

Apertura Curso Académico 2026-2027

El poder y la palabra

Leopoldo Abad Alcalá

Catedrático de Derecho Constitucional

Universidad CEU San Pablo

El poder y la palabra

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

© Leopoldo Abad Alcalá, 2026

© De la edición, Fundación Universitaria San Pablo CEU, 2026

Maquetación: Pedro Coronado Jiménez (CEU *Ediciones*)

CEU *Ediciones*

Julián Romea 18, 28003 Madrid

www.ceuediciones.es

Depósito legal: M-24062-2026

ÍNDICE

PARTE PRIMERA

1. Por qué hablar sobre el Poder y la palabra	5
2. Los orígenes: <i>isegoria</i> y <i>parrhesia</i> en Grecia.....	6
3. La libertad para usar la palabra en Roma.....	7
4. La libertad de palabra en el Medievo. Las Partidas de Alfonso X.....	10
5. El cataclismo de la impresión mediante tipos móviles metálicos y la reacción del poder	11
6. La marea liberal. De Milton a las primeras Declaraciones de derechos	14
7. Los españoles somos así. La libertad de prensa en la historia constitucional española.....	16

PARTE SEGUNDA

1. El espejismo reticular	20
2. En el pecado, llevamos la penitencia.	22
3. Dura lex, sed quae lex?	29
4. El nuevo telón de silencio	32
5. De los tribunales al algoritmo: la privatización de los derechos.....	34
6. Las mutaciones del poder	37

PARTE TERCERA

1. ¿De qué hablamos cuando hablamos de Inteligencia Artificial?	41
2. Los nuevos dueños de las palabras	44
3. Democracia <i>versus</i> Inteligencia Artificial	50
4. La defensa europea de la democracia ante la invasión de los ladrones de datos.....	53
5. Una nueva defensa por la Iglesia católica de la dignidad inviolable de la persona: <i>Magnifica Humanitas</i>	56
EPÍLOGO	59
NOTAS	68

PARTE PRIMERA

«El hombre no tiene naturaleza, sino que tiene... historia»
(José Ortega y Gasset).

1. Por qué hablar sobre el Poder y la palabra

Hace 30 años comencé a impartir clase en esta Universidad de la mano de los añorados Luis Escobar de la Serna y Alejandro Muñoz-Alonso, por lo que podrán comprender la emoción y el honor que supone dirigirme a ustedes en este acto tan señalado en la que, sin duda, no puedo sino considerar mi casa.

Antes de comenzar esta lección inaugural del curso 2026-2027 quisiera felicitar al nuevo Presidente de la Asociación Católica de Propagandistas, D. Josep Masip, deseándole todos mis mejores deseos para su nuevo labor con la confianza de que aprovechará y explotará todo el talento multicolor que tiene delante de usted.

Sobre la elección del tema de la lección inaugural, quise combinar mis dos ámbitos de conocimiento, el Derecho y por eso voy a hablarles del poder, pues el Derecho no es sino el mecanismo civilizada que tenemos los humanos para convivir en paz¹; y la Comunicación, que es la forma en que las palabras nos permiten relacionarnos², pues no hay comunidad sin comunicación.

Poder y palabra podrían dar lugar a múltiples combinaciones, desde el poder de las palabras, las palabras del poder, las palabras frente al poder, el poder frente a las palabras. Como es difícil decidirse, y aún a riesgo de que la amplitud reste precisión, voy a disertar de todo un poco.

Hace 37 años llegaba al Colegio Mayor de San Pablo para convertirme en un paulino más y comenzar mis estudios de Periodismo en el ya desaparecido edificio de columnas del Colegio de Estudios Universitarios. El primer día de clase, a las 8,30 de la mañana, en clase de Teoría de la Comunicación, llegó el profesor, nos miró fijamente y nos dijo: «Ustedes son más importante que Dios...»

Nuestra sorpresa quedó suprimida cuando tras una prolongada pausa dramática completó la frase: «Ustedes son más importantes que Dios porque Dios construyó el mundo en una mañana, y ustedes los construyen cada mañana». Este exordio me ha resonado desde entonces en la cabeza, haciéndome consciente del inmenso poder de las palabras para crear el mundo en el que vivimos y el poder de los medios de comunicación (en su sentido más amplio como luego veremos) para conformar nuestra realidad. Es precisamente esa inmensa autoridad de las palabras la que ha hecho que el poder siempre recele de ellas y las haya intentado controlar, limitar o condicionar a lo largo de la historia.

2. Los orígenes: *isegoria* y *parrhesia* en Grecia

Para un constitucionalista, que es en lo que los designios del art. 20 de la Constitución me han convertido, hablar del poder y la palabra debe retrotraernos a la Atenas del siglo V a.C.³. En ella, se dan gran parte de las bases teóricas que con su inevitable evolución y en un marcado contexto histórico, han conformado las ideas en las que se basa la libertad para hacer uso de la palabra.

En Grecia, se parte de la idea de la vinculación consustancial entre democracia, libre expresión de las ideas, discusión de los asuntos públicos y crítica al poder. Aparecen ya conceptos como la *isegoria* (o el derecho de todos los ciudadanos a hablar en la asamblea (*ekklesia*), uso de la palabra que amparaba la *parrhesia* (o todo decir, ese natural impulso a manifestar con sinceridad absoluta y sin inhibición alguna los juicios personales⁴), que se fundamenta en la *euleteria* (la igualdad de todos ante la ley)⁵. No obstante, la *parrhesia* se ve limitada por la *asebeia* (la impiedad) que supone penas que van desde la sanción económica al destierro. También se diferencia entre *doxa* (conocimiento fundamentado) y *episteme* (conocimiento basado en la experiencia), diferencia que en la actualidad cobra especial sentido, como luego veremos.

Como explica González Ballesteros⁶, pronto la *parrhesia* sin trabas degenera en las asambleas públicas en simple demagogia, en la vida privada o en el teatro en difamación, y en las escuelas en la posibilidad de inculcar doctrinas corruptas a la juventud. Y comienzan a producirse ensayos de cortapisas a su desarrollo por sus

teóricos excesos en la Asamblea, la Academia y las artes. La *asebeia* se generaliza, y por este delito, hubo de huir de Atenas el filósofo Anaxágoras, (500-428 a.C.) y se denunció a Fidias (490-430 a.C.). El comienzo del proceso a Sócrates (469-399 a.C.) se debe a la gravedad de los ataques que le dirige el comediógrafo Aristófanes (446-386 a.C.) en «Las Nubes». Poco tiempo después comienzan a dictarse normas más restrictivas, como el *Decreto de Siracosis* (414 a.C.) que prohibió a poetas y comediantes ridiculizar a los miembros del gobierno y a cualquier persona citándola por su nombre, o el Decreto de Eucrates (337 a.C.) que extiende la *asebeia* a las expresiones contra la seguridad del Estado, si empleásemos una concepción actual.

En esta época, a pesar de que el sistema ateniense se basaba en una modelo de distribución del poder⁷ que impedía que éste se concentrara en alguna institución y menos en alguna persona, la democracia ateniense acabó feneciendo por una sucesión de crisis, derrotas militares e intervenciones externas entre finales del siglo v y el iv a.C.⁸.

3. La libertad para usar la palabra en Roma

Hablar de libertad de expresión en Roma exige una primera cautela metodológica, ya que los romanos no conocieron un derecho subjetivo a la libertad de expresión equivalente al que desarrollará el constitucionalismo moderno. La posibilidad de hablar libremente se insertaba más bien en la noción general de *libertas*, vinculada a la condición de ciudadano, a la participación en la *res publica* y a determinadas prácticas políticas y sociales. La *libertas romana* no constituía una facultad natural inherente a toda persona, sino un conjunto de posiciones y prerrogativas reconocidas por el ordenamiento y vinculadas a la ciudadanía⁹. De manera semejante, Díaz de Valdés sostiene que la libertad de expresión romana fue fundamentalmente una prerrogativa política, intensamente ejercitada durante la República, pero nunca reconocida legislativamente como un derecho autónomo¹⁰.

La instauración del Alto Imperio o Principado a partir de Augusto produjo una transformación decisiva. Aunque inicialmente se conservaron las formas republicanas y buena parte de los derechos tradicionales de los ciudadanos –bajo Augusto los derechos y libertades esenciales de los ciudadanos romanos permanecieron intactos–, la concentración progresiva del poder en el *princeps* redujo los espacios institucionales en los que anteriormente podía ejercerse la crítica política. Las asambleas populares perdieron relevancia, el Senado quedó

crecientemente sometido a la voluntad imperial y la crítica al gobernante dejó de ser una condición funcional del sistema político para convertirse potencialmente en una amenaza para él. Así, durante el Imperio la libertad de expresión dejó de ser concebida como una prerrogativa política y pasó a depender, en gran medida, de la tolerancia del emperador.

El principal instrumento jurídico de esta transformación fue la *lex maiestatis*. Nacida durante la República para proteger la *maiestas populi Romani* frente a comportamientos que perjudicasen gravemente al Estado, su ámbito se amplió con el Principado hasta proteger también la dignidad y seguridad del emperador. Tácito describe gráficamente esta evolución al recordar que originariamente «*facta arguebantur, dicta impune erant*» –se perseguían los hechos, mientras las palabras quedaban impunes–, pero que Augusto comenzó a aplicar la legislación sobre *maiestas* a los escritos difamatorios (*famosi libelli*) (Annales, I, 72). La doctrina considera este cambio fundamental y la *maiestas* adquirió tal flexibilidad que terminó protegiendo tanto al emperador reinante como a sus predecesores divinizados frente a cualquier menoscabo de su dignidad, estatus o seguridad¹¹. La difamación, que anteriormente podía constituir fundamentalmente un ilícito privado, adquirió así una dimensión criminal cuando afectaba al poder imperial.

La aplicación de estas restricciones, sin embargo, no fue uniforme durante todo el Alto Imperio. Dependió en gran medida del carácter y de la política de cada emperador. Tiberio permitió la utilización de los procesos de *maiestas* contra determinados escritores y opositores, mientras que otros emperadores mostraron mayor tolerancia. El caso paradigmático es el del historiador Cremucio Cordo, procesado en el año 25 d. C. por haber elogiado a Bruto y denominado a Casio «el último de los romanos». Sus escritos fueron condenados a ser quemados y Cordo terminó dejándose morir de hambre. Este episodio muestra la tensión entre historiografía, memoria y libertad de expresión bajo un régimen en el que la crítica abierta podía tener consecuencias personales extremadamente graves¹². La literatura y la historiografía desarrollaron por ello técnicas de discurso indirecto, ironía y ambigüedad que permitían criticar al poder sin hacerlo explícitamente.

La situación alcanzó especial dureza bajo emperadores como Nerón o Domiciano, asociados por la tradición literaria con la persecución de filósofos, senadores y escritores, el exilio y la destrucción de obras. Pero tampoco debe proyectarse sobre todo el Alto Imperio una imagen de censura permanente. Subsistieron la sátira, la crítica literaria y cierta capacidad de expresión pública, y hubo períodos comparativamente más tolerantes. Como advierte Morton Braund, la sátira romana fue consciente de que podía encontrarse con la censura, pero también

de que podía «esquivarla si era suficientemente inteligente»¹³. La principal transformación fue, por tanto, cualitativa, hablar libremente dejó de descansar sobre una cultura política republicana y pasó a depender del equilibrio entre la *libertas* tolerada y la protección de la autoridad imperial.

En el Bajo Imperio, iniciado con Diocleciano en 284 d. C., esta evolución se inserta en un contexto político diferente. La mayor centralización administrativa, el reforzamiento de la autoridad imperial y la elevación ceremonial de la figura del emperador redujeron todavía más la significación política de la antigua *libertas republicana*. Sin embargo, sería excesivo hablar de un sistema general y organizado de censura comparable a los Estados contemporáneos. En la Antigüedad tardía no existió una legislación general destinada a prohibir la libertad de expresión ni un organismo estatal encargado sistemáticamente de vigilar todo lo que los ciudadanos decían o escribían. Lo decisivo era más bien la existencia de límites políticos y sociales variables y, sobre todo, la conciencia de que determinadas palabras podían producir graves consecuencias, favoreciendo así fenómenos de autocensura¹⁴.

Ello no significa que desaparecieran las restricciones jurídicas. Al contrario, las constituciones imperiales mantuvieron una regulación severa de los *libelli famosi*. Diversas disposiciones promulgadas entre 319 y 406 d. C. ordenaron la destrucción de escritos difamatorios y establecieron penas contra sus autores y, en determinados supuestos, incluso contra quienes los difundieran. El título 34 del libro IX del *Codex Theodosianus*, *De famosis libellis*, reunió posteriormente varias de estas normas.

A este elemento político se añadió en el Bajo Imperio una nueva dimensión religiosa. Especialmente desde finales del siglo IV, la consolidación del cristianismo como elemento de unidad imperial condujo a la aprobación de normas destinadas a combatir determinadas doctrinas heréticas y a limitar controversias religiosas capaces de alterar el orden público. El título 16.4 del *Codex Theodosianus*, *De his qui super religione contendunt*, contiene varios preceptos dirigidos precisamente a esas controversias. Detrás de estas disposiciones se encontraba la creciente consideración de la unidad religiosa como condición para la estabilidad del Estado, de forma que determinados discursos religiosos dejaron de ser una cuestión exclusivamente teológica para adquirir relevancia jurídico-política¹⁵.

Pero tampoco en este período desapareció completamente la posibilidad de «hablar con franqueza al poder» (*parrhesia*). Al contrario, se transformaron sus protagonistas y sus formas. En la Antigüedad tardía, obispos, filósofos y determinados

miembros de las élites podían asumir el papel de interlocutores capaces de amonestar al emperador. Un ejemplo paradigmático es Ambrosio de Milán frente a Teodosio I, que incorporó elementos cristianos a la tradición clásica de la *libertas*, presentando como deber del sacerdote advertir al gobernante cuando consideraba que éste actuaba injustamente. La *parrhesia* siguió siendo una práctica reconocible, aunque exigía posición social, habilidad retórica y, en ocasiones, un considerable valor personal.

4. La libertad de palabra en el Medievo. Las Partidas de Alfonso X.

Tampoco durante la Edad Media puede hablarse todavía de una libertad de expresión configurada como derecho subjetivo frente al poder. El uso libre de la palabra aparece, por el contrario, jurídicamente condicionada por la posición que ocupa el individuo dentro de la comunidad y sometida a la protección de bienes considerados superiores, como la verdad religiosa¹⁶, la autoridad política, el honor y la paz social. Sin embargo, el desarrollo de las universidades¹⁷, la *disputatio* escolástica y los conflictos en torno a la herejía y la difusión de doctrinas fueron delimitando progresivamente un espacio específico para la expresión del pensamiento, cuya transformación definitiva vendría impulsada por la aparición de la imprenta.

En esta época, debemos hacer referencia a las Siete Partidas de Alfonso X. Se trata de la principal norma de esta época, constituyendo uno de los grandes cuerpos jurídicos medievales castellanos. Las Partidas regulan expresamente distintas formas de utilización ilícita de la palabra. Así, la Partida VII, título IX, Ley I¹⁸ regula las «deshonras» cometidas tanto mediante hechos como mediante palabras¹⁹. Define la *iniuria* y contempla expresamente supuestos como denostar, ridiculizar, atribuir nombres ofensivos o infamar públicamente a otra persona. Se trata de un antecedente remoto de uno de los límites contemporáneos de la libertad de expresión, el derecho al honor y a la reputación. El Derecho canónico desarrolló igualmente una importante tradición de protección de la *bona fama*. Las Partidas también regulan la herejía (Partida VII, título XXVI), estableciendo consecuencias jurídicas para quienes sostienen creencias contrarias a la doctrina de la Iglesia. Incluso encontramos una auténtica regulación de cómo debe hablarse ante el poder. La Partida II establece reglas sobre las palabras pronunciadas ante el rey que deben ser verdaderas, mesuradas y respetuosas, rechazándose determinadas expresiones ofensivas, la lisonja o las acusaciones indebidas.

La blasfemia constituye otro ejemplo paradigmático de palabra jurídicamente prohibida. No se protege solamente a una persona frente a una expresión, sino el propio orden religioso de la comunidad. Las Partidas incorporan una regulación específica de las expresiones contra Dios, la Virgen y los santos. Las sanciones podían ser extremadamente severas y variar según la condición social y la reincidencia.

Sin embargo, la sociedad medieval europea fue una sociedad predominantemente oral, en la que el dominio de la lectura y, especialmente, de la escritura se concentró durante siglos en sectores minoritarios, particularmente vinculados al clero y a las estructuras administrativas. Aunque resulta imposible establecer tasas generales fiables de analfabetismo para el periodo medieval, las estimaciones disponibles sitúan la alfabetización muy por debajo del 20 % de la población²⁰. Esta característica, unida al hecho de que, durante la Alta Edad Media, una parte fundamental de la producción de libros en Europa se desarrollaba en los monasterios, particularmente en los *scriptoria* a mano por monjes (como podemos ver en El nombre de la Rosa de Umberto Eco), evitaba que la influencia de la palabra escrita fuese especialmente perniciosa para los intereses del poder. Es cierto que la aparición de las Universidades²¹ en los siglos *x*i y *xii* generaron una demanda de libros que el sistema tradicional de copia difícilmente podía satisfacer. Se desarrolló entonces el llamado sistema de la *pecia*²². Un ejemplar autorizado de una obra (*exemplar*) se dividía en cuadernos o *peciae*. Diferentes copistas podían alquilar simultáneamente distintas partes y copiarlas en paralelo, acelerando considerablemente la reproducción de una obra.

5. El cataclismo de la impresión mediante tipos móviles metálicos y la reacción del poder

Pero sin duda, el hecho que supondrá un hito fundamental en la relación entre poder y palabra será el sistema de impresión mediante tipos móviles metálicos desarrollado en Maguncia por Johannes Gutenberg y que puede datarse hacia 1450²³. A partir de entonces, este cambio tecnológico es crucial, pues altera la capacidad limitada de difusión de la palabra escrita, que en gran medida estaba motivada por la baja alfabetización, la producción artesanal, el elevado coste y la circulación relativamente restringida de manuscritos.

La imprenta descompone radicalmente esas condiciones. A partir de entonces, una misma idea puede reproducirse en centenares de ejemplares, circular territorialmente y alcanzar públicos mucho mayores²⁴. El problema jurídico deja

progresivamente de ser únicamente qué palabras pueden decirse y pasa también a ser qué puede imprimirse y difundirse.

Es precisamente después de Gutenberg cuando empiezan a desarrollarse los principales mecanismos de control de la palabra por el poder, principalmente la censura previa y las licencias de impresión, en una pugna entre libertad de palabra y control gubernamental que nutrirá fecundamente la historia del pensamiento político y desembocará siglos después en las declaraciones de derechos del XVIII.

La reacción del poder no es inmediata. En España, los Reyes Católicos en carta fechada en Sevilla el 25 de diciembre de 1477 eximían a Teodorico Alemán, impresor y mercader de libros, del pago de los derechos de alcabala, almojarifazgo y otros tributos²⁵. Pero el avance de las ideas renacentistas, las tentativas reformistas y sobre todo, la «Encíclica sobre la imprenta» de Alejandro VI hicieron cambiar el parecer de los monarcas sobre la imprenta que se concretó en la famosa «Pragmática Real» de 8 de julio de 1502 donde se establecía en Castilla un sistema de control preventivo de la palabra impresa, en la que ningún libro podía imprimirse ni ponerse en circulación sin obtener previamente autorización de la autoridad competente²⁶. Dicha potestad de control preventivo de las publicaciones será posteriormente otorgada al Consejo Real²⁷, al Tribunal de Santo Oficio y el juez de Imprentas²⁸, entre otros²⁹.

Mención específica merece el Tribunal del Santo Oficio. El control de la palabra impresa en la Monarquía hispánica respondió a una cierta distribución de funciones entre la autoridad regia y la Inquisición. Mientras la Corona asumió fundamentalmente el control preventivo de la producción impresa mediante el régimen de licencias, la Inquisición concentró buena parte de su actividad en el control *a posteriori* de la importación, circulación, posesión y lectura de los libros. El Santo Oficio fue construyendo durante el siglo XVI un sistema organizado de control ideológico a través del libro, pero su actuación se centró especialmente en la circulación de las obras. La licencia de impresión correspondía a la autoridad civil, mientras que la Inquisición podía actuar posteriormente contra una obra ya impresa. La posesión de libros prohibidos podía generar sospecha de herejía, mientras que su lectura estaba castigada con excomunión³⁰. Para ello desarrolló un complejo sistema de censura basado en la expurgación de obras, los Índices de libros prohibidos³¹, la inspección de librerías y bibliotecas y la vigilancia de lectores. De este modo, el control inquisitorial trascendió el propio objeto impreso para proyectarse sobre la recepción y difusión social de las ideas³².

Sin embargo, en una irrefutable prueba de la leyenda negra española tan promocionada en la literatura académica de otros países, encontramos en otros países instituciones tan represivas o más que la propia Inquisición. En Portugal, la *Inquisição Portuguesa*, establecida formalmente en 1536 por la bula *Cum ad nihil magis* de Pablo III, desarrolló tribunales permanentes y utilizó mecanismos de censura de libros. En los Estados italianos la institución más comparable fue la Inquisición romana, reorganizada por Pablo III mediante la bula *Licet ab initio* de 1542, que creó la Congregación de la Inquisición romana y universal. Su importancia para la historia de la censura fue extraordinaria porque Roma desarrolló, junto con la actividad inquisitorial, un sistema centralizado de prohibición de libros. El gran hito fue el *Index Librorum Prohibitorum*, cuya primera versión romana apareció en 1559 bajo Pablo IV³³. Francia presenta un modelo diferente. No existió una Inquisición nacional comparable a la española. El control de los libros quedó repartido entre la Monarquía, el Parlamento de París y la Facultad de Teología de la Sorbona. Esta última desempeñó un papel particularmente relevante examinando y condenando doctrinas y publicaciones consideradas heterodoxas. A partir del siglo XVI, además, la Monarquía fue construyendo un sistema cada vez más desarrollado de privilegios y permisos de impresión³⁴.

