal de la Unión Europea (TJUE) en el caso Google España de 2014, decisión mediante la cual se reconoce el derecho de las personas a poder borrar información que los buscadores muestren sobre ellas. Además, es necesario recordar cómo en la caracterización del derecho de acceso a la información del SUDH visto previamente, el derecho comprendía: “El derecho de verificar si hay datos personales almacenados en archivos automáticos de datos, de obtener información inteligible sobre cuáles son esos datos y con qué fin se han almacenado, así como que, en su caso, se rectifiquen esos datos” (ver apartado 3.2. de este documento).

La capacidad ilimitada de memoria digital, la universalización de la información y los motores de búsqueda que recogen, registran, organizan y vinculan información, conservan también de manera indefinida esa información sin el consentimiento explícito de la persona usuaria. Son las herramientas de IA las responsables, al igual que ocurre con la moderación del contenido o la desinformación, de cortar contenido e indexarlo a otros en momentos determinados.

---

<sup>72</sup> Véase nota *supra* 68, p. 21-23.

<sup>73</sup> Se entiende por motor de búsqueda, también denominado buscador, el sistema informático diseñado para realizar búsquedas de archivos almacenados en una base de datos a través de *spiders* o *bots*. Cuando el buscador recibe una solicitud, realiza una consulta en la base de datos ofreciendo el resultado en una página de resultados. Si son varios los resultados, se muestran jerarquizados en función de su importancia (o en función de lo que el algoritmo cree que es importante para la persona usuaria). Definición extraída en parte de: <https://www.arimetrics.com/glosario-digital/motor-de-busqueda>

<sup>74</sup> Organización Article 19, *The “Right to be Forgotten”: Remembering Freedom of Expression*, 2016, p. 11-13, disponible en: [https://www.article19.org/data/files/The\\_right\\_to\\_be\\_forgotten\\_A5\\_EHH\\_HYPERLINKS.pdf](https://www.article19.org/data/files/The_right_to_be_forgotten_A5_EHH_HYPERLINKS.pdf)

<sup>75</sup> Miguélez M. G., *El conflicto entre el derecho al olvido y la libertad de expresión e información*, Universidad de León, 2018, p. 33-35, disponible en: <https://buleria.unileon.es/bitstream/handle/10612/11749/Garc%EDa%20Migu%E9lez,%20Miguel.pdf;jsessionid=AB2F8638FB01DD9EFDEDB77A8FE93AC2?sequence=1>

Todo ello tiene implicaciones importantes para la libertad de expresión, entre las que destacan:

- La falta de autodeterminación informativa,<sup>76</sup> entendida como la capacidad de las personas de disponer y tener el control de sus datos y por tanto muy vinculada al derecho a la privacidad.
- El uso posterior que se da a esa información y la posible descontextualización y malinterpretación de la información o de las opiniones vertidas por las usuarias.
- La imposibilidad de dejar atrás el pasado, opiniones o creencias vertidas en el pasado que pueden vulnerar en el presente el derecho a la privacidad o intimidad, así como condicionar la libertad de expresión.

Existen argumentos en favor y en contra al derecho al olvido. Entre los primeros destacan los que opinan que es compatible con el derecho a la libertad de expresión, y que es justo que las personas tengan derecho a decidir qué hacer con esas opiniones que publican y que además no siempre tienen, esas opiniones, valor de interés público. Entre las posturas divergentes hay quien opina que las personas no deben tener el derecho absoluto sobre las informaciones que se vierten sobre ellas, que es importante la preservación de la información, como ya hemos mencionado al inicio, y que el “derecho al olvido” es muy restrictivo para la libertad de expresión frente al derecho de respuesta o corrección.