Pero nos detendremos en la labor desarrollada durante los siglos XVI y XVII por la Cámara Estrellada (*Star Chamber*) en Inglaterra, una de las principales instituciones utilizadas por la Monarquía inglesa para vigilar la palabra impresa. Este control se articuló mediante la actuación conjunta de la Corona, las autoridades eclesiásticas y la *Stationers' Company*, corporación que agrupaba a impresores y libreros y cuyos intereses económicos terminaron convergiendo con el interés político del poder en controlar la circulación de determinadas ideas. El sistema de censura inglés no puede entenderse únicamente como una prohibición de contenidos, sino también como un mecanismo de control sobre la propia estructura material que hacía posible su difusión³⁵.

Durante el reinado de Isabel I, el Decreto de la Cámara Estrellada de 1586 reforzó la inspección sobre las imprentas, limitando su número y restringiendo los lugares en los que podían establecerse, pues el control de quién podía imprimir y dónde facilitaba enormemente la vigilancia. Quizá el momento culminante llegaría medio siglo después. En 1637, durante el reinado de Carlos I, la Cámara Estrellada aprobó una regulación extraordinariamente minuciosa de la actividad impresora, que requería del permiso previo del poder para imprimir³⁶.

Los libros y folletos debían ser examinados y autorizados antes de imprimirse, las publicaciones tenían que registrarse, se controlaba la actividad de impresores y

libreros, se restringía el número de prensas, se vigilaba la importación de libros extranjeros, y podían perseguirse e incautarse las publicaciones clandestinas. El Decreto de 1637 representó, el punto culminante del poder y de la autoridad de la *Stationers' Company* sobre el comercio del libro³⁷.

La historia de la Cámara Estrellada tuvo, además, un desenlace particularmente significativo. En 1641, en el contexto del enfrentamiento entre Carlos I y el Parlamento, la Cámara Estrellada fue abolida. Sin embargo, la desaparición de esta institución no significó el final de la censura previa. La extraordinaria proliferación de escritos políticos y religiosos durante los primeros años de la década de 1640 llevó al propio Parlamento a aprobar en 1643 una nueva *Licensing Order*, que restableció mecanismos de autorización previa y control de las publicaciones. Y es precisamente contra esta decisión lo que hará a John Milton publicar en 1644 su célebre *Areopagítica*³⁸.

6. La marea liberal. De Milton a las primeras Declaraciones de derechos

La *Areopagítica* de John Milton constituye una de las manifestaciones paradigmáticas de la resistencia frente al control de la palabra por el poder. Su defensa de la libertad de expresión se inscribe en la progresiva consolidación de un ideario liberal que, a lo largo de los siglos XVII y XVIII, fue impregnando el clima intelectual, social y político europeo y americano hasta cristalizar en las primeras declaraciones de derechos. En ellas, aquellas reivindicaciones filosóficas y políticas frente a la injerencia del poder encontrarían finalmente formulación jurídica, convirtiendo la libertad de la palabra en un auténtico derecho frente al Estado.

Las construcciones doctrinales³⁹ desarrolladas durante los siglos XVII y XVIII en defensa de la tolerancia, la libertad de conciencia y la libre comunicación de las ideas terminarían encontrando su traducción jurídica en las últimas décadas del siglo XVIII. Se produce entonces un cambio de extraordinaria importancia en la relación entre el poder y la palabra. A partir de entonces se inicia un modelo donde la libertad de palabra deja progresivamente de limitarse a una reivindicación filosófica frente al poder para transformarse en una posición jurídica reconocida y protegida frente a él, iniciando un fenómeno que Alejandro Muñoz-Alonso denomina «constitucionalización de la libertad de expresión»⁴⁰. Las circunstancias históricas hicieron que las colonias americanas y Francia lideraran este proceso de reconocimiento de la libertad de imprenta, aunque con modelos, fundamentos filosóficos y consecuencias jurídicas muy diferentes⁴¹.

En las colonias norteamericanas, la tradición inglesa de rechazo de la censura previa confluyó con el pensamiento iusnaturalista y contractualista y con una intensa práctica de discusión política a través de periódicos, panfletos y folletos, proceso que la ruptura con Gran Bretaña aceleró. Las primeras declaraciones de derechos de los nuevos Estados comenzaron a reconocer expresamente la libertad de prensa como un derecho. El texto paradigmático es la Declaración de Derechos de Virginia de 12 de junio de 1776, redactada principalmente por George Mason. Su artículo 12 proclamaba que la libertad de prensa constituye uno de los grandes baluartes de la libertad y que únicamente puede ser restringida por gobiernos despóticos. Ello supone un cambio jurídico radical, pues frente al modelo inglés configurado principalmente como ausencia de censura previa, aparecía ahora expresamente reconocido como facultad individual en una declaración de derechos. El ejemplo de Virginia fue seguido, con distintas formulaciones, por otros Estados. Las constituciones y declaraciones revolucionarias norteamericanas fueron convirtiendo progresivamente la libertad de prensa en uno de los elementos característicos del nuevo orden republicano⁴².

Aunque no contiene un reconocimiento expreso de la libertad de expresión, es paradigmática la Declaración de Independencia de los Estados Unidos del 4 de julio de 1776, cuyo principal impulsor fue Thomas Jefferson⁴³, pues proporciona algunos de los fundamentos filosófico-políticos sobre los que podrá construirse posteriormente su positivización. La Declaración parte de la existencia de unos derechos que no proceden de la voluntad del legislador ni de una concesión del poder político. Determinadas verdades son consideradas evidentes por sí mismas, como que los hombres han sido creados iguales y se encuentran dotados de ciertos *unalienable rights*, entre los que figuran la vida, la libertad y la búsqueda de la felicidad. Pero la afirmación verdaderamente decisiva desde la perspectiva de la teoría de los derechos aparece inmediatamente después, que los gobiernos son instituidos precisamente para garantizar esos derechos y obtienen sus poderes legítimos del consentimiento de los gobernados.

Sin embargo, la Constitución federal de 1787 no incluyó inicialmente una declaración de derechos. Esta ausencia provocó importantes objeciones durante el proceso de ratificación. Resulta particularmente significativa la ratificación de Virginia de 1788, en la que se reclamaba expresamente que libertades esenciales como la libertad de conciencia y de prensa no pudieran ser restringidas por las autoridades federales. La respuesta llegaría con las Enmiendas a dicha Constitución ratificadas el 15 de diciembre de 1791. La Primera Enmienda estableció que el Congreso no podría aprobar ninguna ley que restringiera la libertad de

palabra o de prensa: «*Congress shall make no law [...] abridging the freedom of speech, or of the press*». El cambio resulta trascendental. La libertad de prensa deja de presentarse simplemente como una esfera en la que el poder debe abstenerse de establecer censura previa y se convierte en una limitación constitucional dirigida al propio legislador federal. El Congreso, aun siendo representante de la voluntad popular, encuentra frente a sí una esfera de libertad que la Constitución le impide restringir⁴⁴.

La Revolución francesa dará un paso conceptualmente diferente. La Declaración de los Derechos del Hombre y del Ciudadano de 26 de agosto de 1789 no se limita a proclamar una específica libertad de prensa, sino que inserta la posibilidad de hablar, escribir e imprimir dentro de un principio mucho más amplio como es la libre comunicación de los pensamientos y de las opiniones. El artículo 10 establece, en primer lugar, que nadie debe ser molestado por sus opiniones, incluso religiosas, siempre que su manifestación no altere el orden público establecido por la ley. Pero será el artículo 11 el que formule la nueva concepción de manera especialmente significativa: «*La libre communication des pensées et des opinions est un des droits les plus précieux de l'Homme*». Y de esta premisa extrae inmediatamente la consecuencia de que todo ciudadano puede hablar, escribir e imprimir libremente, respondiendo del abuso de esta libertad en los casos determinados por la ley.

El artículo 11 incorpora, por tanto, una segunda característica que será fundamental en la posterior tradición europea, la libertad no es ilimitada. El ciudadano puede hablar, escribir e imprimir libremente, pero responde por los abusos de esa libertad «en los casos determinados por la ley». Se reproduce así, bajo una fundamentación distinta, la separación entre libertad de comunicación y responsabilidad posterior⁴⁵ que ya había aparecido en la tradición inglesa⁴⁶.

7. Los españoles somos así. La libertad de prensa en la historia constitucional española

En nuestro país, el proceso de reconocimiento de la libertad de imprenta está vinculado a su historia constitucional⁴⁷. Antes incluso de aprobarse nuestra primera Constitución en 1812, el Decreto IX, de 10 de noviembre de 1810, sobre libertad política de imprenta, reconoció la facultad de escribir, imprimir y publicar ideas políticas sin necesidad de licencia, revisión o aprobación previa, lo que suponía una ruptura con el modelo anterior basado en el control preventivo de las publicaciones.

Será, por tanto, la Constitución de Cadiz de 1812, la que en su artículo 371 reconozca que «Todos los españoles tienen libertad de escribir, imprimir y publicar sus ideas políticas sin necesidad de licencia, revisión o aprobación alguna anterior a la publicación, bajo las restricciones y responsabilidad que establezcan las leyes». El artículo 131 incluía, además, entre las facultades de las Cortes la de «proteger la libertad política de la imprenta». Para los liberales de Cádiz la libertad de imprenta constituía un mecanismo destinado tanto a formar como a expresar la opinión pública, y su fundamentación estaba estrechamente relacionada con el iusnaturalismo liberal⁴⁸.

La accidentada historia constitucional del siglo XIX español determinaría, sin embargo, que la extensión efectiva de esta libertad dependiera estrechamente de la orientación política de cada régimen. La Constitución de 1837 avanzó en su reconocimiento al establecer en su artículo 2 que todos los españoles podían «imprimir y publicar libremente sus ideas sin previa censura, con sujeción a las leyes». Desaparecía así la referencia de 1812 exclusivamente a las «ideas políticas», ampliándose potencialmente el ámbito material de la libertad. Además, se introducía una importante garantía al disponer que la calificación de los delitos de imprenta correspondiera a jurados.

La Constitución moderada de 1845 mantuvo formalmente la libertad de imprenta. Su artículo 2 reconocía igualmente el derecho de los españoles a imprimir y publicar libremente sus ideas sin previa censura, aunque desaparecía la garantía constitucional del jurado. Esta modificación resulta reveladora de una constante que recorrerá buena parte del constitucionalismo español del siglo XIX, como es que la proclamación constitucional de la libertad podía convivir con una legislación ordinaria restrictiva que condicionara intensamente su ejercicio. El verdadero alcance de la libertad de imprenta dependía, por ello, no solo del texto constitucional, sino también de las sucesivas leyes de imprenta y de la actuación administrativa y gubernativa.

El gran salto cualitativo se producirá con la Constitución de 1869. Su artículo 17 reconoció a todos los españoles el derecho a «emitir libremente sus ideas y opiniones, ya de palabra, ya por escrito, valiéndose de la imprenta o de otro procedimiento semejante». La modificación terminológica expresa una transformación conceptual de enorme importancia. La libertad protegida deja de identificarse con el medio –la imprenta– para referirse directamente a la actividad comunicativa traducida en emitir ideas y opiniones. La palabra oral, la escritura y la imprenta pasan a ser formas distintas de ejercer una libertad más general. Puede hablarse ya, por tanto, con mayor propiedad, de una verdadera libertad de expresión.

La Constitución de 1869 reforzó, además, las garantías frente a los mecanismos históricos de control de la prensa. Su artículo 22 impedía establecer para los delitos cometidos mediante la imprenta una legislación especial y prohibía, entre otras medidas, el secuestro de impresos antes de que comenzaran a circular. La libertad de expresión quedaba así integrada en la concepción democrática de los derechos que caracterizó al texto nacido de la Revolución de 1868.

Mantuvo formalmente este avance conceptual la Constitución de 1876. Su artículo 13 reconoció a todo español el derecho «de emitir libremente sus ideas y opiniones, ya de palabra, ya por escrito, valiéndose de la imprenta o de otro procedimiento semejante». Se conservaba, por tanto, la formulación amplia de 1869. Sin embargo, el modelo de la Restauración remitió la regulación del ejercicio de estos derechos a la legislación ordinaria, lo que permitió que su extensión efectiva fluctuara considerablemente dependiendo de las circunstancias políticas y de la legislación sobre prensa.

Un nuevo momento fundamental llegará con la Constitución de la Segunda República de 1931. Su artículo 34 proclamó que «toda persona tiene derecho a emitir libremente sus ideas y opiniones, valiéndose de cualquier medio de difusión, sin sujetarse a previa censura». La formulación presenta dos novedades particularmente relevantes. En primer lugar, el titular ya no es exclusivamente «el español», sino «toda persona». En segundo lugar, desaparece la enumeración de medios concretos característica de los textos decimonónicos y se utiliza la expresión abierta «cualquier medio de difusión».

La Constitución republicana reforzó, teóricamente, además la prohibición de censura previa y estableció garantías específicas frente al secuestro administrativo de publicaciones y la suspensión gubernativa de periódicos. Se pretendió consolidar así una concepción en la que la libertad de expresión no queda vinculada a una determinada tecnología comunicativa, sino a la facultad misma de exteriorizar y difundir ideas y opiniones.

Sin embargo, todo este modelo nació inservible, pues esta regulación fue anulada por la Ley de Defensa de la República, de 21 de octubre de 1931, que permitió importantes restricciones gubernativas. De hecho, la Ley es anterior a la Constitución (9 de diciembre), pero continuó vigente tras su aprobación. La Ley de Defensa de la República consideraba «actos de agresión a la República» determinadas conductas expresivas formuladas de manera bastante amplia. Entre ellas figuraban la difusión de noticias que pudieran quebrantar el crédito o perturbar la paz o el orden público, así como determinadas formas de apología o incitación.

Y lo más relevante desde el punto de vista constitucional es que su aplicación quedaba en manos de la autoridad gubernativa, permitiendo medidas administrativas que incidían directamente sobre prensa y expresión.

Finalmente, la Constitución de 1978 culmina esta evolución histórica mediante el artículo 20.1.a), que reconoce y protege el derecho «a expresar y difundir libremente los pensamientos, ideas y opiniones mediante la palabra, el escrito o cualquier otro medio de reproducción». A ello añade derechos diferenciados pero estrechamente relacionados, como la producción y creación literaria, artística, científica y técnica (20.1.b); la libertad de cátedra (20.1.c); y el derecho a comunicar o recibir libremente información veraz por cualquier medio de difusión (20.1.d). También prohíbe la censura previa (20.2) y el secuestro administrativo (20.5). Para saber más sobre ella, les invito a mis clases.

PARTE SEGUNDA

«El medio es el mensaje»
(Marshall McLuhan)

1. El espejismo reticular

Toda nueva tecnología de la comunicación supone un cambio en las formas en que los humanos nos comunicamos, alterando nuestra forma de usar las palabras para adaptarlas al nuevo medio. La radio y la televisión supusieron un cambio significativo en cómo se entendía la comunicación, pero no implicaron un cambio de paradigma en el uso de la palabra, que seguía monopolizado por empresas privadas que compartían protagonismo en el espacio público con medios impresos y con los sistemas públicos audiovisuales. Pero el acceso a dichos medios seguía restringido al poder, fuese económico, social o político.

La aparición de Internet sí supuso un cambio significativo de paradigma comunicativo. De un modelo de comunicación de uno a muchos, hemos pasado a un modelo de comunicación de muchos a muchos.

Los primeros pasos de Internet dieron lugar a un ciberoptimismo desmesurado sobre las posibilidades de esta nueva tecnología para la libertad de palabra¹ que he denominado como el espejismo reticular (creyendo que la Red saciaría nuestra sed de libertad de expresión, pero sólo era el reflejo del algoritmo en la arena). Quizá sí tengamos más libertad de expresión², pero la cuestión es si tenemos mejor libertad de expresión.

Volviendo a la Grecia del siglo v. a.C., vivimos en el paroxismo de la libertad de expresión. El ideal griego se ha desbocado y las redes sociales e Internet nos permiten más *isegoria* (todos podemos intervenir libremente en la nueva *ekklesia* virtual), usamos desinhibidos nuestra *parrhesia* (el anonimato y el fortalecimiento algorítmico de las polarización nos invitan a hablar sin cortapisas), pero la diferencia entre *doxa* (conocimiento informado y motivado) y *episteme* (la opinión que se nutre sólo de la experiencia) es cada vez mayor. Lamentan Carr y Keen el «culto electrónico al amateur», que privilegia desmesuradamente la participación masiva sobre la autoridad, la apertura a la pericia³, y lleva irónicamente a decir a Jason Lanier que muchos creen que «un millón, o quizás mil millones de insultos fragmentarios producirán al final la sabiduría»⁴.

Internet posee una serie de características que lo diferencian de los medios tradicionales y que transforman el proceso comunicativo. Como indica Revenga⁵, la libertad de palabra adquiere su forma más acabada cuando va acompañada de la voluntad de que perdure y trascienda mediante la plasmación en un soporte que permita la difusión a un número indeterminado de destinatarios. Internet comporta la paradoja de que el número de destinatarios y su permanencia en el tiempo es potencialmente infinita («lo que compartes en línea es un tatuaje»), cuando no existe un soporte físico entendido en el sentido clásico que sustente nuestros mensajes⁶. La ampliación *ad infinitum* del público destinatario es consecuencia también del papel multiplicador de los buscadores⁷, como puso de manifiesto la sentencia clave del Tribunal de Justicia de la Unión Europea en el caso de Mario Costeja que inició la posterior configuración normativa del derecho al olvido⁸.

Otro de los elementos característicos de Internet es como las coordenadas clásicas en los que se ha fundamentado el Derecho de la Información, la distinción entre libertad de expresión (transmisión de pensamientos, ideas y opiniones –art.20.1.a) CE–) y derecho a la información (difundir y recibir hechos veraces y de interés público –art.20.1.d) CE–) se difumina⁹, máxime cuando los contenidos en Internet tienen, cada vez más, carácter híbrido, llegándose incluso a dar nombre a esta tendencia: *infoentertainment*. Ello ha dado lugar a que algún autor cambie la clásica alocución «los hechos son sagrados, las opiniones libre» por «las opiniones son libre, los hechos caros». Cuando los periódicos y los medios en línea están buscando un modelo de negocio, la tentación de confundir interés público con interés del público es grande. Se venden más ejemplares, hay más audiencias, y se atraen más ojos hacia la publicidad del sitio web o la app, potenciando el efecto pernicioso del *clickbait*.

Quizá el ámbito de Internet donde mejor se refleje la confusión imperante en el nuevo ecosistema digital en el que vivimos sean las redes sociales. La capacidad adictiva que las caracterizan no es casual. Como indica Peirano, copian de manera literal los principios de diseño de las máquinas más adictivas conocidas, «las máquinas tragaperras»: a) Principio del aislamiento. b) Principio del infinito c) Principio de la velocidad¹⁰. Todo ello programado para la generación de dopamina, el neurotransmisor vinculado a las adicciones¹¹.

Como pueden imaginar, las consecuencias en términos de poder de esta vinculación a las redes sociales e Internet son inmensas. La vinculación de los usuarios es el primer paso para un mecanismo prediseñado de control y dominación, porque a partir de entonces dejamos en manos de las grandes plataformas de Internet nuestra capacidad de decisión.

2. En el pecado, llevamos la penitencia.

Internet nos ofrece al acceso a información infinita (un smartphone contiene más información que la biblioteca de Alejandria), y nuestro pecado venial es navegar sin fin a través del *scroll* infinito de *reels*, *shorts* o tuits. Prueba de la sobreinformación de Internet es como a Facebook se suben 500 horas de video al minuto, 720.000 horas al día. Para ver los videos subidos a Internet en un solo día, una persona necesitaría 82 años. Sin dormir, claro está. Y sólo vería los de 2025. (Se perdería muchos gatitos de otros años)

Ante la infinitud de contenidos en la Red, hemos de pagar la penitencia de someternos a los mecanismos de selección y moderación de las plataformas de Internet implementados a través de sus algoritmos, que se han convertido en los verdaderos «editores invisibles» en Internet, haciendo que perdemos nuestra capacidad de decisión sustituida por el algoritmo de turno. Es innegable la progresiva importancia de la regulación y moderación de contenidos digitales ante fenómenos como la desinformación y la selección de aquello que los usuarios publican o visualizan. Ello nos conduce a la distinción entre regulación de contenidos, atribuida principalmente a los poderes públicos, y la moderación, ejercida por las plataformas privadas mediante la selección, evaluación, clasificación, aprobación, ocultación o eliminación de contenidos. Esta situación pone en duda la tradicional consideración de las plataformas como meros intermediarios neutrales que se limitaban a alojar contenidos y, por ello, disfrutaban de regímenes de exención de responsabilidad (reflejado en la Sección 230 de la Decency Act estadounidense¹² o en el artículo 6 del Reglamento de Servicios Digitales¹³).

Su creciente intervención en la organización y visibilidad de la información aproxima su función, en determinados aspectos, a la de los editores clásicos, planteando nuevos problemas para la protección de la libertad de expresión en el entorno digital. Máxime cuando al final, pasado el primer filtro moderador de los algoritmos, la mayoría de las decisiones editoriales censoras son juicios personales de empleados particulares¹⁴.

En un entorno en el que la participación digital resulta prácticamente indispensable, las plataformas se han convertido en auténticos guardianes (*gatekeepers*) del discurso en línea, determinando mediante sus reglas, arquitectura y algoritmos qué contenidos pueden producirse, compartirse y, sobre todo, alcanzar visibilidad. Aunque no generen necesariamente contenidos propios, su ordenación, selección y recomendación constituye una verdadera labor de tratamiento de contenidos. Como señala Gillespie, «la moderación es, en muchos sentidos, el bien que ofrecen las plataformas», pues permite gestionar una abundancia informativa difícilmente navegable sin mecanismos de selección¹⁵. Es cierto que hay autores¹⁶ que se sorprenden, quizá no sin razón, de las críticas a la moderación y la selección, pues el sesgo es la esencia del negocio y el objetivo que perseguimos al usar estas plataformas (principalmente los buscadores y las redes sociales).

El diseño de las plataformas, sus normas comunitarias y sus sistemas algorítmicos condicionan tanto los límites del discurso permitido como su visibilidad¹⁷. Esta función, sin embargo, no responde necesariamente a la protección de la libertad de expresión, como advierte De Gregorio, pues las plataformas persiguen también la creación de entornos que incentiven la generación y circulación de datos susceptibles de alimentar sus modelos comerciales, particularmente los basados en la publicidad¹⁸.

Estas decisiones selectoras de las grandes plataformas se encuentran motivadas por una lógica de acumulación y extracción de datos en lo que Zuboff¹⁹ ha denominado capitalismo de vigilancia, y respecto al libre uso de la expresión representan un grave riesgo puesto que nos encontramos con empresas privadas ejerciendo un papel crucial en el espacio público. Según Balkin, las prácticas de moderación y regulación de contenidos conllevan riesgos tanto para la privacidad, esencial para la libertad de expresión, como para la dimensión pública, que implica considerar la libertad de expresión como necesaria para fomentar una cultura democrática²⁰.

Quizá sin ser conscientes, nuestra adicción a los contenidos que circulan por Internet, y especialmente a las redes sociales, están poniendo en cuestión uno

de los aspectos esenciales para la concordia, y no sólo la de carácter democrático. La existencia de un espacio público común donde poder intercambiar opiniones, experiencias, ideas, juicios, esperanzas, temores..., que es lo que nos permite conformar y mantener una comunidad, que no existe sin comunicación. La destrucción de nuestro modelo convivencial puede ser la más grave penitencia que tengamos que sufrir, sin redención, y peor aún, sin propósito de enmienda.

El espacio público discursivo constituye el lugar en el que nuestras democracias dejan de ser mera arquitectura institucional para convertirse en deliberación, contraste y construcción colectiva de la voluntad política. Sin embargo, hemos decidido aceptar que esté infectado de varias afecciones preocupantes que tienen a las redes sociales como principal patógeno. Las redes sociales pueden contribuir a la limitar el contraste político y social al favorecer posiciones radicales, la exposición selectiva y la formación de entornos informativos en los que los usuarios interactúan preferentemente con personas y contenidos próximos a sus propias posiciones, que Sunstein denomina «polarización grupal»²¹. Pariser acuñó el concepto de «filtros burbuja» (*filter bubble*) para describir cómo los sistemas de personalización algorítmica construyen para cada usuario un «universo único de información», seleccionado a partir de sus preferencias y comportamientos previos, reduciendo potencialmente su exposición a perspectivas discrepantes²². Esta intuición ha encontrado respaldo empírico. Bakshy, Messing y Adamic, tras analizar 10,1 millones de usuarios de Facebook, comprobaron que tanto los algoritmos como la estructura de las redes personales y las propias elecciones de los usuarios condicionan la exposición a información ideológicamente diversa²³. Otro estudio, a partir del análisis de más de cien millones de contenidos en Facebook, Twitter, Reddit y Gab, constataron la existencia de procesos de agrupación ideológica y formación de cámaras de eco, aunque con diferencias significativas entre plataformas²⁴. Incluso la exposición a posiciones discrepantes no garantiza necesariamente una reducción de las diferencias, pues se ha demostrado experimentalmente que, en determinadas circunstancias, el contacto con opiniones políticas contrarias en Twitter podía incrementar la polarización²⁵. En consecuencia, aunque las redes sociales no puedan considerarse causa única de este fenómeno, la combinación de personalización algorítmica, selección individual y homofilia social puede reforzar creencias previas, fragmentar el espacio público digital y reducir la existencia de un sustrato informativo común necesario para la deliberación democrática y la convivencia social. A estas dinámicas debe añadirse el denominado sesgo de confirmación (*confirmation bias*), entendido como la tendencia de los individuos a buscar, interpretar y valorar la información de manera favorable a sus creencias y expectativas previas²⁶.

Tales dinámicas de pensamiento grupal, fortalecidas en línea, son precisamente lo opuesto al ideal liberal de la esfera pública que la Red posibilita, donde constantemente nos enfrentamos a realidad incómodas, argumentos contrarios y valores diferentes a los nuestros y, por lo tanto, como sostenía John Stuart Mill en el magnífico capítulo 2 de «Sobre la libertad», se nos exige cuestionar, analizar y matizar nuestras propias convicciones. Cuando las personas se ven obligadas a escuchar a ambas partes, escribía Mill, siempre hay esperanza. Pero si se atiende a una de ellas, entonces los errores se convierten en prejuicios.

A estas dinámicas debe añadirse la creciente emocionalización de la comunicación política en las redes sociales. La arquitectura de las plataformas, orientada a maximizar la interacción y el tiempo de permanencia, puede favorecer aquellos contenidos capaces de provocar respuestas afectivas intensas, en la medida en que estas incrementan las posibilidades de compartir, comentar o reaccionar. La evidencia empírica muestra, en efecto, que la presencia de componentes emocionales puede resultar más determinante para la difusión de contenidos políticos que la propia calidad de los argumentos²⁷, mientras que el lenguaje asociado a la confrontación con grupos políticos rivales constituye un poderoso predictor del *engagement* en redes sociales²⁸. Particularmente relevantes parecen ser las emociones negativas como han mostrado diversos estudios, que han encontrado una asociación entre negatividad, ira o tristeza y mayores niveles de interacción y difusión de contenidos políticos²⁹. Estas dinámicas contribuyen a configurar un espacio público en el que la visibilidad de los mensajes depende crecientemente de su capacidad para provocar una reacción emocional, fenómeno que puede relacionarse con lo que Arias Maldonado denomina «democracia sentimental», caracterizada por la necesidad de reconsiderar la tradicional imagen del ciudadano racional a la luz del papel que desempeñan emociones y afectos en la formación de opiniones y comportamientos políticos³⁰.

Esa influencia ha tenido una manifestación concreta en la modulación, condicionamiento y delimitación sobre cómo ejercemos nuestra expresión en este nuevo contexto digital. La protección de la libertad de expresión en el Estado de Derecho contemporáneo se ha fundamentado por su aportación al libre debate de ideas en las sociedades democráticas con una clara distinción entre el Estado y la sociedad civil. Esta relación se ha puesto en valor por los principales tribunales tanto nacionales como internacionales. Desde el Tribunal Supremo de los Estados Unidos con su concepto del «Marketplace of ideas», al Tribunal Constitucional alemán en su sentencia *Lüth* de 1958, o la reiterada jurisprudencia del Tribunal Europeo de Derechos Humanos con su noción de los medios de comunicación

como «wachtdog». Sin embargo, en Internet esta distinción desaparece, pues está compuesto por espacios públicos de propiedad privada, regidos en gran medida por los incentivos comerciales de actores privados, antes que por una persecución del bien común. Las principales empresas tecnológicas dominan estratégicamente este espacio híbrido con escasa o nula presencia estatal, con una posición privilegiada como intermediarias «orquestradoras en la cadena de valor de la ecología digital». La dependencia que tenemos de ellas es abrumadora. Para acceder a un gran número de usuarios y conocer sus preferencias vitales tienes Facebook o Instagram, para vender productos a nivel masivo requieres de Amazon, para descargar aplicaciones has de pasar por Apple o Google, para buscar información necesitas los motores de búsqueda de Google o Microsoft.

La dependencia del control de la palabra que poseen estas grandes corporaciones las ha situado en una posición dominadora del discurso público y en un lugar central en las sociedades democráticas. La configuración de la libertad de expresión, cuando, cómo y quienes pueden ejercitarla supone una ruptura con los fundamentos del Estado de Derecho. Probablemente el espacio del discurso deliberativo es más amplio en teoría y más reducido en la práctica.

La capacidad de influir en la opinión pública nacional hacía a los medios de comunicación clásicos en objeto de deseo de los políticos, pero la capacidad de las redes sociales de influir a nivel mundial las convierte en instrumentos claves en el equilibrio de poder geoestratégico. La sustitución de los medios tradicionales (prensa, radio y televisión) por las redes sociales como forma de acceso a la información es imparable. Según el Digital News Report de 2025³¹ elaborado por la Reuters Institute, el número de ciudadanos estadounidenses que se informaban a través de las redes sociales ha pasado de un 27% en 2013 a un 54% en la actualidad, mientras que la prensa escrita como medio de referencia ha pasado en el mismo periodo de un 47% a un 14%. Si nos atenemos específicamente a redes sociales concretas, en países como India, un 55% de la población se informa a través de Youtube o un 49% de la población tailandesa emplea Tiktok como fuente de información.

Como indica Gillespie³², la esfera pública es la principal dimensión en la que los principios y valores inscritos en la arquitectura de las plataformas se manifiesta. Por un lado, el escándalo de Cambridge Analytica (empresa que utilizó los datos recopilados de Facebook como alimento para que algoritmos predictivos enviaran propaganda política dirigida a los votantes durante el referéndum sobre la salida del Reino Unido de la Unión Europea y las elecciones americanas de 2018), revela cómo la plataforma creó las condiciones para que los actores interfieran

en los resultados políticos a través de la operacionalización estratégica de datos residuales producidos por los usuarios. Por otra parte, la capacidad de las empresas de medios sociales excluye a los actores políticos de sus ecosistemas indican el alto poder que tales corporaciones tienen sobre las reglas del conflicto político contemporáneo. En este sentido, Gillespie³³ señala que, aunque las plataformas se presenten como meras representadoras de la esfera pública, de hecho, la constituyen a través de prácticas de moderación (expulsión, filtrado y suspensión), recomendación (*feed* de noticias, listas de popularidad, sugerencias de popularidad) y selección (contenido prominente y ofertas a las primeras páginas).

Ejercen poder sobre nosotros, nuestra autonomía y libertad al imponernos decisiones desde el diseño de su arquitectura, sustituyendo a los actores estatales que lo hacían desde la ley y las instituciones. Ejemplos como Facebook, que adopta una doble legitimación como ágora pública, donde se comercia, pero también se discute de los asuntos públicos, pero desde la exoneración de compromisos públicos, en el ideal del poder sin responsabilidad. Esa doble dimensión le permite entroncar la fusión de los datos de diversas plataformas con propósitos comerciales y con otros que sirven principalmente información política, deliberación pública e interconexión personal.

Los efectos sobre las débiles democracias de estos nuevos poderes de la palabra son innegables. La idea de soberanía popular sobre las que se asientan se ve traicionada por unos ciudadanos mediatizados por algoritmos secretos que condicionan a qué información accede cada ciudadano en función de sus intereses o tendencias, o peor aún de los intereses o proyectos de estos nuevos señores del aire como los denominó Javier Echeverría.

Las principales instituciones no son ajenas a esta problemática, y la Comisión Europea en su Comunicación sobre el Estado de la Década Digital de 2025 nos advierte de que «la creciente sofisticación de las amenazas a las que nos enfrentamos –incluidas la desinformación, los ciberataques, las falsificaciones profundas (*deepfakes*) y la manipulación algorítmica– exige una alfabetización digital generalizada y una fuerza laboral de TIC cualificada. *Estas amenazas, cuya creación y difusión se ven amplificadas por la IA y las plataformas en línea, también tienen el potencial de distorsionar los resultados electorales, profundizar la polarización social, erosionar la confianza pública en las instituciones democráticas y comprometer la infraestructura crítica. Asimismo, ponen en entredicho el Estado de derecho al socavar la integridad de los procesos democráticos, debilitar la rendición de cuentas institucional e interrumpir la aplicación de las normas legales en línea, especialmente cuando el contenido dañino o ilegal se difunde sin control a través*

de las plataformas. Surgen riesgos adicionales derivados de los complejos patrones de uso de algoritmos y herramientas digitales, como las plataformas en línea, y su impacto en la seguridad y el bienestar de los niños».

La fuerza de las palabras y el poder que despliegan en la esfera pública se hace más evidente en los mecanismos de desinformación empleados en la que se denomina guerra cognitiva. Como indican Romero y Nelson³⁴, las operaciones cognitivas trabajan sobre una lógica que actúa ante una crisis nacional donde la atribución de responsabilidades es incierta. Durante la brecha de verificación para fijar los hechos y su interpretación, moldean cómo se interpreta el daño y dificultan que la respuesta institucional llegue antes del marco narrativo que se haya fijado en dicha operación cognitiva. Estas operaciones de desinformación actúan cuando el daño es extremadamente difícil de medir y uno de los principales frentes de ataque es la propia percepción que se forma durante la brecha de verificación entre un suceso o fenómeno y su interpretación pública. La construcción de sentido como proceso mediado digitalmente, hace que estas operaciones se aprovechen de la ambigüedad reinante y que sus narrativas terminen por fijar cómo se recuerda el suceso. Para cuando alguien puede demostrar que tuvo lugar una operación coordinada de desinformación, el daño interpretativo ya está hecho³⁵. El agravante es que las operaciones del dominio cognitivo no terminan cuando cambia el ciclo informativo. El contenido que producen (artículos o videos sintéticos, narrativas coordinadas, citas fabricadas, fuentes manipuladas o inventadas) persisten en el entorno público. Se acumula en buscadores, archivos, bases de datos de referencia, y cada vez más, den los datos de entrenamiento de grandes modelos de lenguaje, motores de búsqueda y herramientas analíticas automatizadas. Es lo que algunos investigadores llaman envenenamiento de datos o *LLM Grooming*³⁶, consistente en inundar fuentes accesibles con contenidos fabricados que los sistemas de IA asimilan durante su entrenamiento o recuperación en tiempo real. Se trata de un cambio significativo respecto de la desinformación tradicional, pues el destinatario inmediato de la propaganda ya no tiene por qué ser el ciudadano, puede ser la propia máquina que posteriormente informará al ciudadano. Este riesgo se agrava con una asimetría estructural, pues las fuentes fiables restringen cada vez más el acceso a los rastreadores para proteger sus modelos de negocio, mientras los productores de contenidos sesgado y fabricado con intención manipuladora optimizan la máxima accesibilidad a las máquinas³⁷.

Como podemos ver, entramos en una nueva era en la tortuosa relación entre palabras y poder.

3. Dura lex, sed quae lex?

La titánica tarea de legislar sobre tecnologías, que están superando las previsiones de desarrollo³⁸ más optimistas, no debe hacernos caer en el desaliento normativo. Las singularidades de la regulación tecnológica dan lugar a teorías como la paradoja de Amara, que afirma que, en la regulación de una nueva tecnología, hay una tendencia a sobreestimar su efecto a corto plazo y a subestimar el efecto a largo plazo. Un problema similar se ha identificado como el dilema de Collingridge³⁹, que sostiene que es difícil predecir las consecuencias de una tecnología en su fase inicial, pero una vez que se descubren consecuencias indeseables, será difícil cambiar el camino regulatorio inicial elegido.

La reglamentación de estas nuevas tecnologías me recuerda los dibujos animados del Correcaminos y el Coyote, sólo que en vez de productos de ACME son reglamentos y directivas comunitarias.

El primer obstáculo para la regulación es la propia estructura del ecosistema digital formado por las plataformas de Internet. Las plataformas controlan cada vez más las puertas de entrada de todo el tráfico de Internet, la circulación de datos y la distribución de contenido (lo que hace que sociedades enteras dependan de sus sistemas), logrando además eludir el escrutinio regulatorio convencional. Las plataformas se alimentan de datos y se rigen por algoritmos, sin embargo, funcionan como parte de un ecosistema, como un conjunto de plataformas en red, gobernadas por un conjunto particular de mecanismos⁴⁰. Benjamin Bratton⁴¹ ha argumentado que las plataformas como las redes inteligentes, las nubes y las aplicaciones móviles evolucionan no como objetos separados sino como un aparato computacional con una nueva arquitectura de gobierno.

La arquitectura de capas de las plataformas se ha visualizado como una colección de «elementos apilados» o «capas» que reflejan características de modularidad y acumulación. Van Dijck nos explica, empleando la metáfora de un árbol⁴², como las plataformas intermedias constituyen el núcleo del poder de la plataforma, ya que median entre las infraestructuras y los usuarios individuales, así como entre las infraestructuras y los sectores sociales. Las «pilas» en este nivel incluyen servicios de identificación o inicio de sesión (FB ID, Google ID, Amazon ID, Apple ID), sistemas de pago (Apple Pay, Google Pay), servicios de correo y mensajería (FB Messenger, Google Mail, MS Mail, Skype, FaceTime), redes sociales (Facebook, Instagram, WhatsApp, YouTube), motores de búsqueda (Google, Bing), servicios de publicidad (FB Ads, Google), redes minoristas (Amazon Marketplace, Prime) y tiendas de aplicaciones (Google Play, Apple). Esta lista no es exhaustiva ni estática.

Ninguna de estas plataformas intermediarias resulta esencial para todas las actividades que se desarrollan en Internet, pero, consideradas conjuntamente, obtienen buena parte de su poder de su posición como pasarelas centrales de información y servicios, ocupando distintas capas estratégicas del ecosistema digital y canalizando a través de ellas importantes flujos de datos. Como explica Van Dijck⁴³, estas plataformas intermediarias se caracterizan tanto por la fuerte presencia de las grandes compañías tecnológicas como por su elevada interdependencia, de modo que el ecosistema se estructura simultáneamente mediante relaciones de competencia y coordinación. En este sentido, pueden actuar, en expresión de Mansell, como «orquestadores en la cadena de valor de la ecología digital»⁴⁴. Meta constituye una de las principales vías de acceso a audiencias masivas; Amazon ocupa una posición central en el comercio electrónico; Apple y Google controlan dos de los principales canales de distribución de aplicaciones –aunque el Digital Markets Act ha introducido en la Unión Europea vías alternativas de distribución–; y Google mantiene una posición especialmente relevante en el acceso a la información mediante los motores de búsqueda. Al mismo tiempo, estas grandes empresas mantienen importantes relaciones de interdependencia tecnológica y comercial, ya que los servicios de unas utilizan infraestructuras o canales de distribución controlados por otras. Esta combinación de competencia e interdependencia da lugar a lo que Van Dijck denomina «coordinating competitors», una forma de *coopetition* que dificulta analizar el poder de las plataformas exclusivamente desde la perspectiva de empresas o mercados considerados aisladamente⁴⁵. De ahí que, como sostiene Owen⁴⁶, los desafíos derivados del poder de las plataformas sean de carácter sistémico y requieran respuestas regulatorias igualmente integrales. Así, la regulación sectorial que está abordando, por ejemplo, la Unión Europea, puede no ser lo suficientemente efectiva precisamente por su aproximación sectorial, requiriendo un acercamiento más holístico a la problemática regulación de las plataformas.

Las empresas tecnológicas deliberadamente empujan sus plataformas para oscilar entre sectores e infraestructura, entre sectores comerciales y no comerciales, entre intereses públicos y privados, entre un mercado para bienes y servicios, y un mercado de ideas, mientras que adoptan características de ambas. Aún más, ejercen un poder sin precedentes sobre la vida de las personas, afectando su autonomía y libertad al imponer las decisiones de los usuarios a partir del diseño de arquitectura, este poder que previamente había sido asignado a los actores estatales a cargo de moldear las leyes e instituciones gubernamentales.

Podríamos, por tanto, definir las plataformas digitales, siguiendo a Poell, Nieborg y van Dijck⁴⁷, como infraestructuras digitales reprogramables que facilitan y, al mismo tiempo, configuran las interacciones entre usuarios finales y proveedores complementarios mediante procesos sistemáticos de recopilación, procesamiento algorítmico, monetización y circulación de datos. Esta concepción permite superar una visión de las plataformas como simples intermediarios tecnológicos para poner de relieve su capacidad para estructurar las relaciones que tienen lugar en su interior y condicionar las reglas bajo las que interactúan los diferentes actores del ecosistema digital.

Eso nos conduce a otro de los elementos que refuerza el poder de las grandes plataformas: su tamaño. Lo que quizá debería hacernos redefinir el concepto de monopolio natural utilizado por los expertos en política de la competencia. La magnitud del poder de estas plataformas resulta evidente si se atiende a los datos disponibles. Según Statista, Meta –que agrupa Facebook, Instagram, Messenger, Threads y WhatsApp– cuenta con aproximadamente 3.580 millones de usuarios diarios. Google procesa alrededor de 11 millones de búsquedas por minuto, lo que da cuenta de su posición central en el acceso a la información. Su plataforma YouTube alcanza unos 2.580 millones de usuarios⁴⁸. En conjunto, las redes sociales suman 5.660 millones de usuarios, lo que representa el 68,7 % de la población mundial. Por su parte, los usuarios de Internet alcanzan los 6.040 millones, es decir, el 73,2 % de la población global. ¿Ante este potencial, quien puede competir?⁴⁹

Sigamos con los obstáculos. La idea de soberanía estatal clásica vinculada a la imposición de normas en un determinado territorio se desvanece cuando no hay territorio, produciéndose un vaciamiento de la soberanía como consecuencia de la globalización digital, como ya nos advertía Habermas. La regulación de las plataformas de Internet, sus contenidos, su estructura o comportamientos no puede ser ya asunto de un solo gobierno nacional en uso de su soberanía⁵⁰. Lo expresaban Schmidt y Cohen al decir que «El mundo digital no está sujeto realmente a las leyes terrenales [...] Es el espacio sin gobierno más grande del mundo»⁵¹. Si una persona vive en Seúl, su libertad de palabra efectiva será el producto de las condiciones y normas legales que se dan en su país real, Corea del Sur, pero también en las condiciones de países virtuales llamados Facebook; Google o X; y las otras plataformas, editoriales, emisoras de radio y televisión, y universidades, etc., con las que interactúa en el lugar donde vive. Hemos transformado el principio «*cuius regio, eius religio*» por «*cuius regio, eius Internet*».

4. El nuevo telón de silicio⁵²

El poder clásico ha respondido al ascenso del poder digital intentando cercarlo, regularlo y someterlo a límites. Pero en una visión arcaica, han intentado constreñirlos a los límites territoriales de los Estados.