---

<sup>76</sup> Íbidem, p. 36-39.

cap. 5

# Conclusión

A lo largo del documento hemos tratado de identificar y analizar de forma general algunos impactos en la libertad de expresión relacionados con el uso de las tecnologías digitales en general y la IA en particular. Parte de estos impactos suponen o conllevan innegables ventajas para aquellas personas que pueden acceder a las tecnologías que, sin ser todas, son muchas más que hace unas décadas, cuando estas no estaban tan extendidas. El acceso a internet, pero también a equipos como computadores u ordenadores, *smart phones* y *tablets*, ha permitido y permite una mayor y más rápida conectividad entre las personas incluso en lugares antes impensables; una mayor facilidad para expresar opiniones de manera más abierta; el acceso a un número abundante de información, conocimientos y creencias, incluso en tiempo real; formas diversas de archivar y compartir esa opinión y mayores oportunidades de dar voz a personas a las que se les ha dificultado tradicionalmente.

Otras tecnologías digitales pueden generar impactos que nos sitúan frente a importantes retos, desafíos e incluso vulneraciones para los derechos humanos en general y la libertad de expresión en particular, que debemos analizar y abordar. No son cuestiones nuevas y, por tanto, existen ya alternativas y propuestas de solución que, aun siendo parciales, es bueno conocer para que la rápida evolución de estas tecnologías no nos abrume, no nos paralice hasta el punto de que nos impida identificar los diversos usos que finalmente se les dan. Porque el problema no es su existencia, sino cómo se utilizan. Es por ello que cuanto más las conozcamos, mejor sabremos cómo aprovecharlas y mejor también podremos enfrentar estos desafíos.

La “alfabetización digital” en el sentido más amplio del término, cobra especial importancia. Es necesario, no solo disponer de las capacidades para saber utilizar y aprovechar lo que el mundo digital nos ofrece, que es mucho, sino que es urgente conocer, de la forma más completa posible, cómo es ese entorno digital. Esto nos permitirá beneficiarnos de los impactos positivos y mitigar los negativos y supondría además, el paso previo para la construcción o fortalecimiento de redes, alianzas o estrategias conjuntas en la sociedad que incentiven este conocimiento, promuevan la investigación, generen el debate e intercambio y finalmente faciliten la acción.

No es fácil. Como hemos visto, existe una gran opacidad y falta de transparencia sobre todo en relación al funcionamiento de la IA, que abarca cada vez más ámbitos en la toma de decisiones. Esto tiene relación directa con una creciente incertidumbre en cuanto al rol, tanto del sector privado como de los gobiernos, que dificulta a su vez, la asignación de responsabilidades y limita a la hora de legislar. Las personas tienen derecho a saber qué pasa con estas tecnologías, cómo operan y cómo les afecta y son los estados quienes deben dar respuesta a este derecho, quienes deben proteger y garantizar la totalidad de derechos, y las empresas deben a su vez responsabilizarse y rendir cuentas. Porque un aspecto clave que no debemos olvidar, sobre el que hemos hecho hincapié a lo largo del texto, es que detrás de la IA, detrás de los algoritmos, siempre está la mente humana. La IA está ganando terreno en la toma de ciertas decisiones, en su automatización y abarca cada vez más sectores, pero lo que subyace detrás son los problemas ya conocidos que se amplifican en el formato digital.

El derecho a la libertad de expresión recogido en el SUDH impone a los estados obligaciones negativas de abstenerse de tomar medidas que afecten al ejercicio de la libertad de opinión y expresión, pero también obligaciones positivas de promoverla y, por tanto, las tecnologías digitales, incluyendo la IA, deben diseñarse, elaborarse, producirse y desplegarse siendo compatibles con estas obligaciones.

La complejidad de cada uno de los retos mencionados en este documento y de los factores relacionados, unido a la rápida y constante evolución que sufren las tecnologías digitales, requeriría un análisis individualizado para determinar, caso por caso, cómo afectan cada una de ellas a los derechos humanos. La realización de auditorías y actividades de monitoreo que determinen específicamente cómo son esos impactos, revisando la adecuación de la regulación existente, sería lo ideal para, por un lado, disponer de un diagnóstico lo más aterrizado posible, y, por otro, reconducir o ajustar esa regulación.

La adopción de un enfoque de derechos humanos en los análisis que se lleven a cabo permitiría además poner especial atención a la situación de las personas y los grupos vulnerabilizados y discriminados porque suelen ser, como casi siempre, los que más sufren los impactos negativos de los vínculos entre el uso de las tecnologías digitales y la libertad de expresión.