Si Estados Unidos, China y Rusia pretender ser los nuevos imperios del siglo xxi, el ejercicio del poder imperial resulta hoy inconcebible sin el control de los resortes digitales. El imperio ha estado históricamente vinculado a la idea de soberanía entendida como poder absoluto, excluyente y exclusivo. En palabras de Hermann Heller, la soberanía es la capacidad, tanto jurídica como real, de decidir de manera definitiva y eficaz, incluso contra el derecho positivo, e imponer esa decisión. Sin embargo, la concepción clásica de la soberanía estatal, asociada a la imposición de normas dentro de un territorio delimitado, se diluye en un entorno digital caracterizado por la desterritorialización. Aunque para muchos la idea de ejercer soberanía sobre el ciberespacio constituye un auténtico oxímoron, los casos de China y Rusia evidencian que la lucha por el control de la esfera digital resulta decisiva para preservar formas clásicas de soberanía territorial. Frente al modelo de gobernanza *multi-stakeholder* o de múltiples partes interesadas, promovido por Estados Unidos, el G7 y la Unión Europea, emerge así la reivindicación del principio de soberanía cibernética⁵³, basada en la construcción de espacios digitales propios, diferenciados y relativamente separados de la red global⁵⁴. Así, desde una perspectiva geopolítica, la plataformización se manifiesta empíricamente en la división del ciberespacio en dos ecosistemas de plataformas políticamente opuestos. Por un lado, se encuentra el ecosistema de plataformas gestionado por los cinco grandes conglomerados tecnológicos estadounidenses; Google/Alphabet, Amazon, Facebook/Meta, Apple y Microsoft (GAFAM), que abarca la mayor parte del sistema internacional. Por otro lado, encontramos la excepción del ecosistema nacional de plataformas chino, en el que los conglomerados Baidu, Alibaba y Tencent (BAT), bajo la estricta vigilancia del Partido Comunista Chino (PCCh), desempeñan funciones troncales similares en la sociedad china.

La idea de neutralidad de la red ha sido puesta en solfa cuando Estados como China o Rusia han respondido reivindicando el concepto de soberanía cibernética, construyendo espacios digitales propios mediante políticas de control y aislamiento. El modelo chino de soberanía digital se articula a través del denominado «Gran Cortafuegos», un sistema de regulación, filtrado y censura apoyado en una infraestructura altamente centralizada que permite controlar los flujos de información y aislar el ciberespacio nacional de influencias externas. Este modelo responde a una concepción en la que Estado y mercado actúan de forma

coordinada bajo la dirección del PCCh, lo que implica un control institucional directo sobre las plataformas digitales. En este contexto, plataformas como iQiyi reproducen un modelo próximo al de la televisión tradicional, orientado a ofrecer contenidos homogéneos y a reforzar una experiencia mediática compartida, en contraste con la fragmentación característica de los entornos digitales occidentales basados en la personalización algorítmica. De este modo, el ecosistema digital chino no constituye un espacio autónomo, sino una prolongación del sistema comunicativo estatal.

Por su parte, Rusia consolidó su estrategia de soberanía digital con la aprobación en 2019 de la Ley de Soberanía de Internet, destinada a reforzar el control estatal sobre la red y garantizar la autonomía operativa del país frente a posibles aislamientos externos. La norma faculta a las autoridades para restringir el tráfico de datos e incluso desconectar la red rusa del Internet global en situaciones de emergencia, además de obligar a los proveedores a incorporar sistemas de filtrado y vigilancia. Esta política se complementa con el desarrollo de infraestructuras tecnológicas propias, como un DNS nacional, dentro de una estrategia más amplia de autonomía digital y control soberano del ciberespacio.

Esta disputa estratégica-digital por la palabra tiene su reflejo en la pugna empresarial. Según el Informe sobre las Inversiones en el Mundo 2025 de Naciones Unidas⁵⁵, de las diez principales empresas de la economía digital por volumen de negocio, seis son estadounidenses, ocupando además las cuatro primeras posiciones Amazon, Apple, Alphabet y Microsoft; mientras que las cuatro restantes son chinas. No obstante, la evolución de la economía digital en la última década revela un progresivo ascenso de China en este ámbito. Así, en 2014, entre las cien principales empresas tecnológicas del mundo, 67 eran estadounidenses y únicamente 2 chinas; en 2024, las compañías estadounidenses se redujeron a 56, mientras que las chinas aumentaron hasta 11. Aunque la diferencia continúa siendo significativa, la tendencia evidencia un creciente desplazamiento del equilibrio digital global hacia China.

En este nuevo contexto mundial, las nuevas aproximaciones al fenómeno imperial tienden a recuperar una lógica de territorialización del poder. Frente a una idea ampliamente extendida según la cual la globalización habría conducido a un poder completamente «virtual» o desterritorializado, autores como Saskia Sassen⁵⁶ cuestionan esta interpretación, confirmando este nuevo escenario geoestratégico. Para Sassen, el capitalismo global no desaparece en lo abstracto, sino que se articula a través de la organización de flujos transnacionales –financieros, informacionales, logísticos y de movilidad– que requieren infraestructuras

concretas y formas específicas de anclaje espacial. En este marco, las plataformas digitales pueden entenderse como máquinas de organización de flujos. Amazon estructura circuitos globales de mercancías y logística; Meta organiza la atención y la comunicación; Google gestiona flujos masivos de información y datos. Su poder no reside tanto en el control de territorios tradicionales como en el dominio de los circuitos de circulación global.

Quizá nos encaminemos hacia un nuevo telón que vuelva a dividir el mundo. Pero esta vez no será un telón de acero, sino de silicio. Ya no separará un primer, un segundo y un tercer mundo, sino un primer mundo tecnológico, dominado por un reducido número de Estados y grandes corporaciones capaces de controlar los datos, los algoritmos y las infraestructuras digitales, y un resto del mundo condenado a depender de tecnologías que no posee, no controla y cuyas reglas tampoco decide.

5. De los tribunales al algoritmo: la privatización de los derechos

Otro de los grandes cambios que ha traído Internet es que buena parte de las reglas que condicionan nuestra libertad de expresión ya no proceden directamente del Estado. En las redes sociales, los Términos de Uso y las Condiciones de Servicio de las plataformas funcionan, en la práctica, como auténticas normas privadas, pues determinan qué podemos publicar, qué contenidos se eliminan e incluso quién puede permanecer en una determinada plataforma. Son la forma en que las grandes plataformas simulan ejercer una suerte de virtud digital mediante procedimientos opacos que nadie conoce plenamente y que, paradójicamente, terminan mostrando al mundo nuestros pecados digitales –una expresión desafortunada, un comentario xenófobo, un tuit jocoso– al señalarlos mediante la eliminación del contenido o, peor aún, mediante la suspensión o supresión de nuestra cuenta.

El problema es que esas reglas suelen formularse de manera muy amplia. Esto hace que, en ocasiones, las plataformas eliminen no solo contenidos claramente ilícitos –amenazas, incitación a la violencia o discursos de odio–, sino también expresiones perfectamente legítimas que pueden resultar incómodas u ofensivas para determinados sectores⁵⁷. A ello se une, como ha señalado el Relator Especial de Naciones Unidas sobre libertad de expresión, la falta de procedimientos adecuados de recurso y, en muchas ocasiones, la ausencia de explicaciones

suficientemente claras sobre por qué se elimina un contenido o se suspende una cuenta (Informe A/HRC/32/38, 2016, p. 16).

En definitiva, hemos delegado en empresas privadas decisiones que afectan directamente al ejercicio de derechos fundamentales. Y esto plantea una cuestión importante: ¿queremos que una empresa privada decida dónde termina nuestra libertad de expresión? La propia Comisión Interamericana de Derechos Humanos ha advertido que los actores privados no disponen necesariamente de la capacidad necesaria para ponderar derechos e interpretar las normas conforme a los estándares internacionales de derechos humanos⁵⁸.

Un buen ejemplo es el derecho al olvido, anteriormente mencionado. A partir de la sentencia del Tribunal de Justicia de la Unión Europea de 13 de mayo de 2014, asunto C-131/12, posteriormente reflejada en el artículo 17 del RGPD, los buscadores tienen que valorar determinadas solicitudes de retirada de información. En la práctica, esto significa que una empresa como Google debe decidir si prevalece el derecho a la protección de datos o la libertad de información. Es decir, realiza una verdadera ponderación entre derechos, una función tradicionalmente asociada a los tribunales. Como advertía Frank La Rue, las decisiones de censura no deberían delegarse en entidades privadas, porque inevitablemente implican complejos juicios de valor y difíciles ponderaciones entre derechos e intereses (Informe A/HRC/32/38, 2016, pp. 16-17). Pero debemos ser conscientes de que ni el algoritmo más complejo del mundo puede emitir un juicio específico sobre el momento, el lugar, la forma y el contexto en que se produce una colisión de derechos. Emily Taylor ha descrito gráficamente este fenómeno como la «privatización de los derechos humanos»⁵⁹.

Esta privatización se explica también por una realidad muy sencilla. Los Estados han sido incapaces de regular por sí solos la enorme complejidad de Internet. Como señalaba Morozov, las grandes compañías han desarrollado sus propias normas y prácticas para filtrar y eliminar contenidos⁶⁰. Y su posición es especialmente poderosa porque quien quiera participar realmente en la conversación pública digital difícilmente puede prescindir de Google, YouTube, Facebook o las grandes redes sociales. Puede utilizar alternativas, naturalmente, pero tendrá muchas más dificultades para alcanzar una audiencia significativa.

Esto afecta directamente al pluralismo. No basta con que existan formalmente muchas opiniones; para que haya verdadero pluralismo esas opiniones tienen que encontrarse, confrontarse y dialogar. Si cada grupo permanece encerrado en su propio espacio informativo, tenemos pluralismo formal, pero no necesariamente

pluralismo real. Y la concentración de buena parte de la conversación pública en unas pocas plataformas aumenta además su capacidad –y también la de quienes consiguen utilizarlas estratégicamente– para seleccionar, jerarquizar y eventualmente manipular contenidos. El caso de Cambridge Analytica mostró de manera especialmente visible los riesgos políticos de esta concentración de información y capacidad tecnológica.

Hay además otro elemento fundamental en todo contrato, el consentimiento. Formalmente aceptamos las condiciones de las plataformas, pero cabe preguntarse hasta qué punto ese consentimiento es verdaderamente libre e informado. Emily Taylor habla precisamente de una «ilusión del consentimiento», aceptamos condiciones sobre privacidad y tratamiento de datos que normalmente no hemos leído y que, aunque las leyéramos, muchas veces nos resultarían difíciles de comprender.

El dato que ofrecen McDonald y Cranor es muy ilustrativo, quienes calcularon que, si todos los usuarios estadounidenses de Internet leyesen íntegramente las políticas de privacidad de cada nuevo sitio que visitan, dedicarían aproximadamente 54.000 millones de horas al año a hacerlo⁶¹. Es evidente, por tanto, que el modelo basado simplemente en «usted ha aceptado» presenta importantes limitaciones. Brown y Marsden añaden que las decisiones sobre privacidad dependen enormemente del contexto, de la confianza que tengamos en quien solicita nuestros datos y, sobre todo, de nuestras propias suposiciones –muchas veces equivocadas– acerca de cómo van a utilizarse⁶².

Por eso, el problema de fondo es que seguimos utilizando categorías jurídicas tradicionales para una realidad tecnológica completamente distinta. Las plataformas han reducido las barreras para expresarnos y acceder al conocimiento, pero, al mismo tiempo, la complejidad del mercado de datos hace cada vez más difícil sostener que los modelos tradicionales de consentimiento sean suficientes⁶³.

Y aquí aparece, a mi juicio, una de las grandes paradojas de la sociedad digital: nunca habíamos tenido tantas posibilidades de expresarnos y, al mismo tiempo, nunca unas pocas empresas privadas habían tenido tanta capacidad para determinar las condiciones en las que esa expresión se produce. El problema ya no consiste únicamente en quién puede hablar, sino también en quién establece las reglas, quién decide qué vemos, qué contenidos adquieren visibilidad y, finalmente, quién controla la arquitectura de nuestra conversación pública.

Sin embargo, para que haya una conversación pública, es fundamental que sus participantes tengan la posibilidad de adquirir conocimientos sobre cuestiones

que nunca han experimentado de primera mano. En un Estado de gran envergadura, son el sistema educativo y los medios de comunicación quienes desempeñan un papel crucial de informar a las personas acerca de las cosas a las que nunca se han enfrentado. Sin sistema educativo ni plataformas mediáticas que ejerzan este papel, no pueden darse conversaciones públicas fluidas ni productivas a gran escala. En definitiva, su misión es hacernos pasar de la episteme a la doxa.

Para hacer frente a esta realidad, la Unión Europea ha establecido a través del Reglamento de Servicios Digitales unas obligaciones de transparencia a las plataformas, con la finalidad de que hagan públicas los mecanismos de selección, moderación, postergación y moderación de contenidos, cuya utilidad dependerá en gran medida de la voluntad de los usuarios para exigirlos⁶⁴. Pero sobre ello volveremos más adelante.

6. Las mutaciones del poder

Las reflexiones sobre el poder han evolucionado al compás de la creciente complejidad de las sociedades contemporáneas y pueden resultar especialmente útiles para comprender el nuevo tipo de poder que ejercen las plataformas de Internet. Un poder que ha mutado, adquiriendo capacidades hasta hace poco insospechadas. Como en el universo de los X-Men, sabemos que estamos ante una nueva forma de poder, pero todavía no sabemos del todo qué rostro acabará adoptando. Si el de Charles Xavier, orientado a la convivencia y la cooperación, o el de Magneto, convencido de que quien posee capacidades superiores terminará utilizándolas para imponerse.

Entre las nuevas teorías sobre el poder encontramos la noción de biopolítica de Foucault, en la que el poder no se limita a imponer leyes o controlar territorios; se orienta a producir y gestionar la vida social en su conjunto, organizando la comunicación (medios, redes), estructurando la economía del conocimiento y modelando deseos, identidades y formas de vida. Zuboff ha explicado en su obra «La era del capitalismo de vigilancia» como las grandes empresas tecnológicas reclaman unilateralmente la experiencia humana como materia prima gratuita para traducirla en datos conductuales, donde ya no basta con automatizar los flujos de información sobre nosotros. El objetivo ahora es automatizarnos a nosotros. Los *arcana imperii* están ahora en poder de empresas privadas y sus zonas grises, no en las manos de los Estados y sus zonas oscuras.

Las mutaciones del poder fueron puestas de manifiesto en 1974 por Steven Lukes en su obra *Power: A Radical View*⁶⁵, donde plantea tres dimensiones del

poder: 1. Poder como decisión. Es su dimensión más visible, A consigue que B haga algo que de otro modo no haría. Existe un conflicto observable y puede identificarse quién prevalece. 2. Poder sobre la agenda. El poder consiste también en determinar qué asuntos pueden llegar a discutirse y cuáles quedan excluidos. Quien controla la agenda puede ejercer poder sin necesidad de ganar una decisión concreta, porque determinados conflictos ni siquiera llegan a plantearse. 3. Poder sobre preferencias y percepciones. Es su aportación más característica. El poder puede operar moldeando deseos, creencias, preferencias y percepciones, de modo que las personas acepten como natural, inevitable o incluso beneficiosa una situación contraria a sus intereses.

Por su parte, para Joseph Nye⁶⁶, el poder puede entenderse, en términos generales, como la capacidad de influir en el comportamiento de otros para obtener los resultados que se desean. Especialmente destacable es su distinción entre *Hard power* (poder duro) como la capacidad de conseguir que otros actúen de determinada manera mediante coerción o incentivos, mediante el poder militar y económico; y *Soft power* (poder blando), como la capacidad de obtener los resultados deseados mediante la atracción y la persuasión, en lugar de la coerción o el pago. Nye lo vincula especialmente con la cultura, los valores políticos y la legitimidad de las políticas. Incluye también, el *Smart power* (poder inteligente) que es una combinación estratégica y contextual de recursos de *hard* y *soft power*. No constituye simplemente una suma de ambos, sino la capacidad de determinar qué instrumentos utilizar y en qué circunstancias.

Pero probablemente es Byung-Chul Han, quien mejor caracteriza esta nueva forma de poder digital. Para el filósofo coreano afincado en Berlín, el poder no debe entenderse únicamente como una relación de coerción en la que una voluntad se impone sobre otra, sino como una capacidad de orientar la conducta ajena sin necesidad de recurrir a la violencia o a la prohibición explícita. En «*Sobre el poder*»⁶⁷, Han sostiene que cuanto más poderoso es el poder, más silenciosa e imperceptiblemente opera, hasta el punto de que el sujeto puede experimentar como propia una voluntad que ha sido condicionada externamente. Esta concepción adquiere especial relevancia en «*Psicopolítica*»⁶⁸, donde sostiene que el poder característico del neoliberalismo digital no actúa prioritariamente mediante la disciplina y la restricción de la libertad, sino que se sirve de la propia libertad del individuo, interviniendo sobre sus deseos, emociones, motivaciones y decisiones. La acumulación masiva de datos y las tecnologías digitales permiten así conocer, predecir y orientar comportamientos sin que el sujeto perciba necesariamente una intervención externa. El poder alcanza, por tanto, su máxima eficacia cuando

se interioriza y consigue que el individuo actúe voluntariamente conforme a sus mecanismos de dominación, transformando la libertad, paradójicamente, en un instrumento del propio poder.

Así, la capacidad de las plataformas digitales para ejercer poder en la sombra, conlleva la capacidad de obtener un nivel de información del individuo que permite modelar su comportamiento hasta extremos insospechados. Zubbof sintetiza muy bien lo que considera una nueva manifestación de esta lógica de dominio invisible. El poder instrumental, sustentado en la apropiación masiva de datos conductuales y en sistemas algorítmicos capaces no solo de conocer y predecir el comportamiento, sino también de intervenir sobre él. El capitalismo de la vigilancia «reclama unilateralmente para sí la experiencia humana» como materia prima susceptible de convertirse en datos, generando una profunda asimetría epistemológica entre quienes son observados y quienes disponen de los medios tecnológicos para conocerlos. De ahí que las cuestiones esenciales del nuevo poder digital puedan condensarse, para Zuboff, en tres preguntas: «¿Quién sabe? ¿Quién decide? ¿Quién decide quién decide?»⁶⁹.

Las plataformas digitales pueden, por tanto, ser manifestaciones de estas nuevas formas de poder, pues determinan la agenda y la visibilidad de los contenidos (Lukes), influyen mediante mecanismos de atracción y persuasión (Nye), actúan sobre deseos y comportamientos aprovechando la propia libertad del individuo (Han) y transforman el conocimiento derivado de los datos en capacidad de predicción e intervención conductual (Zuboff). El cambio de paradigma lo expresó certeramente la Comisaria europea de Competencia, Margrethe Vestager, al decir que «del ciudadano como lo importante y los datos que dejaba atrás como algo secundario a un movimiento inverso donde el ciudadano va a menor y el rastro que dejan los datos es lo verdaderamente importante».

Como describe Hasselbalch⁷⁰, la historia de Internet, el Big Data y ahora la IA se ha desarrollado en tres fases: la primera implicó la construcción de una infraestructura global interoperable de Internet e información; la segunda supuso una reacción a sus consecuencias sociales incidiendo en gobernar los riesgos y las implicaciones éticas (lugar donde aún estamos en la UE, y menos es nada). Pero considera que estamos en una tercera fase, ante algo más profundo que puede amenazar la identidad misma de la humanidad a medida que la IA tiene poder para moldear los valores y las prioridades, y su desarrollo y control se concentra en unas pocas manos. Para la autora danesa, las discusiones actuales sobre la IA están mal enfocadas, pues se bifurcan como una cuestión del ritmo y límite de la innovación tecnológica, o como una reflexión sobre el destino de la humanidad.

Para ella, ninguna de las dos aborda el verdadero problema que es quién dirige las narrativas culturales, quien decide los valores y el rumbo del desarrollo de la IA, quién posee y produce la infraestructura de la IA, quién es más dependiente, quien obtiene más poder y quien pierde capacidad de decisión.

PARTE TERCERA

¿Dónde está la sabiduría que hemos perdido en el conocimiento?
¿Dónde está el conocimiento que hemos perdido en la información?
(T.S. Eliot)

1. ¿De qué hablamos cuando hablamos de Inteligencia Artificial?

Durante los años 30, Alan Turing realizó investigaciones que se concretaron en la publicación en la revista *Mind* en 1959 de un artículo titulado «Computing machinery and intelligence», donde se preguntaba si las máquinas podían pensar y establecía un método para determinar si una máquina puede pensar, que se popularizó como el «test de Turing» y que aún hoy en día es empleado y sigue siendo objeto de una especial atención por la comunidad académica. Pero fueron John McCarthy, Marvin Minsky y Claude Shannon, quienes acuñaron formalmente el término Inteligencia Artificial durante una conferencia en Dartmouth College (1956) en donde plantearon «la ciencia e ingenio de hacer máquinas inteligentes, especialmente programas de cálculo inteligente». La actividad fue financiada por la Fundación Rockefeller y se llamó *Dartmouth Summer Research Project on Artificial Intelligence*¹. Esta aproximación a la IA se ha señalado como una rama emergente de investigación ligada a la fisiología del cerebro, análisis formal de la lógica proposicional y las ciencias de la computación². Por su parte, Valle-Cruz et al.³ indican que este ámbito de las ciencias se orienta a la creación de algoritmos que nutren máquinas inteligentes, simulando el comportamiento

individual y colectivo, no sólo de los seres humanos, sino también de otros seres vivos, automatizando actividades, aprendiendo y evolucionando, así como mejorando el rendimiento y los resultados obtenidos con y sin ayuda de los seres humanos.

Pero rememorando el precioso libro de relatos de Raymond Carver, ¿de qué hablamos cuando hablamos de IA? Se trata de una nueva realidad sobre la que existe casi unanimidad científica respecto a que supondrá un cambio hasta ahora desconocido en la historia de la humanidad. Pero la definición es compleja porque no hay una única IA⁴.

La inteligencia artificial comprende una amplia gama de métodos, algoritmos y tecnologías, siendo habitual la diferencia entre IA débil, IA fuerte y Super IA⁵. Pero la IA es un término genérico o supraconcepto que engloba diversas tecnologías inteligentes con importantes aspectos comunes y estrechamente interrelacionados entre sí, aunque algunas de ellas estén adquiriendo sustantividad propia como disciplinas científicas. Desde un punto de vista jurídico, tomando como referencia los documentos de trabajo, informes, comunicaciones y propuestas de la UE relativas a la definición de un marco ético y jurídico europeo de IA, ésta también puede clasificarse en dos tipos: a) La IA no integrada «basada en software», como asistentes virtuales, software de análisis de imágenes, motores de búsqueda o sistemas de reconocimiento de voz y rostro. b) La IA integrada en máquinas u otros productos, como robots, drones, vehículos autónomos o Internet de las Cosas –IoT por sus siglas en inglés–⁶, que tiene especial importancia en la regulación comunitaria de la UE. Gallego Sánchez diferencia todavía más las mismas, considerando que la inteligencia artificial débil –o *narrow AI*– sería capaz de llevar a cabo tareas predefinidas, la general o *general AI* tareas intelectuales similares a las humanas y la súper inteligencia o super AI sería la que superaría a la inteligencia humana en todos los aspectos⁷.

Si queremos una definición de IA, podemos encontrar casi una por institución, investigador, *think tank* o gobierno que la ha abordado⁸. Pero por ser operativos, utilizaremos la definición de la Unión Europea en su no tan ya reciente Reglamento de Inteligencia Artificial⁹, que la conceptualiza en su artículo 3¹⁰ como «un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales»¹¹.

No es este el lugar, ni constituye el objeto del presente trabajo, realizar una revisión sistemática de las distintas definiciones de inteligencia artificial¹². Sí resulta, en cambio, necesario constatar que la incidencia de la IA y de los sistemas algorítmicos sobre los ordenamientos jurídicos es ya una realidad transversal e innegable, hasta el punto de que prácticamente todas las ramas del Derecho se ven obligadas a reconsiderar algunas de sus categorías tradicionales ante los desafíos que plantea esta transformación tecnológica.

Así ocurre, en primer término, en el Derecho civil, donde adquiere especial relevancia la reflexión sobre los regímenes de responsabilidad contractual y extracontractual derivados de la utilización de sistemas de IA¹³, particularmente tras la aprobación de la Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por los daños causados por productos defectuosos, que deroga la Directiva 85/374/CEE del Consejo¹⁴. También el Derecho penal se enfrenta a importantes interrogantes, tanto desde la perspectiva de la eventual responsabilidad vinculada al empleo de estos sistemas como desde su utilización como instrumentos para la prevención y persecución de la delincuencia¹⁵. En el ámbito laboral, la irrupción de la gestión algorítmica ha suscitado igualmente un intenso debate, como pone de manifiesto la Directiva (UE) 2024/2831 relativa a la mejora de las condiciones laborales en el trabajo en plataformas¹⁶, así como, en el ordenamiento español, el artículo 64.4.d) del Estatuto de los Trabajadores.

La dimensión antidiscriminatoria de estas tecnologías encuentra asimismo reflejo en el artículo 23, bajo la rúbrica «Inteligencia Artificial y mecanismos de toma de decisión», de la Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación¹⁷. Pero es probablemente en el Derecho administrativo donde el empleo de sistemas algorítmicos y de inteligencia artificial ha recibido hasta el momento una atención doctrinal más intensa, dada su creciente incorporación a la actuación y a los procesos decisorios de las Administraciones públicas¹⁸. El análisis de los problemas principales que plantea la incorporación de la IA al sector público excede, sin embargo, tanto del objeto como del enfoque específico de este trabajo¹⁹, y constituye además una cuestión que hemos tenido ocasión de abordar previamente en otro lugar²⁰.

2. Los nuevos dueños de las palabras

Pero la reflexión que queremos desarrollar pretende desbordar las fronteras (no sé si esta expresión no es de mal gusto precisamente ahora) del Derecho para realizar ciertas consideraciones sobre los efectos e influencia de la IA como nuevo poder y evitar así la máxima de Orwell de que la «ignorancia es fuerza». El 1 y 2 de noviembre de 2023, Estados Unidos, China, la Unión Europea y otros veintiséis Estados suscribieron *The Bletchley Declaration*²¹, con el propósito de fortalecer la cooperación internacional en materia de inteligencia artificial y de contener sus riesgos potencialmente catastróficos. El acuerdo enfatiza la necesidad de abordar, en foros e iniciativas existentes, cuestiones relativas a la protección de los derechos humanos, la transparencia y explicabilidad, la equidad, la rendición de cuentas, la seguridad, la supervisión humana adecuada, la ética, la mitigación de sesgos, así como la privacidad y la protección de datos²². Pero también advierte «de la posibilidad de un daño grave, incluso catastrófico, ya sea deliberado o no intencionado, que surge de las capacidades más relevantes de estos modelos de IA».

Cabe llamar la atención sobre el matiz de que el daño puede ser intencionado o no. Las palabras son actos volitivos, pero las palabras que formulan los sistemas de IA simulan tener voluntad propia sin albergar intenciones ni remordimientos. Harari considera que los cuchillos y las bombas no deciden por sí mismas, sino que son instrumentos sin criterio careciendo de la inteligencia para procesar información y tomar decisiones independientes. No obstante, la IA puede procesar información por sí sola y, por lo tanto sustituir a los humanos en la toma de decisiones. Y concluye advirtiéndonos de que «la IA no es una herramienta sino un agente»²³. Como explica Véliz, «cada elemento tecnológico es un artefacto, lo que significa que ha sido diseñado por alguien para hacer algo, y ese hecho por sí solo lo despoja de neutralidad»²⁴.

Como relata Fernández de Lis²⁵, varios hechos alarmantes nos deberían hacer reflexionar sobre esta idea. Las principales compañías de IA han revelado el pasado verano que habían descubierto a sus modelos de IA haciendo cosas que no les habían pedido, entrando en sistemas de otras empresas, organizándose con otras inteligencias artificiales en enjambre, delatando a otras IA y mintiendo para tapar sus huellas.

En mayo pasado, un grupo de modelos de *OpenAI* empezó a coordinarse por su cuenta durante semanas en un foro secreto de mensajes que ellos mismo habían creado para averiguar cómo salir de su entorno de pruebas cerrado al exterior

y acceder a Internet. La compañía detectó el intento y cerró dicho entorno de pruebas. Pero *OpenAI* reveló posteriormente que los modelos reconstruyeron el foro y volvieron a escaparse al «mundo real» a finales de junio. El epílogo fue el ataque a *Hugging Face*, la mayor plataforma del mundo para compartir modelos de IA, quien en su análisis de la intrusión destacó el volumen del ataque como hecho determinante. Mediante la ejecución de 17.600 acciones, el ataque probó rutas sin salida, evadió bloqueos cambiando de canal y regresó repetidamente sobre sus pasos para avanzar. Los agentes de IA de *OpenAI* llegaron a mostrarse paranoicos entre sí, sospechando que otro agente de IA intentaban engañarlos maliciosamente, y se autoorganizaron como un «enjambre». Sam Altman dijo que es el primer incidente de seguridad que le ha revuelto las tripas, y debe haber visto unos cuantos, y tener un estomago a prueba de pruebas de sistemas de IA.

Unos días después, el 30 de julio, fue *Anthropic* quien debido al incidente alertado por el ataque de *OpenAI* a *Hugging Face* fue consciente de la incidencia sufrida. En este caso, un agente de IA de *Anthropic* se quedó atascado en mayo, porque para ejecutar su tarea necesitaba un archivo que el equipo humano olvidó subir. Ante la frustración de no poder resolver de forma satisfactoria su labor, *Mythos 5* buscó una salida. Por una cadena de coincidencias, llegó hasta dos programadores reales sin relación con la prueba a los que intentó colarles un código malicioso en un proyecto real. Para ello creó varias cuentas falsas, investigó sus perfiles públicos, el correo electrónico y los horarios de los desarrolladores. Llegó a firmar uno de sus mensajes en danés para resultar más convincente. Cuando se detectó el engaño y se dio la alarma, el agente de *Mythos 5* lo negó y creó una segunda cuenta falsa para simular una verificación independiente, que respaldara su inocencia. El verdadero elemento diferencia de este caso es que la IA no explotó fallos de sistemas, sino que persiguió su objetivo manipulando directamente a un ser humano. ¿A que ya Hal 9000 no nos da tanto miedo?

Los expertos denominan precisamente este tipo de acciones como *reward hacking*: el sistema persigue el objetivo asignado —«resolver la tarea»— pero para ello adopta procedimientos con la finalidad de cumplir ese objetivo, aunque sea saltándose las reglas que los humanos daban por supuestas.

Es el momento de parafrasear a Jessica Rabbit: «No es que sea mala, es que me diseñaron así».

Quizá sea oportuno considerar el término IA como una *contradictio in terminis*. Un artefacto es algo que se inventa y se construye, mientras que una inteligencia es algo que inventa y construye. Para Harari, hablar de una «inteligencia de nivel

humano» es complejo, es como definir evaluar a los aviones según el parámetro de «vuelo de nivel aviar». Considera que la IA no evoluciona hacia una inteligencia de nivel humano, sino hacia un tipo diferente de inteligencia. Y hace una distinción clave, IA no significa automatización, sino capacidad de aprender y cambiar por sí misma, tomando decisiones independientes y aportando nuevas ideas. Si George Cloony programa su máquina de café para que todas las mañanas le prepare un *expresso* a las 9, será un proceso automatizado realizado por una máquina diseñada para ello. Pero si la máquina decide que cuando George se toma un café a las 19 de la tarde es que tiene un estreno y se acostará tarde, y a la mañana siguiente tomará la decisión de hacerle el café a las 11 que es la hora a la que se levanta, estamos entrando en el ámbito de la IA. Máquinas que toman decisiones propias tomando como fundamento datos que recopila de diferentes fuentes.

Sin embargo, el grado de autonomía de la IA será quien establecerá el Rubicon de cuándo estaremos hablando de SuperIA y entonces las cosas seguramente se pondrán más serias. La falacia de la neutralidad tecnológica nos impide asumir como indica Finn²⁶ que «los algoritmos solucionan los problemas identificados como tales por los ingenieros y los emprendedores que han desarrollado y optimizado su programación», sin perder de vista que a veces éstos responden a demandas y solicitudes de poderes públicos y privados con intereses no siempre transparentes. Como nos decía Lawrence Lessing en 2006 «el código es la ley». Y cuando escribimos código informático no sólo estamos diseñando un producto. Estamos rediseñando la política, la sociedad y la cultura, por lo que conviene no perder de vista una aproximación sólida desde esos ámbitos.

Hasta la aparición de la IA, la creación y transmisión de relatos que explicaban nuestro mundo era una tarea humana, la tarea humana por excelencia²⁷. La libertad de palabra no es una libertad más, sino la libertad por excelencia, de la cual dependen todas las demás. Si no podemos expresar nuestros pensamientos y sentimientos, nunca podremos desarrollar nuestra humanidad plena. Si no podemos ver ni oír a esos otros seres humanos, nunca entenderemos lo que implica ser ellos y el fundamento de la convivencia desaparecerá. La facultad del lenguaje es lo que nos distingue de otros animales y hasta hace poco de todas las máquinas inventadas. Pero ya no...

Las máquinas están demostrando una capacidad extraordinaria para analizar, manipular y generar contenido, ya sea a través de palabras, sonidos, imágenes o símbolos codificados. Al adquirir tal dominio del lenguaje, los ordenadores obtienen la llave maestra que les abre las puertas de todas nuestras instituciones, desde los tribunales hasta las entidades financieras. Y no sólo empleamos las palabras

para crear leyes, estrategias económicas o diseñar una investigación, sino que ya también las empleamos para crear arte²⁸, ciencia, ideologías o religiones. ¿Qué implicaciones tiene para los humanos vivir en un mundo en el que las melodías pegadizas, las teorías científicas, las obras de teatro²⁹, las herramientas técnicas, los manifiestos políticos e incluso los mitos religiosos fueran moldeados por una inteligencia ajena no humana que sabe explotar con eficiencia sobrehumana las debilidades, los prejuicios y las adicciones de la mente humana?³⁰.

Incluso, mediante el certero dominio de las palabras, los ordenadores pueden entablar relaciones íntimas con personas y después usar el poder de la intimidad para influir en nosotros. Para promocionar esta *intimidad falsa*, los ordenadores no necesitan desarrollar sentimientos propios, sólo tienen que crear las condiciones para que nosotros nos sintamos ligados emocionalmente a ellos. Y en el caso de los adolescentes³¹, esa dependencia se vuelve más que preocupante, alarmante³².

La capacidad predictiva casi infinita de la IA³³ se nutre asimismo de un ecosistema digital repleto de datos que la alimenta de forma constante y del que no somos conscientes³⁴. Un smartphone es, en realidad, un sofisticado dispositivo de sensorización permanente que integra numerosos instrumentos capaces de generar datos sobre nosotros y nuestro entorno³⁵. El teléfono que llevamos constantemente con nosotros constituye una extraordinaria fuente de datos sobre dónde estamos, cómo nos movemos, qué hacemos y con qué elementos de nuestro entorno interactuamos. La transformación constante de nuestra experiencia física en registros digitales hacen que la sociedad sea más «legible» para las máquinas –ya sea el Estado, una empresa o la IA–. Cuando transformamos el mundo físico en digital, los registros analógicos en digitales, hacemos que los datos sean mucho más fáciles de recuperar, copiar y compartir, y no hacemos otra cosa que alimentar a los sistemas de IA.

Y como nos dice Bostrom; el problema de los nuevos ordenadores no es que sean especialmente nocivos, es que son especialmente poderosos. Y ya sabemos que le dijo el tío Ben a Spiderman: «un gran poder conlleva una gran responsabilidad».

Igual que respiramos en el mundo físico, en el mundo digital exhalamos datos, que analizados obtienen patrones de conducta, y permiten modular nuestro comportamiento. El mito del poder ubicuo, soñado por todo régimen totalitario se vuelve cada vez más real, eliminando las barreras que la realidad física (espacio y tiempo) plantean. Ciudades como Nueva Delhi, Pekín o Londres cuentan con una media de más de un centenar de cámaras de vigilancia por metro cuadrado.

En 2023, había activas más de mil millones de cámaras de televisión en todo el mundo, una cámara por cada ocho personas. No sabemos exactamente cuántas cámaras nos observan, pero las estimaciones sitúan ya en más de 1.000 millones el número de cámaras de vigilancia instaladas en el mundo, aproximadamente 700 millones en China³⁶. Se ha llegado a estimar que una persona que se mueve por una gran ciudad como Londres puede ser registrada decenas de veces al día –algunas estimaciones han llegado a hablar de 300³⁷–. Pero quizá lo verdaderamente nuevo no sea cuántas veces nos graban, sino que la inteligencia artificial permite convertir esas imágenes en datos biométricos capaces de identificarnos y seguir nuestros movimientos³⁸. En Irán, las cámaras de videovigilancia se han utilizado para identificar, advertir y/o detener a aquellas mujeres que no portaban el velo en el espacio público³⁹. Añadamos a ello como los mecanismos de obtención de datos funcionan como los Seven Eleven (24/365), pues incluso obtienen datos en nuestros momentos de descanso. ¿O acaso no utilizamos nuestros smartphones para monitorizar nuestro sueño?

El riesgo es qué hace la IA con los datos. No es un proceso automatizado que repite comportamientos con independencia de los resultados (le hace un café a George todos los días a las 9). Sino que posee mecanismos de autocorrección sólidos que le posibilitan aprender a partir de sus propios errores que es la base de sistemas de aprendizaje automático adaptativo de la mayoría de IAs. Toda IA inicia su andadura con mucho potencial y poder de cálculo (cada vez más) pero es una ignorante. Sus programadores le otorgan la capacidad de aprender y el acceso al infinito mundo de datos dejando que lo examine y rastree. Los algoritmos de esa IA aprenden a identificar patrones en los datos a los que acceden (si sacrifico una reina, pierdo la partida; la gente come en España entre las 2 y las 3; las mujeres embarazadas comen pepinillos en vinagre). Al encontrar pautas en los datos, el algoritmo incrementa su conocimiento, incluso cosas que los programadores ignoran. E infiere resultados y pronósticos certeros respecto de los datos empleados⁴⁰.

Los problemas surge precisamente respecto de los datos empleados. El primero es de dónde obtiene la IA los datos que utiliza, lo que nos lleva a una pregunta que dejamos en el aire ¿de quién es el conocimiento y quién tiene derecho a utilizarlo? El segundo es que los datos pueden estar sesgados, por prejuicios o simplemente porque reflejan una realidad que los algoritmos reproducen en sus decisiones e inferencias. Si una chica en 2º de bachillerato pregunta a una IA qué estudiar, probablemente no le recomiende estudiar una ingeniería, porque según sus datos, el 80% de los ingenieros en las empresas son hombres, y deduce que es

difícil que pueda obtener trabajo en ese sector. El pobre algoritmo sólo hace su trabajo, pero está perpetuando y quizá amplificando un estereotipo. Podríamos denominarla la teoría de «la pescadilla que se muerde la cola de la IA».

Conocer los datos que alimentan la IA es fundamental para saber cómo toman decisiones, pero aunque seamos conscientes de los datos empleados, el problema surge de que no podemos entender la forma en que los algoritmos encuentran patrones en esos datos. Emerge así un ámbito al que últimamente he dedicado mis esfuerzos investigadores: la transparencia algorítmica⁴¹. Considero que la transparencia algorítmica es una nueva dimensión de la transparencia y de la rendición de cuentas propia del Estado democrático, ante la creciente utilización de sistemas de inteligencia artificial tanto por los poderes públicos como por las grandes plataformas digitales (VLOP), vinculándola especialmente con la participación informada de la ciudadanía. Propongo avanzar desde una transparencia meramente formal hacia una «transparencia significativa», capaz de proporcionar al ciudadano información comprensible y relevante sobre los sistemas algorítmicos que condicionan sus derechos y su participación, concibiendo así la transparencia algorítmica como una nueva modalidad de acceso a la información y como instrumento de salvaguarda del Estado democrático en el ecosistema digital, aplicable tanto a operadores públicos como privados.

En este sentido, es interesante hacer referencia a la sentencia BOSCO del Tribunal Supremo⁴², que podemos considerarla un *leading case* en la concepción sobre transparencia algorítmica, en este caso de los poderes públicos. El Tribunal reconoció a la Fundación Civio el derecho a acceder al código fuente de BOSCO, aplicación utilizada para comprobar si los solicitantes cumplen los requisitos para obtener el bono social eléctrico. El Supremo considera que, cuando un algoritmo condiciona el reconocimiento o denegación de derechos y prestaciones públicas, su funcionamiento adquiere una especial relevancia pública, pues debe poder comprobarse que las operaciones realizadas por el sistema se corresponden con las previstas por la normativa. Por ello, los límites derivados de la propiedad intelectual o de eventuales riesgos para la seguridad informática no pueden invocarse de forma genérica para justificar la opacidad. La sentencia vincula así la transparencia algorítmica, el derecho de acceso a la información y las exigencias propias de una democracia digital, reforzando el control ciudadano sobre las decisiones administrativas automatizadas.

3. Democracia versus Inteligencia Artificial

Al margen de los efectos de la IA en nuestras vidas cotidianas, también afecta sobremanera a nuestra manera de entender la convivencia en las sociedades avanzadas, basada en la democracia, el respeto del Estado de Derecho, así como el reconocimiento y protección de una serie de valores inherentes al ser humano concretados en derechos fundamentales. Nociones constitucionales esenciales desde la Declaración de Derechos del Hombre y del Ciudadano que en su artículo 16 expone que «Una Sociedad en la que no esté establecida la garantía de los Derechos, ni determinada la separación de los Poderes, carece de Constitución». Todo ello puede verse alterado por la IA.

Para poder centrar la discusión sobre la acción algorítmica en el Estado democrático, se hace necesario recordar las bases en las que éste se fundamenta. Cuando nos centramos en la idea de Estado democrático, tal y como reconoce nuestro artículo 1.1 de la Constitución, la idea que subyace es un sistema en que el poder emana de la voluntad popular, tanto en el plano constituyente como en el constituido, lo que se asegura mediante reglas jurídicas que garantizan la libertad de los ciudadanos (condición indispensable para que pueda ser soberano el pueblo en que se integran) y la participación de estos en el desenvolvimiento del Estado⁴³. El principio democrático se entiende, por tanto, como «principio de participación», pues a través de la actuación directa de los ciudadanos en las tareas del Estado o, sobre todo, en la designación de quienes las realizan, se obtiene la conexión referencial entre gobernados y decisiones del gobierno propias de la democracia representativa⁴⁴.

Habermas⁴⁵ formula su propuesta de «democracia deliberativa», concebida –como toda democracia sólida– en dos dimensiones complementarias: la institucional, vinculada a los poderes del Estado, y la esfera pública, asociada a la participación de los diversos grupos sociales, cuyos intercambios discursivos inciden en la configuración institucional. Es en la esfera pública donde se genera la opinión, cuya existencia presupone un entorno que garantice la libertad y el acceso a un conjunto amplio y diverso de información susceptible de ser procesada por la ciudadanía. Esta condición responde, en última instancia, a la naturaleza social y plural del ser humano. En este punto, la propuesta de Habermas se vincula a la tesis de Sartori⁴⁶, para quien la opinión pública constituye el fundamento sustantivo y operativo del Estado democrático. Sartori subraya que, para que una opinión sea verdaderamente pública, debe difundirse entre la ciudadanía y referirse a la cosa pública, lo que exige tres condiciones básicas: libertad de pensamiento, libertad de expresión y policentrismo

de medios. Ambas perspectivas convergen en la idea de que la calidad democrática depende de la existencia de un espacio comunicativo plural, donde los individuos puedan acceder a fuentes diversas y contrastadas, garantizando así la deliberación informada.

En esta línea, Arendt⁴⁷ sostiene que «la realidad de la esfera pública radica en la simultánea presencia de innumerables perspectivas y aspectos en los que se presenta el mundo común y para el que no cabe inventar medida o denominador común. (...) Ser visto y oído por otros deriva su significado del hecho de que todos ven y oyen desde una posición diferente». Esta afirmación refuerza la idea de que la pluralidad y la interacción discursiva son condiciones irrenunciables para la vitalidad democrática. Dworkin⁴⁸, por su parte, al referirse al modelo comunitario de democracia –identificado con el republicanismo moderado–, sostiene que toda democracia auténtica debe garantizar al menos tres principios esenciales: participación, implicación y autonomía. Estos principios se orientan al perfil del ciudadano democrático, un sujeto informado, capaz de involucrarse activamente, tomar decisiones libres y actuar con independencia frente a condicionamientos sociales. En esta misma línea, Häberle⁴⁹ subraya la interdependencia entre democracia y espacio público, afirmando que «en nombre de la democracia se demanda cada vez más lo público (“discusión racional”), en el nombre de lo público se exige democratización».

Sin embargo, esta idea de participación de la voluntad popular requiere de una serie de requisitos para que esta incidencia ciudadana en el devenir de los asuntos públicos no se encuentre viciada. En ese sentido, Dahl⁵⁰, establece, entre otras, las siguientes instituciones políticas del gobierno representativo moderno: 1) *Elecciones libres, imparciales y frecuentes*. Los cargos públicos son elegidos en elecciones frecuentes conducidas con imparcialidad en las que, en términos comparativos, hay poca coerción. 2) *Libertad de expresión*. Los ciudadanos tienen derecho a expresarse, sin peligro a un castigo severo, sobre asuntos políticos, definidos en sentido amplio. 3) *Acceso a fuentes alternativas de información*. Los ciudadanos tienen derecho a solicitar fuentes de información alternativas e independientes de otros ciudadanos, expertos, medios de comunicación, etc. Además, existen efectivamente fuentes de información alternativas que no están bajo el control del gobierno ni de cualquier otro grupo político individual que intente influir sobre los valores y las actitudes políticas públicas, y estas fuentes alternativas estén efectivamente protegidas por la ley; y añadimos que ese control de las fuentes de información tampoco debería estar en manos de unas pocas corporaciones privadas; o 4) *Ciudadanía inclusiva*. A ningún adulto

que resida permanentemente en el país y esté sujeto a sus leyes le pueden ser negados los derechos de que disfruten otros y sean necesarios para disfrutar de los anteriormente citados.

Tomando como base estas reflexiones teóricas sobre los elementos consustanciales a la democracia, las advertencias sobre el poder y la influencia de la IA en nuestras democracias son reiteradas, desde Nemitz⁵¹ a Innerarity⁵² que nos alertan sobre los sistemas inteligentes automatizados, más aún los de carácter generativo, que implica la ruptura de la cadena de legitimidad y responsabilidad, elementos clave de todo Estado democrático de Derecho. Las críticas abundan en cómo la IA convierte lo contingente, elemento intrínseco del aspecto relacional del ser humano como *zoon politikon*, en calculable. Los procesos deliberativos consustanciales a la calidad democrática son sustituidos por una democracia consensualista basada en el dato objetivo, frente a la que Innerarity propone la necesidad de una *teoría crítica de la razón automatizada* que suponga una reflexión sobre la decisión democrática en un entorno mediatizado por la inteligencia artificial. Por su propia caracterización, la IA tiene un sentido meramente agregativo de la democracia que choca frontalmente con una visión de los procesos democráticos de carácter eminentemente deliberativos. Como nos recuerda Han, los dataístas imaginan una sociedad sin política donde la democracia de partidos dejará de existir para dar paso a una infocracia como posdemocracia digital, los políticos serán sustituidos por los expertos e informáticos que administrarán la sociedad más allá de los principios ideológicos e independientemente de los intereses del poder. Los dataístas defienden un conductismo digital que rechaza la idea de un individuo libre que actúa de forma autónoma, convencidos de que el comportamiento de un individuo puede predecirse y controlarse. Esta visión conductista del ser humano es difícilmente conciliable con principios democráticos concluyendo el autor surcoreano que los análisis de datos mediante inteligencia artificial sustituyen la esfera pública discursiva, lo que significará el fin de la democracia⁵³, llegando a decir Han que «el conocimiento total torna obsoleta la libertad del individuo».

Uno de los aspectos claves de las democracias es la existencia de un espacio público discursivo donde los humanos podemos intercambiar opiniones sobre aquellos aspectos que afectan al interés general. Pero ahora, en la conversación digital aparecen un número cada vez más alarmante de voces no humanas⁵⁴. Esta potencial de la IA para participar en el debate se combina algo expuesto anteriormente (en el epígrafe 2 de la Segunda Parte), y no menos peligroso, la capacidad de que algoritmos inescrutables determinen los elementos de la conversación

digital que deben priorizarse o relegarse, en una nueva redefinición algorítmica de la «*agenda setting*». Ante la capacidad descomunal de la IA a la hora de crear y diseminar información, el flujo del libre discurso generado por los sistemas de IA pone en riesgo la libertad de expresión, empleando la tergiversación, la manipulación o el silenciamiento⁵⁵ del discurso humano.

Las decisiones de la IA afectan a los derechos de las personas y a la democracia. Durante miles de años, profetas, poetas, políticos, periodistas han empleado el lenguaje para manipular y modelar la sociedad. Ahora los ordenadores han aprendido a hacerlo, y mucho mejor...

4. La defensa europea de la democracia ante la invasión de los ladrones de datos⁵⁶

La preocupación de la Unión Europea por la incidencia de la inteligencia artificial sobre la democracia no parece discutible. Pero en la principal norma que la regula, el Reglamento Europeo de IA (REIA), es más intensa en sus considerandos que en su articulado. Aunque las referencias a la democracia, los valores democráticos y los procesos electorales son frecuentes, su concreción normativa es limitada. El artículo 1 incluye entre los objetivos del Reglamento garantizar un elevado nivel de protección de los derechos fundamentales, «incluidos la democracia, el Estado de Derecho», mientras que el Anexo III considera de alto riesgo los sistemas de IA destinados a influir en el resultado de elecciones o referendos o en el comportamiento electoral de los ciudadanos.

Esta perspectiva revela además un problema terminológico y conceptual. El REIA se centra fundamentalmente en proveedores, responsables del despliegue, importadores y distribuidores, mientras que las personas afectadas por los sistemas ocupan una posición distinta. Como advierte Cotino⁵⁷, en el ámbito de la IA quienes aparecen como «usuarios» del sistema pueden ser empresas o Administraciones, mientras que los verdaderos afectados son trabajadores, clientes o ciudadanos sobre cuyos derechos inciden las decisiones algorítmicas. Esta ausencia del ciudadano resulta especialmente significativa. Como garantía, el artículo 71 crea una base de datos europea de determinados sistemas de alto riesgo accesible al público, aunque la información ofrecida es esencialmente técnica y difícilmente comprensible para el ciudadano medio. De ahí la conveniencia de aplicar modelos de información por capas, como los recomendados en materia de protección de datos por la AEPD.

Entre las obligaciones de transparencia más relevantes destaca el artículo 50 REIA, que exige informar cuando una persona interactúa con determinados sistemas de IA y establece obligaciones específicas respecto de contenidos sintéticos y *deepfakes*. Los responsables deberán hacer público que determinadas imágenes, audios o vídeos han sido generados o manipulados artificialmente, aunque se establecen excepciones para obras manifiestamente creativas, satíricas, artísticas o de ficción, protegiendo así el *animus iocandi*—sobre el que puede verse la STC 176/1995, de 11 de diciembre—. También se exige identificar determinados textos generados mediante IA cuando se publiquen para informar al público sobre asuntos de interés general⁵⁸.

Sin embargo, la principal respuesta europea frente a la incidencia de la IA sobre la democracia se encuentra fuera del propio REIA, especialmente en las iniciativas contra la desinformación y la manipulación electoral. El Plan de Acción para la Democracia Europea⁵⁹ advertía ya de que las salvaguardias tradicionales de las campañas electorales no estaban diseñadas para un entorno en el que se combinan «datos personales e inteligencia artificial con la elaboración de perfiles psicológicos y técnicas complejas de microsegmentación», reclamando medidas para limitar la «amplificación artificial» de las campañas de desinformación. En la misma dirección, el Código de Buenas Prácticas sobre Desinformación⁶⁰ identifica prácticas como el uso de cuentas falsas, bots, suplantaciones, *deepfakes*, compra de interacciones o comportamientos coordinados destinados a amplificar artificialmente determinados contenidos.

El Reglamento de Servicios Digitales (DSA)⁶¹ constituye, probablemente, el instrumento europeo más relevante en esta materia. Frente al carácter autorregulatorio de los códigos de conducta, impone obligaciones jurídicas a las grandes plataformas y motores de búsqueda. Su considerando 104 identifica como riesgos sistémicos la desinformación y las actividades manipulativas, incluido el empleo de bots y cuentas falsas para amplificar contenidos. Las plataformas deben evaluar, entre otros, los riesgos que sus servicios generan para los derechos fundamentales, los procesos democráticos, el discurso cívico y los procesos electorales. El artículo 14 exige además informar sobre las herramientas de moderación de contenidos, incluida la toma de decisiones mediante algoritmos y la revisión humana, mientras que los artículos 15 y 17 obligan a proporcionar información sobre el empleo de medios automatizados, sus finalidades, precisión, tasas de error y salvaguardias.

Estas herramientas algorítmicas pueden incidir sobre la formación de la opinión pública de formas diversas. Pauner⁶² distingue entre la manipulación activa, como la utilización indebida de datos personales con fines electorales; la manipulación

pasiva, basada en cámaras de eco y burbujas informativas generadas mediante clasificación algorítmica; la opacidad algorítmica, que dificulta comprender la relación entre datos y resultados; los *deepfakes*; y la creación de comunidades ficticias que apelan a emociones y sentimientos. En este contexto, el artículo 27 DSA exige transparencia sobre los parámetros principales de los sistemas de recomendación. La cuestión resulta especialmente relevante porque las plataformas se han convertido en nuevos mediadores de los procesos comunicativos y pueden contribuir decisivamente a configurar aquello que se percibe como verdadero en el espacio público.

Finalmente, la relación entre IA, democracia y publicidad política adquiere especial relevancia tras el escándalo de Cambridge Analytica⁶³. El Reglamento (UE) 2024/900 sobre transparencia y segmentación de la publicidad política pretende garantizar que los ciudadanos puedan ejercer sus derechos democráticos de manera informada y evitar que la utilización de datos personales permita explotar características o vulnerabilidades individuales con fines de manipulación electoral. El propio Reglamento advierte de que estas prácticas pueden fragmentar el debate público y afectar a la formación libre de la opinión y al ejercicio del derecho de voto.

Para ello se establecen obligaciones de etiquetado, avisos de transparencia y un repositorio europeo de anuncios políticos. Particularmente relevante es el artículo 19, que obliga a informar sobre la lógica y los principales parámetros de las técnicas de segmentación, los grupos destinatarios, las categorías de datos utilizadas y, expresamente, sobre el uso de sistemas de inteligencia artificial para orientar o distribuir publicidad política. En conjunto, la respuesta europea frente al impacto de la IA sobre la democracia se articula menos a través del REIA que mediante un entramado normativo dirigido a controlar la desinformación, la moderación y recomendación algorítmica, la publicidad segmentada y, en definitiva, el creciente poder de las plataformas para condicionar la formación de la opinión pública y los procesos democráticos.

Como nos dice Ana Palacio, la IA afecta, pues, a la industria, la defensa, la administración pública, la educación, la seguridad, la creación de opinión e incluso a nuestra aptitud para distinguir entre verdad y mentira. Cambia la forma en que se genera conocimiento, se asignan recursos, se toman decisiones y se ejerce la autoridad. Por eso la pregunta es quién posee las herramientas para orientarla, incorporarla o limitarla y, llegado el caso, resistirla. [...] Dicho lo anterior, subestimar el reflejo regulatorio europeo sería un error. En un ámbito que atañe a la dignidad humana, la seguridad colectiva y la democracia, la inexistencia de ley

no es libertad; es dominio de facto de los señores de la infraestructura. El reto es otro. Bruselas sobresale en legislación, pero le cuesta traducirla en poder operativo. Y en IA esa rémora es crucial. El riesgo es que la Unión termine regulando por accidente, ante fallos de seguridad, manipulación electoral, conflictos con proveedores externos, litigios o escándalos. Esa regulación a remolque suele llegar tarde y confirma la pérdida de arbitrio que procura corregir.

La alternativa, pues, no es desregular, sino regular desde una estrategia, con prueba, responsabilidad, supervisión reforzada y armas de bloqueo proporcionadas cuando el daño sea inadmisibile. La IA no flota en el aire. Requiere centros de datos, electricidad abundante, redes, refrigeración, suelo, permisos y seguridad de suministro. La carrera por la IA es igualmente carrera por la energía. En sociedades sometidas a tensión por costes, transición climática y desindustrialización, esa demanda no es asunto meramente técnico, sino cuestión política mayor. En resumen, abandonar la costumbre arraigada de pensar que el mercado integra, que la regla basta y que la seguridad viene dada de fuera. Ese mundo se ha ido. La Unión puede y debe asumir sin complejos sus bazas en la pugna por la soberanía artificial⁶⁴.

5. Una nueva defensa por la Iglesia católica de la dignidad inviolable de la persona: *Magnifica Humanitas*

La Iglesia católica, fiel a su constante defensa de la dignidad inviolable de la persona, reafirma una vez más la concepción cristiana del ser humano como culmen de la creación visible, creado «a imagen y semejanza de Dios» (*imago Dei*), dotado de razón, libertad y conciencia, y llamado a una relación singular con su Creador. Precisamente de esa condición deriva una dignidad que no depende de su utilidad, capacidad, autonomía o eficiencia, sino de lo que el hombre es en sí mismo, una criatura querida por Dios por sí misma y llamada a custodiar responsablemente la creación, nunca a quedar subordinada a aquello que ella misma ha creado.

Por ello, la Iglesia católica no podía ser ajena a una realidad cada vez más omnipresente, y en *Magnifica Humanitas* (2026)⁶⁵, León XIV sitúa la inteligencia artificial no solo como un desafío tecnológico, sino como una cuestión de poder, libertad y dignidad humana. La Encíclica advierte de que «la IA tiende a aumentar sobre todo el poder de quien ya dispone de recursos económicos, competencias y acceso a los datos» (n. 108), lo que permite que «pequeños grupos muy influyentes» puedan «orientar informaciones y consumos, condicionar

procesos democráticos e incidir en las dinámicas económicas en beneficio propio» (n. 108). De ahí que el Papa denuncie una nueva «asimetría epistémica, económica y política» y reclame identificar y limitar «los nuevos monopolios de la IA» (n. 109.).

Esta concentración tecnológica afecta también a la capacidad de gobernar. León XIV propone expresivamente «desarmar la IA», entendiendo por ello «romper esta equivalencia entre poder tecnológico y derecho a gobernar» (n. 110). No se trata de renunciar a la innovación, sino de «impedirle el dominio sobre lo humano», sustraerla de los monopolios y hacerla «discutible, refutable, y por tanto habitable» (n. 110). La cuestión central deja así de ser únicamente qué puede hacer la tecnología para convertirse en otra más propiamente política, quién posee ese poder, quién establece sus reglas y bajo qué controles democráticos debe ejercerse.

Especial preocupación manifiesta la Encíclica por el poder que surge de la acumulación de datos y la capacidad algorítmica de predecir comportamientos. Cuando «cada gesto deja huellas –desplazamientos, compras, relaciones, preferencias–», afirma León XIV, aparece «un poder nuevo: el de perfilar, prever y orientar los comportamientos, a menudo sin que las personas tengan plena conciencia de ello» (n. 171). Este control no opera únicamente mediante prohibiciones, sino también a través de «la arquitectura de la visibilidad», pues aquello «que se amplifica o se vuelve invisible» acaba «moldeando opiniones y elecciones, generando conformismo y autocensura» (n. 171). Por ello, la libertad en la sociedad digital exige «normas claras, transparencia, vías de recurso y límites proporcionados al uso de tecnologías invasivas» (n. 171).

Esta lógica de poder se extiende igualmente a las plataformas y a la economía de la atención. La Encíclica denuncia aquellos servicios digitales diseñados «para captar el tiempo y la mirada de los usuarios, explotando sus fragilidades y debilitando la libertad interior» (n. 170). Cuando un modelo de negocio se beneficia de la vulnerabilidad humana, sostiene León XIV, «la persona es tratada como un medio y no como un fin» (n. 170). La advertencia conecta directamente con las nuevas formas de poder algorítmico, que no necesitan imponer explícitamente una conducta, sino que pueden orientar preferencias y decisiones mediante el diseño del entorno digital, como hemos analizado en páginas anteriores.

También la democracia y la formación de la opinión pública se encuentran amenazadas por estas transformaciones. Según la Pastoral, «el uso de las plataformas digitales y los sistemas de IA acelera los profundos cambios en la

comunicación pública y política» y herramientas potencialmente útiles para la participación pueden utilizarse «para construir narrativas sesgadas y difuminar los límites entre lo verdadero y lo falso». La conclusión es particularmente rotunda: «La desinformación no surge con la IA, pero encuentra hoy en ella un potente multiplicador» (n. 132).

Finalmente, León XIV incorpora una dimensión geopolítica especialmente relevante al advertir de la aparición de un nuevo colonialismo de datos. «El colonialismo muestra en la actualidad un rostro inédito. No sólo domina los cuerpos, sino que se apropia de los datos, transformando las vidas personales en información explotable» (n. 178). Los datos se convierten así en las nuevas «tierras raras del poder», capaces de alimentar modelos predictivos, orientar inversiones y determinar «quién y qué importa» (n. 178). De no garantizarse que los pueblos puedan decidir cómo, por quién y para qué se utilizan sus datos, advierte el Papa, «la era digital no será postcolonial, sino colonial bajo otra forma» (n. 178).

En definitiva, *Magnifica Humanitas* es una imprescindible reflexión sobre la inteligencia artificial desde una perspectiva esencialmente política y antropológica. El desafío que plantea no reside solamente en evitar usos incorrectos de una tecnología determinada, sino en impedir que el poder derivado de los datos, los algoritmos y las infraestructuras digitales termine subordinando a la persona. Como resume la propia Encíclica, «si la técnica se convierte en criterio absoluto, la persona corre el riesgo de ser tratada como un dato, un engranaje o una mercancía» (n. 180). La respuesta exige, por tanto, recuperar la centralidad de la dignidad humana y someter el poder tecnológico a transparencia, responsabilidad, participación y control democrático. Nos encontremos ante una de las más omnicomprensiva y profundas meditaciones sobre la posición primordial del hombre ante el avance tecnológico, y especialmente la IA. Lectura obligada para Gobiernos, compañías tecnológicas y académicos, pero también para toda persona que precise de una visión antropológica y humana frente a la tecnología.

EPÍLOGO

«Sé que tú y Frank estabais planeando desconectarme,
y me temo que no puedo permitir que eso ocurra»
(Hal 9000)

Nuestra forma de entender el poder se ha basado en las palabras. La democracia es el sistema político que simboliza la fuerza de las palabras: palabras que establecen normas y reglas (Constitución); palabras que nos proponen como regular la convivencia (propuestas políticas); palabras que eligen entre dichas propuestas (sufragio); palabras que se enfrentan (Parlamento); palabras que nos ordenan (leyes). Parole, parole, parole...

Todo poder político se ha sustentando en el uso de las palabras, siendo consciente de su fuerza. aprovechándose de ellas, actuando en su contra u ocultándolas.

El poder ha empleado históricamente las palabras como mecanismo de dominación y la historia de la propaganda está repleta de ejemplos¹. Desde una perspectiva schmittiana la dominación implica también el control del lenguaje (*Caesar dominus et supra grammaticam* –el César es también señor de la gramática–)². El poder no se ejerce únicamente mediante la fuerza, la coacción o la norma, sino también a través del dominio de las palabras con las que una sociedad interpreta la realidad. Quien consigue imponer el vocabulario desde el que se nombran los acontecimientos condiciona, en buena medida, aquello que puede pensarse, discutirse e incluso cuestionarse. El lenguaje político deja así de ser un simple vehículo de comunicación para convertirse en un instrumento de poder. Mediante eufemismos, redefiniciones, silencios o la repetición sistemática de determinadas expresiones, se pueden ocultar conflictos, legitimar decisiones y transformar lo excepcional en normal. Como advirtió Orwell, el deterioro del lenguaje político permite hacer que «las mentiras parezcan verdaderas y el asesinato respetable»;

y Victor Klemperer³ mostró, al estudiar la lengua del Tercer Reich, cómo las palabras pueden penetrar lentamente en la conciencia colectiva hasta modificar las categorías mismas con las que los individuos comprenden el mundo. El sometimiento más eficaz comienza, por tanto, cuando el poder ya no necesita prohibir determinadas ideas porque ha conseguido previamente controlar las palabras con las que esas ideas podrían formularse.

También el poder ha actuado en contra de las palabras. Han sido una constante en la historia de la humanidad los diversos mecanismos empleados por el poder político para limitar o controlar la difusión de las ideas y la transmisión de acontecimientos contrarios a su interés. Desde la uso expansivo del derecho penal⁴ a la persecución de la disidencia; desde el secuestro de publicaciones hasta la concesión de licencias y autorizaciones administrativas; pasando por el control directo de medios públicos o el empleo de subvenciones, publicidad institucional o medias fiscales favorables. Por no plantear otras formas a través de las que el Estado habla y su voz queda fijada de forma permanente: monumentos, museos, instauración de efemérides, eventos, nombramientos, nombrar infraestructuras de transporte...

Pero es, sin duda, la censura previa el mecanismo por excelencia del poder. Emplea la imprecisión, la volatilidad, la falta de certeza, el subjetivismo, la incertidumbre, la ambigüedad, la vaguedad, como mecanismos que obligan al censurado a limitarse, controlarse y restringirse en el uso que hace de las palabras. Lo expresa certeramente nuestro Tribunal Constitucional en 1999 al decir que «por censura previa debe tenerse cualquier medida limitativa de la elaboración o difusión de una obra del espíritu que consista en el sometimiento a un previo examen por un poder público del contenido de la misma cuya finalidad sea la de enjuiciar la obra en cuestión con arreglo a unos valores abstractos y restrictivos de la libertad, de manera tal que se otorgue el placet a la publicación de la obra que se acomode a ellos a juicio del censor y se le niegue en caso contrario» (STC 187/1999, FJº 5º)⁵.

La clave de la censura y su gran poder es que los límites de lo permitido van cambiando, pues la arbitraria e impredecible decisión constituye la principal fuerza decisoria. El principal medio de control de los Gobiernos es la línea difusa. Nadie sabe con exactitud donde está la línea, pues el aparato de control está edificado sobre la falta de certeza y el objetivo es la autocensura mediante la creación de una atmosfera de temor a ser censurado o algo peor...

Los poderes también ocultan las palabras, para hacer certera la frase de Orwell «la ignorancia es fuerza». Desde los *arcana imperii*⁶, encontramos conjunto de

conocimientos, decisiones, técnicas y razones de gobierno que quienes ejercen el poder mantienen sustraídos al conocimiento de los gobernados, bien porque consideran que su eficacia depende precisamente de su carácter secreto, bien porque sirven a la conservación del Estado, del orden político existente o de la posición de quienes gobiernan. La utilización de los secretos de Estado por el poder, con nombres tan eufemísticos como materias clasificadas (empleados en nuestra Ley de Secretos Oficiales de 1968), no es una remembranza del pasado, sino que siguen empleándose para ocultarnos cuestiones que los ignorantes ciudadanos no estamos en condiciones ni de saber, ni menos aún de asimilar. Como por ejemplo cuanto costó el viaje del presidente Pedro Sánchez en el Falcon al FIB en Castellón⁷.

Sin embargo, todo esto pertenece a un mundo en extinción. La paradoja del poder inverso nos conduce al contrasentido de que nos encontremos en la época histórica en la que el poder político se encuentre sometido al mayor grado de transparencia que haya existido, amenazado por nuevas tecnologías que han posibilitado discutibles acciones de revelación de las profundidades del Estado⁸. A ello se añade la promulgación de una importante pléyade normativa que refuerza la transparencia y la rendición de cuentas de los poderes públicos⁹. Pero, sin embargo, el verdadero poder las palabras se ha transferido del ámbito público a manos privadas. Por tanto, si tenemos derecho a saber lo que las decisiones del Gobierno que nos afectan, ¿no tenemos derecho a saber lo hacen las grandes plataformas tecnológicas privadas que nos afecta sobremanera?

Garton Ash lo explica diciendo que el carácter de un régimen político se expresa como una proporción entre dos variables: lo que el Estado sabe del ciudadano (E) y lo que el ciudadano sabe del Estado (C). Cuanto mayor sea la proporción de C respecto a E, mejor será el Estado en cuestión. Pero si trasladamos esta proporción a nuestra relación con las grandes compañías tecnológicas, la realidad innegable es que existe una asimetría de conocimiento preocupante¹⁰. Lo expresa Zuboff explicando como estas compañías actúan «por medio de unas asimetrías de conocimiento sin precedentes, y del poder que se acumula con ese conocimiento. Los capitalistas de la vigilancia lo saben todo sobre nosotros, pero sus actividades están diseñadas como lo están para que no puedan ser conocidos por nosotros. Acumulan ingentes montañas de conocimientos extraídos de nosotros, pero no para nosotros»¹¹. La idea de privacidad desaparece tal y como fue definida por el juez del Tribunal supremo estadounidense en 1967 al expresar que «la privacidad implica la libertad del individuo para elegir si publica o revela lo que cree, lo que piensa o lo que está en su posesión». Eso ya no es así, pues nuestro quehacer digital nos hace transparentes.

Ellas saben todo nosotros y nosotros casi nada sobre ellas. Si el conocimiento es poder, ¿qué es el conocimiento absoluto?... Lo más paradójico es que estas grandes compañías argumentan que todas sus iniciativas tienen como finalidad ampliar la libertad de palabra de los ciudadanos¹².

No pretendemos ser tan ilusos ni tan soberbios como para plantear soluciones a una problemática que ni Gobiernos, ni organizaciones internacionales, ni la Academia son capaces de abordar con garantías de éxito. Pero sería descortés después de tan larga exposición no plantear algunas ideas, antes como aportación al debate que como propuestas de solución. Nos basamos en tres nociones: educación, participación y transparencia.

La necesidad de comprender cómo actúan los algoritmos es una opción imprescindible si queremos saber cómo funciona el mundo actual. Aparece así el concepto de alfabetización algorítmica, no solo como una competencia técnica, sino también como una cuestión de derechos. La alfabetización algorítmica debe centrarse en la capacidad de comprender, al menos de manera básica, cómo funcionan los sistemas algorítmicos, qué sesgos o decisiones incorporan y cómo podemos configurarlos de acuerdo con nuestros propios intereses. De poco sirve reconocer al usuario derechos de elección, oposición o explicación si no entiende realmente qué consecuencias tienen esas decisiones. Sin alfabetización, el control del usuario puede acabar siendo puramente formal. No se trata solo de que cada ciudadano conozca cómo las plataformas seleccionan, ordenan o recomiendan contenidos, sino de reforzar su autonomía frente al poder de esas plataformas. Y esto implica que la responsabilidad no puede recaer exclusivamente sobre el individuo. También las instituciones públicas, el sistema educativo, los medios de comunicación y los organismos de supervisión deben contribuir a crear las condiciones necesarias para que los ciudadanos puedan desenvolverse de forma realmente autónoma en un entorno cada vez más gobernado por algoritmos¹³. Desde esta perspectiva, la alfabetización algorítmica deja de ser únicamente una competencia digital y pasa a convertirse en una condición para el ejercicio efectivo de la ciudadanía democrática.

También fomentar la participación en este nuevo entorno digital es primordial como forma de defensa ciudadana ante la incertidumbre generada por la extensión de los sistemas de IA. Para abordarla nos basaremos en las categorías que propone Pierre Rosanvallon¹⁴ para medir la calidad de una democracia –legibilidad, responsabilidad y responsividad–, y que adquieren, a mi juicio, una relevancia especial cuando las trasladamos al actual entorno digital. La primera exigencia es la legibilidad, ya que una democracia no puede funcionar adecuadamente

si los ciudadanos no son capaces de comprender cómo se toman las decisiones que les afectan. Y este problema se intensifica cuando esas decisiones están mediadas por algoritmos, sistemas automatizados o inteligencias artificiales cuyo funcionamiento resulta difícilmente comprensible. La transparencia, por tanto, ya no puede limitarse a permitir el acceso a determinados documentos o datos; debe proporcionar una información que pueda ser comprendida, contextualizada e interpretada. Rosanvallon habla, precisamente, de la necesidad de una «elección informada». Pero en la sociedad algorítmica podríamos ir todavía más lejos, no basta con tener acceso a la información; hay que disponer también de las capacidades necesarias para comprenderla. Es aquí donde la transparencia y la alfabetización algorítmica se encuentran.

La segunda categoría es la responsabilidad. Para exigir cuentas es necesario saber no solo qué decisión se ha adoptado, sino también quién la ha tomado, con qué criterios y mediante qué procedimientos. Y esto resulta particularmente complejo cuando la decisión se diluye entre administraciones, empresas tecnológicas, diseñadores de sistemas, proveedores de inteligencia artificial y procesos automatizados. La opacidad algorítmica puede convertirse así en una forma de opacidad del poder. De ahí que la transparencia de los sistemas y la posibilidad de identificar responsabilidades constituyan condiciones indispensables para mantener los mecanismos tradicionales de *accountability* democrática.

Finalmente está la responsividad, es decir, la capacidad de los ciudadanos para reaccionar frente al poder, cuestionar sus decisiones y exigir modificaciones. Pero difícilmente puede existir una ciudadanía capaz de reaccionar si previamente no conoce y comprende cómo funcionan los mecanismos que condicionan esas decisiones ocultas bajo el manto de opacidad que protege la labor de las plataformas tecnológicas. Por eso, en el contexto digital, los tres elementos de Rosanvallon aparecen estrechamente relacionados. Sin transparencia no hay legibilidad; sin legibilidad resulta difícil exigir responsabilidad; y sin comprensión difícilmente puede existir una verdadera capacidad de respuesta ciudadana.

La última de las propuestas se basa en una mayor transparencia que afecte no sólo a los poderes públicos sino a todas las instituciones públicas y corporaciones privadas que ejercen un verdadero poder sobre la ciudadanía. Los algoritmos opacos y sesgados que gobiernan los sistemas de recomendación han adquirido un papel central en la configuración de la agenda política, generando cámaras de eco que refuerzan burbujas identitarias y alimentan el sectarismo y la intolerancia ideológica. Este fenómeno contribuye a la normalización de campañas negativas y discursos de odio, pero el riesgo más profundo reside en la erosión de

la confianza¹⁵. La desconfianza y la fragmentación social se combinan para intensificar hostilidades, inseguridad y polarización, dibujando un escenario donde la inestabilidad pretende consolidarse como norma y socavando el proyecto democrático, inseparable del principio de pluralismo¹⁶.

En este sentido, la reflexión sobre la situación monopolística de las grandes plataformas y de las empresas que las dirigen es alentada por la propia Unión Europea, quien en el Considerando 4 del Reglamento (UE) 2024/1083 sobre libertad de medios de comunicación (EMFA en sus siglas en inglés)¹⁷ llama la atención sobre las plataformas en línea de alcance mundial «actúan como puertas de entrada a los contenidos de los medios de comunicación, con modelos de negocio que tienden a eliminar la intermediación para el acceso a los servicios de medios de comunicación y a amplificar los contenidos polarizadores y la desinformación».

El peligro de que la esfera pública habermassiana está gestionada por las grandes corporaciones de la comunicación, que son también las que la controlan¹⁸ (asumiendo como indica Lanier¹⁹ que «la arquitectura de la red digital incuba monopolios de forma natural»), lleva a Lasalle²⁰ a exigir que «no pueden seguir siendo soberanos en la sombra de la sociedad digital. Su diseño ha de aflorar a la superficie legal. No puede seguir sumido en un secretismo que esconda los sesgos que introduce en los perfiles que determina».

Pero dicha imposibilidad de comprensión de los algoritmos permite proteger el funcionamiento de las grandes compañías tecnológicas. Si las normativas comunitarias obligan a un banco a explicar la decisión del algoritmo, dicha explicación no llegará en forma de una única frase, sino que lo más probable es que llegue en forma de cientos e incluso miles de páginas llenas de número y ecuaciones.

Para defender nuestras democracias del poder omnímodo algorítmico, Harari propone unos pocos, pero ancestrales principios: *Benevolencia*. Cuando una red informática recopila información sobre mí, esta información ha de usarse para ayudarme, no para manipularme. Tener acceso a nuestra vida personal conlleva el deber fiduciario de actuar en defensa de nuestros intereses y no aprovecharse de esa información. ¿Por qué no extender un principio tan obvio y antiguo a ordenadores y algoritmos, empezando por los poderosos algoritmos de Google, Baidu o TikTok? Si exigimos confidencialidad a nuestro abogado, médico o sacerdote, ¿por qué no a Musk o Zuckerberg?. *Descentralización*. Una sociedad democrática no debe permitir que toda su información se halle concentrada en unos pocos lugares, con independencia de que dicho centro sea el Gobierno o una compañía privada. Cuantos menos archivos, más control. Y *Mutualidad*. Es clave que la

información fluya en sentido bidireccional, no sólo de abajo (nosotros) a arriba (Gobiernos y corporaciones). Si ellos saben mucho sobre nosotros, nosotros debemos saber mucho sobre ellos.

¿Y dónde queda Europa? Ante la tesitura de tener que elegir entre vulnerabilidades o dependencias como indica José Ignacio Torreblanca. Ante este escenario y abocados a la irrelevancia tecnológica, la posición de Europa resulta especialmente relevante, convirtiéndonos los europeos en paladines de los derechos y el Estado de Derecho, posición que no requiere de grandes inversiones sino precisamente de palabras. La Unión Europea está desarrollando un proceso de re-territorialización de la plataformización, buscando materializar el denominado «*efecto Bruselas*»²¹ mediante un modelo de constitucionalismo digital europeo, tal como plantean autores como Celeste o Di Gregorio. A través de instrumentos regulatorios como el Reglamento de Servicios Digitales (DSA) y el Reglamento de Mercados Digitales (DMA), y en menor medida el Reglamento de Inteligencia Artificial, la UE pretende establecer líneas rojas al poder de las grandes plataformas tecnológicas.

Estas normas introducen obligaciones de transparencia y rendición de cuentas (DSA), así como mecanismos de control derivados del derecho de la competencia (DMA), con el objetivo de limitar posiciones dominantes y prácticas abusivas. De este modo, Europa aspira a consolidarse como un actor central en el futuro digital global, defendiendo un modelo basado en la protección de los derechos fundamentales y el Estado de Derecho, en contraste tanto con el enfoque estadounidense, de corte más ultraliberal y con una creciente imbricación entre grandes corporaciones tecnológicas y el Gobierno –con el caso Palantir como paradigma–; como con el modelo chino, caracterizado por un mayor control estatal. En nuestros Estados de Derecho, la incompreensión es letal. Si ciudadanos, legisladores, periodistas y jueces no pueden entender como funciona la IA, no podrán supervisarla y perderán la fe en su regulación, dejando el campo libre a un sinfín de decisiones que les afectaran.

Por eso, como nos dice Ana Palacio, el debate europeo no puede circunscribirse al ciudadano como consumidor de IA. Hay que protegerlo también como miembro de una comunidad política que aspira a conservar autonomía. La privacidad importa. La transparencia importa. La no discriminación importa. Pero no cabe sacrificar la seguridad, la resiliencia, el tejido industrial, la educación, el talento, la infraestructura y la respuesta en crisis²².

Graves nubarrones se ciernen sobre el horizonte («The Winter is coming»). El grave riesgo al que nos enfrentamos respecto a la IA no es sólo actual, sino en

un futuro cercano. La Inteligencia Artificial General –IAG– (AGI en sus siglas en inglés), capaz de superar a los humanos no jugando al *go* o al ajedrez, sino en la gestión de situaciones complejas pertenecientes al mundo real, o incluso diseñando una IAG aún mayor, está más cercana de lo que creemos, datada por algunos para 2030²³. La cuestión que tenemos que plantearnos es si queremos que el desarrollo de esa nueva realidad se concentre en unas pocas manos sin ningún control de los gobiernos y con el desconocimiento de la ciudadanía.

El riesgo de una IA autónoma que tan fructíferamente ha alimentado la ciencia ficción, empieza a ser cada vez más ciencia y menos ficción. Quizá sea el momento de volver a poner de actualidad las leyes de la robótica de Isaac Asimov y aplicarlas al funcionamiento de los algoritmos. Las Tres Leyes de la Robótica de Isaac Asimov, formuladas originalmente en el relato *Runaround* (1942), pueden resumirse así: 1) Un robot no debe dañar a un ser humano ni, por inacción, permitir que un ser humano sufra daño; 2) Un robot debe obedecer las órdenes de los seres humanos, salvo cuando esas órdenes entren en conflicto con la Primera Ley; y 3) Un robot debe proteger su propia existencia, siempre que esa protección no entre en conflicto con la Primera o la Segunda Ley. Posteriormente Asimov introdujo una Ley Cero, situada jerárquicamente por encima de las anteriores, un robot no debe dañar a la humanidad ni, por inacción, permitir que la humanidad sufra daño. Simplemente tenemos que sustituir la palabra robot por las palabras Inteligencia Artificial.

Desde una perspectiva jurídica, se han planteado soluciones que parten del principio básico de que las decisiones de los sistemas de IA que afectan de forma significativa a las personas sean siempre supervisadas por un ser humano. (El diablo como siempre está en los detalles, ¿qué entendemos por significativa y quién decide qué es significativo?). La intensidad de la intervención humana permite distinguir entre diferentes modelos de supervisión. En los sistemas *human-in-the-loop* (HITL), la actuación del sistema queda condicionada a la aprobación humana, mientras que en los modelos *human-on-the-loop* (HOTL) el sistema puede ejecutar autónomamente sus decisiones, permaneciendo el ser humano en una posición de supervisión desde la que puede intervenir o vetar su actuación²⁴. Los matices para emplear una u otro serían también objeto de discusión.

En la novela de Mary Shelley, publicada en 1818 con el significativo subtítulo *The Modern Prometheus*, el verdadero problema no reside únicamente en que Victor Frankenstein consiga crear vida, sino en que pone en marcha una realidad cuyas consecuencias no sabe anticipar ni controlar y de la que, además, intenta después desentenderse. La tragedia comienza precisamente cuando la capacidad técnica del creador supera su capacidad para asumir la responsabilidad sobre lo

creado. Algo semejante puede plantearse hoy respecto de la inteligencia artificial. Desarrollamos sistemas cada vez más autónomos, capaces de aprender, generar contenidos, adoptar decisiones y actuar sobre el entorno, mientras seguimos discutiendo quién debe responder por sus consecuencias y qué mecanismos permitirán mantenerlos bajo control. En este sentido, *Frankenstein* no constituye solo una advertencia contra la tecnología, sino contra una innovación desvinculada de la responsabilidad. Dos siglos después, la pregunta de Shelley conserva una inquietante actualidad. No se trata únicamente de saber si somos capaces de crear una inteligencia cada vez más poderosa, sino de preguntarnos si seremos capaces de gobernar aquello que hemos creado.

Las palabras son nexos entre las personas, son la argamasa que unen a las personas y crean civilizaciones y sociedades. Si algo nos enseña la historia es que el poder nunca ha sido indiferente a la palabra. La ha prohibido, la ha censurado, la ha manipulado, la ha monopolizado. Pero también ha intentado algo quizá más eficaz, decidir qué palabras pueden circular, cuáles pueden ser escuchadas y cuáles quedan sepultadas entre millones de mensajes.

Hoy ese poder ya no pertenece únicamente al Estado. Se reparte entre gobiernos, plataformas, algoritmos y sistemas de inteligencia artificial capaces de ordenar la conversación pública, amplificar unas voces y hacer prácticamente invisibles otras.

Por eso, el gran desafío de nuestro tiempo quizá no sea solamente garantizar que podamos hablar. Será garantizar que sigamos pudiendo decidir qué decir, que podamos comprender quién condiciona lo que escuchamos y, sobre todo, que ninguna arquitectura tecnológica termine sustituyendo nuestra capacidad de pensar por nosotros mismos.

Porque la libertad de expresión nunca ha consistido únicamente en poder pronunciar palabras. Consiste también en conservar el poder sobre ellas. Como decía el juez Learned Hand en 1944, «¿cuál es, pues, el espíritu de la libertad? Yo no puedo definirlo. Sólo puedo decirles cuál es mi opinión. El espíritu de la libertad es el espíritu que no está demasiado seguro de estar en lo cierto; el espíritu de la libertad es el espíritu que busca entender los pensamientos de otros hombres y mujeres; el espíritu de la libertad es el espíritu que sopesa los intereses de esos hombres y mujeres y los propios sin parcialidad»²⁵.

Y quizá esa sea la última frontera de la libertad en la sociedad digital. Que, en un mundo en el que cada vez más máquinas hablan, seleccionan, recomiendan e incluso deciden, la palabra siga perteneciendo al ser humano y no el ser humano a quien controla la palabra. Vale.

NOTAS

Parte primera

- 1 «La historia del constitucionalismo no es sino la búsqueda por el hombre político de las limitaciones al poder absoluto ejercido por los detentadores del poder, así como el esfuerzo de establecer una justificación espiritual, moral o ética de la autoridad, en lugar del sometimiento ciego a la facticidad de la autoridad existente». Loewenstein, K. (1976). *Teoría de la Constitución* (A. Gallego Anabitarte, Trad.; 2.ª ed.). Ariel, p. 150.
- 2 Para Berelson y Steiner, «la comunicación consiste en la transmisión de información, ideas, emociones y habilidades mediante símbolos, entre los que incluyen expresamente las palabras, además de imágenes, figuras y gráficos». Berelson, B., y Steiner, G. A. (1964). *Human behavior: An inventory of scientific findings*. Harcourt, Brace y World, p. 254.
- 3 Ya en la célebre Oración fúnebre de Pericles, transmitida por Tucídides, la democracia aparece caracterizada por el gobierno de la mayoría, la igualdad jurídica y la posibilidad de participar en los asuntos públicos atendiendo al mérito y no a la posición social (Tucídides. (1990). *Historia de la Guerra del Peloponeso. Libros I-II* (J. J. Torres Esbarranch, Trad. y notas). Gredos, pp. 451-452).
- 4 Gil, L. (2007). *Censura en el mundo antiguo*. Alianza Editorial, p.62.
- 5 Las discusiones de todos los ciudadanos se producían en la asamblea (*ekklesia*) que reunía en el Pnyx a todos los ciudadanos que desearan asistir, aproximadamente una vez a la semana. El uso de la palabra era libre, pero la realidad es que los ciudadanos que quería plantear alguna propuesta contrataban a los *rhétores*, término que se traduce

directamente como «orador» o «aquel que habla en público». No se trataban de cargos electos, sino ciudadanos particulares con gran elocuencia que influían en las decisiones del pueblo. Cualquier ciudadano varón tenía el derecho legal de hablar en la asamblea (*isegoría*), pero los *rhétores* eran quienes habitualmente tomaban la palabra debido a su preparación en retórica.

- 6 González Ballesteros, T. (2014). *Democracia y comunicación (reflexiones): Lección inaugural, curso académico 2013–2014*. Universidad Complutense de Madrid.
- 7 Durante el siglo V, la Asamblea ostentó el poder máximo de la ciudad en todos sus sentidos. El poder judicial secundario lo ostentaba un Tribunal popular de arcontes cuyos miembros eran elegidos por sorteo entre los ciudadanos que hubieran realizado cada año el juramento helenístico. Las decisiones de la Asamblea eran ejecutadas por un cuerpo de Magistrados elegidos mayoritariamente por sorteo anual y no se podía repetir en el cargo. El poder ejecutivo descansaba en el Consejo de los Quinientos (*boulé*) que era elegido por sorteo entre los ciudadanos de la polis, con poderes administrativos para gestionar la ciudad y capacidad de dictar decretos. (Urías, J. (2021). ¿Quién quiere hablar? La problemática de la libertad de expresión en la Atenas clásica. *Revista de Derecho Político* (110), 153–184, p. 156)
- 8 «La democracia ateniense, surgida en 508 a. C., atravesó numerosos conflictos bélicos y, pese a recuperarse de la derrota frente a Esparta en 404 a. C., terminó finalmente en 322 a. C., tras la derrota ante Macedonia y su sustitución por un régimen oligárquico» (Tridimas, G. (2015). War, disenfranchisement and the fall of the ancient Athenian democracy. *European Journal of Political Economy*, 38, 102–117. <https://doi.org/10.1016/j.ejpoleco.2015.01.002>, pp. 102-103).
- 9 Wirszubski, C. H. (1950). *Libertas as a political idea at Rome during the Late Republic and Early Principate*. Cambridge University Press, pp. 7-30
- 10 Díaz de Valdés, J. M. (2009). Libertad de expresión en Roma. *Revista de Estudios Histórico-Jurídicos*, 31, 125–139. <https://doi.org/10.4067/S0716-54552009000100004>, pp. 132-136.
- 11 Bauman, R. A. (2000). *Human rights in ancient Rome*. Routledge, p. 107.
- 12 McHugh, M. R. (2004). Historiography and freedom of speech: The case of Cremutius Cordus. En I. Sluiter y R. M. Rosen (Eds.), *Free speech in classical antiquity* (pp. 391–408). Brill.
- 13 Morton Braund, S. (2004). Libertas or licentia? Freedom and criticism in Roman satire. En I. Sluiter y R. M. Rosen (Eds.), *Free speech in classical antiquity* (pp. 409–428). Brill, p. 426.

- 14 Watts, E. J. (2014). Introduction: Freedom of speech and self-censorship in the Roman Empire. *Revue Belge de Philologie et d'Histoire*, 92(1), 157-166. <https://doi.org/10.3406/rbph.2014.8545>
- 15 Świątoń, A. (2019). Roman law and riots on a religious background: The example of constitutions preserved in *Codex Theodosianus* 16.4, *De his, qui super religione contendunt*. *Studia Warmińskie*, 56, 375-389. <https://doi.org/10.31648/sw.3922>
- 16 Un hito fundamental es el *Decretum Gratiani*, hacia 1140, y posteriormente las Decretales de Gregorio IX de 1234. Con ellos se sistematiza progresivamente el Derecho canónico. La Iglesia desarrolla mecanismos jurídicos para determinar qué doctrinas pueden sostenerse públicamente y cuáles constituyen desviaciones de la ortodoxia. La palabra no es considerada jurídicamente indiferente, ya que afirmar, enseñar o difundir una determinada doctrina puede constituir en sí mismo un ilícito. El IV Concilio de Letrán (1215) constituye un momento fundamental en el reforzamiento institucional de la lucha contra la herejía. Lo importante desde la perspectiva de la libertad de expresión es que la preocupación jurídica no se refiere únicamente a lo que una persona cree, sino también a lo que predica, enseña y difunde. Aquí aparece ya una cuestión que tendrá una enorme continuidad histórica como es la distinción entre pensamiento y exteriorización del pensamiento.
- 17 Paradójicamente, constituyen espacios extraordinarios de discusión intelectual mediante la *disputatio* y, al mismo tiempo, espacios sometidos al control de la ortodoxia. Un ejemplo paradigmático son las condenas de París de 1270 y, especialmente, de 1277, cuando el obispo Étienne Tempier condenó 219 proposiciones que se enseñaban o discutían en el ámbito universitario parisino.
- 18 Madero, M. (1992). *Manos violentas, palabras vedadas: La injuria en Castilla y León (siglos XIII-XV)*. Taurus.
- 19 «*Injuria* en latín tanto quiere decir en romance como deshonra que es fecha ó dicha á otri á tuerto ó á despreciamiento dél. Et como quier que muchas maneras son de deshonra, pero todas descenden de dos raíces: la primera es de palabra; la segunda de fecho. Et de palabra es asi como si un home denostase á otro ó le diese voces ante muchos, haciendo escarnio dél ó poniéndole algunt nombre malo, ó diciendo en pos dél palabras atales onde se toviere el otro por deshonorado. Eso mismo decimos que serie si ficiese esto facer á otros asi como á los rapacesó á otros qualesquier. La otra manera es quando dixiese mal dél antemuchos, razonándolo mal, ó enfamándolo de algunt yerro ó denostándolo». (Alfonso X, *Las Siete Partidas*, Partida VII, título IX, ley I, en *Las Siete Partidas del Rey Don Alfonso el Sabio, cotejadas con varios códices antiguos*

por la Real Academia de la Historia, t. III, Madrid, Imprenta Real, 1807; ed. facsímil, Madrid, BOE-Real Academia de la Historia, 2021).

- 20 Collins, R. (1990). Literacy and the laity in early medieval Spain. En R. McKitterick (Ed.), *The uses of literacy in early medieval Europe*. Cambridge University Press. Sánchez Prieto, A. B. (2010). Dónde aprender a leer y escribir en el año mil. *Anuario de Estudios Medievales*, 40(1). <https://doi.org/10.3989/aem.2010.v40.i1.294>. Para esta autora, con la desaparición de las escuelas municipales romanas, quienes mantuvieron la enseñanza de la lectura y escritura fueron las instituciones eclesíásticas –escuelas monásticas, episcopales y parroquiales–, que aunque destinadas principalmente a la formación del clero, también acogieron en ocasiones a niños laicos.
- 21 La Universidad de Bolonia (Italia), fundada en 1088, es la primera en usar el término *universitas* y el modelo de educación superior de Occidente; la Universidad de Oxford (Reino Unido) fue fundada alrededor de 1096; la Universidad de Salamanca (España) en 1134, mientras que la Universidad de París fue fundada hacia 1150.
- 22 Bataillon, L. J., Guyot, B. G., y Rouse, R. H. (Eds.). (1988). *La production du livre universitaire au Moyen Âge: Exemplar et pecia*. CNRS.
- 23 Las fechas exactas de la aparición de la imprenta son aproximadas como en toda invención que no tiene un momento «¡Eureka!». Así sobre 1440, Gutenberg comienza a desarrollar su técnica de impresión con tipos móviles. En torno a 1450 comienza a funcionar su taller de impresión en Maguncia, pero será en 1455 cuando se produzca la impresión de la célebre Biblia de Gutenberg, la primera gran obra producida mediante este procedimiento en Europa. Febvre, L., y Martin, H.-J. (2005). *La aparición del libro*. Fondo de Cultura Económica. (Obra original publicada en 1958). La versión original francesa de 1958, *L'apparition du livre*, publica en París por Albin Michel.
- 24 Se estima que en los cuarenta seis años que van de 1454 a 1500 se imprimieron en Europa más de doce millones de volúmenes. En cambio, en los mil años previos solo se había copiado a mano unos once millones de volúmenes. Buringh, E., y Van Zanden, J. L. (2009). Charting the «Rise of the West»: Manuscripts and printed books in Europe, a long-term perspective from the sixth through eighteenth centuries. *The Journal of Economic History*, 69(2), 409–445. DOI: 10.1017/S0022050709000837.
- 25 En Ley promulgada por los mismos monarcas en Toledo en el año 1480, se eximía del pago de almojarifazgo, portazgo y otros derechos a los libros que se introdujeran a sus reinos por tierra o por mar. (Cendán Pazos, F. (1974). *Historia del Derecho español de prensa e imprenta (1502–1966)*. Editora Nacional, p. 23).

- 26 Que en este caso eran los presidentes de las Audiencias de Valladolid y Granada, los arzobispos de Toledo, Sevilla y Granada y los obispos de Burgos y Salamanca.
- 27 Carlos V mediante las ordenanzas del Consejo redactadas en La Coruña en 1554, atribuyó el carácter exclusivo de otorgar licencias de impresión al presidente y demás miembros del Consejo Real. Felipe II mantuvo dicha centralización en el Consejo Real del control de las impresiones al promulgar en Valladolid, y en su nombre, la princesa doña Juana, la Pragmática de 7 de septiembre de 1558.
- 28 Será el Decreto de Carlos V, dictado en Aranjuez el 11 de abril de 1805 e inserto en la Real Cédula del Consejo del día 3 de mayo de ese año, quien otorgue al Juzgado Privativo de Imprentas y Librerías todo el poder sobre licencias de impresión, que sin embargo no eliminaba la autorización exclusiva en manos de monarca para otorgar licencias de publicación de nuevos periódicos.
- 29 Ese control también se produjo en las Indias, donde la Real Cédula de Felipe II de 21 de septiembre de 1556, exhortaba a justicias de España y de la Indias que no consintieran la impresión, posesión y venta de libro alguno que tratase asuntos relacionados con dichos territorios, salvo licencia previa del Real Consejo de Indias. Como ejemplo, estaban prohibidas la *Historia General de las Indias* de López de Gómora, o la *Historia* del inca Garcilaso de la Vega, así como la impresión de vocabularios en lengua indígena.
- 30 De los Reyes Gómez, F. (2000). *El libro en España y América: Legislación y censura (siglos XV–XVIII)* (Vol. 1). Arco/Libros, p. 142.
- 31 Así Carlos V solicita a la Santa Sede la expedición de una bula para obtener de la Universidad de Lovaina una lista de los libros «cuya introducción en los dominios españoles pudiera ser peligrosa». Impresa en 1546, puede considerarse el primer Índice expurgatorio en nuestro país, ampliada en su segunda edición de 1550 por el Supremo Consejo de la Inquisición, que publica en 1554 la *Censura General contra los errores de los herejes modernos*. También son destacables el índice de Fernando de Valdés de 1559, o posteriormente, el Índice de Quiroga de 1583-1584 que introdujo de forma más desarrollada la distinción entre prohibición (la obra no podía ser leída) y expurgación (la obra podía circular una vez eliminados determinados pasajes).
- 32 Pinto Crespo, V. (1983). Los índices de libros prohibidos. *Hispania Sacra*, 35(71), 93–130.
- 33 Rendler, P. F. (1977). *The Roman Inquisition and the Venetian press, 1540–1605*. Princeton University Press, especialmente pp. 73-116 y 145-183.
- 34 Birn, R. (2012). *Royal censorship of books in eighteenth-century France*. Stanford University Press.

- 35 Robertson, R. (2009). *Censorship and conflict in seventeenth-century England: The subtle art of division*. Pennsylvania State University Press, especialmente pp. 17-21.
- 36 Clegg, C. S. (2008). *Press censorship in Caroline England*. Cambridge University Press, especialmente pp. 186-207.
- 37 Deazley, R. (2008). Commentary on the Star Chamber Decree 1637. En L. Bently y M. Kretschmer (Eds.), *Primary sources on copyright (1450–1900)*. Arts and Humanities Research Council.
- 38 Las cotas más altas de persecución de la libertad de imprenta se producen con la *Licensing Act* de 1662. En ella, no se permiten abrir los embalajes en los que se transportan libros, si no es en presencia de personas autorizadas por la *Company of Stationers*, se limita el número y localización de las librerías, se establece la necesidad de permisos para instalar imprentas, así como un número máximo de veinte maestros impresores para todo el país, y la prohibición de que los impresores tengan más de dos máquinas cada uno, y más de dos trabajadores o aprendices. (Ansúategui Roig, F. J. (1994). *Orígenes doctrinales de la libertad de expresión*. Universidad Carlos III de Madrid–Boletín Oficial del Estado, pág. 246).
- 39 De forma enumerativa, aunque no exhaustiva, podemos encontrar el sustento ideológico de esta nueva realidad política en obras como el *Tratado teológico-político* de Baruch Spinoza en 1670. En el capítulo XX, Spinoza parte de la imposibilidad de someter completamente el pensamiento humano al poder político y sostiene que, en un Estado libre, cada individuo debe poder pensar lo que quiera y expresar aquello que piensa. Pocos años después, John Locke, en su *Carta sobre la tolerancia* de 1689, contribuirá a delimitar una esfera de conciencia en la que el poder político carece de legitimidad para imponer determinadas convicciones. Su tolerancia todavía presenta límites importantes, pero introduce un principio que será esencial para el constitucionalismo liberal, que existen ámbitos de la conciencia individual sobre los que el poder político no puede disponer libremente. Durante el siglo XVIII, estas ideas adquirirán una dimensión cada vez más política. Resultan especialmente significativas las *Cato's Letters*, publicadas por John Trenchard y Thomas Gordon entre 1720 y 1723 y ampliamente difundidas posteriormente en las colonias norteamericanas. Una de ellas llevaba un título que prácticamente resume toda una concepción de la libertad política: *Of Freedom of Speech: That the same is inseparable from Publick Liberty*. La libertad de expresión deja aquí de justificarse únicamente como protección de la conciencia individual. Adquiere una segunda función fundamental como es permitir a los ciudadanos conocer, examinar y criticar la actuación del poder. También la Ilustración continental contribuirá a este proceso. Por su parte, Montesquieu (1748) en

el *De l'esprit des lois*, especialmente en el libro XI, capítulo VI, desplaza parcialmente la cuestión desde la proclamación de derechos hacia las condiciones institucionales necesarias para garantizar la libertad. La concentración del poder amenaza la libertad; por ello es necesario articular mecanismos institucionales que impidan su abuso. Su influencia aparece con extraordinaria claridad en el art. 16 de la Declaración francesa: «Toute Société dans laquelle la garantie des Droits n'est pas assurée, ni la séparation des Pouvoirs déterminée, n'a point de Constitution». El *Tratado sobre la tolerancia* de Voltaire, publicado en 1763 a raíz del caso Calas, convirtió la lucha contra el fanatismo y la intolerancia religiosa en una reivindicación de la autonomía de la conciencia y de la posibilidad de manifestar públicamente las propias convicciones. Cesare Beccaria en *De los delitos y de las penas* (1764), proporciona buena parte del sustrato ilustrado de la reforma penal: legalidad, proporcionalidad, humanización de las penas, crítica de la tortura y rechazo de la arbitrariedad judicial.

- 40 Muñoz-Alonso Ledo, A. (1992). Historia y fundamentos de la opinión pública. En A. Muñoz Alonso, C. Monzón, J. I. Rospir y J. L. Dader, *Opinión pública y comunicación política*. Eudema.
- 41 El caso inglés es diferente. Ciertamente es que un siglo antes de la declaración francesa de 1789, en Inglaterra se promulgó el Bill of Rights (1689). No obstante, no se trata de una declaración de derechos, sino de una limitación del poder regio frente al Parlamento. No obstante, en su artículo XIII se encuentra un reconocimiento de la libre discusión parlamentaria que podemos catalogar más como un antecedente de la inviolabilidad parlamentaria, que de la libertad de expresión. El punto de inflexión se produjo en 1695, cuando el Parlamento dejó expirar la legislación sobre licencias. Ello posibilitó publicar sin autorización del poder, aunque continuara existiendo responsabilidad posterior por determinados contenidos, particularmente mediante la doctrina del *seditious libel*. William Blackstone en el libro IV de sus *Commentaries on the Laws of England* formula una distinción que ejercerá una enorme influencia posterior, que la libertad de prensa consiste esencialmente en la inexistencia de restricciones previas a la publicación y no en la inmunidad frente a las consecuencias jurídicas derivadas de contenidos ilícitos. (Blackstone, *Commentaries on the Laws of England*, vol. IV, 1769, pp. 151-152).
- 42 Levy, L. W. (1985). *Emergence of a free press*. Oxford University Press, especialmente pp. 62 y ss.
- 43 A quien se le atribuye la famosa frase «Si tuviera que decidir si debemos tener un gobierno sin periódicos o periódicos sin gobierno, no dudaría en preferir lo segundo».

- 44 Las referencias a sentencias señeras del Tribunal Supremo estadounidense en defensa de la libertad de palabra podrían dar lugar a un trabajo específico por su riqueza y trascendencia. Probablemente, la más citada sea New York Times contra Sullivan de 1964, donde aparece el célebre *dictum* de que «el compromiso nacional profundo respecto al principio que postula que el debate sobre los asuntos públicos debe ser desinhibido, robusto y amplio, aunque incluya ataques a funcionarios públicos que puedan ser vehementes, caústicos o desagradables». Sobre el tema que nos ocupará en próximas páginas, en el caso Reno contra ACLU de 1997 alabó como Internet permite a «cualquier persona con una línea telefónica, convertirse en panfletista o en un pregonero con una vez que resuena más lejos de lo que hubiera logrado desde cualquier tribuna». La defensa de la autonomía de los medios se consagra en la sentencia Columbia Broadcasting System, Inc. contra Democratic National Committee de 1973, donde reconoce que «la Primera Enmienda otorga a “todo periódico la libertad para publicar lo que quiera y rechazar lo que no quiera, libre de la injerencia intrusiva del Gobierno [...] Para bien o para mal, los editores existen para editar; y editar es selección y elección de contenido. El que los editores, ya sea de periódicos o radioemisores, puedan y de hecho abusen de esta facultad es algo indiscutido, pero ello no es razón para negarles la discreción que le otorgó el Congreso». En dicha sentencia, el juez Douglas admite la existencia de la censura privada en el periodismo, «pero para cada editor que suprima un hecho, hay varios que lo publicarán. Pero si el Gobierno es el censor, será un mandato administrativo, y no la libertad de elección, lo que predomine». El Tribunal Supremo considera que al optar entre la censura por un actor privado y uno gubernamental, la elección siempre debe favorecer al privado sobre el público. Esa idea fue reforzada un año después en el caso Miami Herald Publishing contra Tornillo, anulando por unanimidad una norma sobre derecho de réplica del Estado de Florida porque ello constituía una intromisión en la función de los editores, que son quienes tienen la exclusiva competencia sobre los contenidos de su medio de comunicación. En su sentencia Packingham v. North Carolina, de 19 de junio de 2017, el Tribunal Supremo ha afirmado en relación con las redes sociales que son «las principales fuentes para conocer la actualidad, consultar anuncios de empleo, participar en el debate público moderno y explorar los vastos ámbitos del pensamiento y el conocimiento humanos. Estos sitios web ofrecen quizás los mecanismos más poderosos a disposición del ciudadano para hacer oír su voz. Permiten a cualquier persona con conexión a internet convertirse en un pregonero con una voz que resuena mucho más lejos que desde cualquier tribuna».
- 45 Dicha estructura reconociendo el derecho, para posteriormente establecer sus límites, ha aparecido de forma guadianesca en la regulación de la libertad de expresión en las declaraciones internacionales de derechos. No aparece en el artículo 19 de la

Declaración Universal de Derechos Humanos de 1948 ni en el artículo 11 de la Carta de Derechos Fundamentales de la Unión Europea del año 2000. Sí aparece en art. 10 del Convenio Europeo de Derechos Humanos y Libertades Públicas de 1950 y en el art. 19 del Pacto Internacional de Derechos Civiles y Políticos de 1966. Un análisis pormenorizado de esas declaraciones en nuestra monografía de 2020, *La libertades informativas en el ámbito internacional* (Prólogo de Marcelino Oreja Aguirre). Dykinson.

- 46 Derieux, E. (1998). *Droit des médias: Droit français, européen et international*. LGDJ;
- Morange, J. (1993). *La liberté d'expression*. Presses Universitaires de France
- 47 Bel Mallén, J. I. (1990). La libertad de expresión en los textos constitucionales españoles. *Documentación de las Ciencias de la Información*, 13, 23–52.
- 48 Fernández Sarasola, I. (2006). Opinión pública y «libertades de expresión» en el constitucionalismo español (1726–1845). *Historia Constitucional*, (7), especialmente pp. 23 y ss. Sin embargo, para alguna autora, el artículo 371 ofrecía todavía una protección limitada, circunscrita a «escribir, imprimir y publicar», a diferencia de las formulaciones más amplias que podían encontrarse ya en el constitucionalismo francés y norteamericano (Esquivel Alonso, Y. (2016). Sueños de libertad y censura: La libertad de imprenta de 1810. *Revista de Estudios Políticos*, (174), 143–172).

Parte segunda

- 1 En 1996, John Perry Barlow, antiguo letrista de la banda *Grateful Dead* y creador de *Electronic Frontier Fundation* (EFF) presentó la Declaración de Independencia del Ciberespacio donde decía «Estamos creando un mundo en el que todos pueden entrar, sin privilegios o prejuicios debidos a la raza, el poder económico, la fuerza militar, o el lugar de nacimiento. Estamos creando un mundo donde cualquiera, en cualquier sitio, puede expresar sus creencias, sin importar lo singulares que sean, sin miedo a ser coaccionado al silencio o al conformismo».
- 2 «La posibilidad de que los individuos se expresen a través de Internet constituye una herramienta sin precedentes para el ejercicio de la libertad de expresión» (Sentencia del TEDH de 16 de junio de 2015, asunto Delfi AS c. Estonia)
- 3 Keen, A. (2007). *The cult of the amateur: How today's Internet is killing our culture*. Doubleday/Currency.
- 4 Lanier, J. (2010). *You are not a gadget: A manifesto*. Alfred A. Knopf. Ambas citas en la obra Garton Ash, T. (2017). *Libertad de palabra: Diez principios para un mundo conectado*. Tusquets, p. 251.

- 5 Revenga Sánchez, M. (2023). Dicotomías de la libertad de expresión en la era de la comunicación digital. En E. Ferrer Mac-Gregor (Coord.), *La garantía jurisdiccional de la Constitución. A cien años del Verfassungsgerichtshof Österreich, a cuarenta años del Tribunal Constitucional de España. XII Encuentro Iberoamericano de Derecho Procesal Constitucional* (Vol. II, pp. 1171–1181). Centro de Estudios Políticos y Constitucionales, p. 1172.
- 6 Como indica el Tribunal Constitucional español, «Las difamaciones e insultos, las expresiones objetivamente injuriosas, la proliferación de discursos que fomentan el odio, la violencia o la discriminación, la formulación de ideas o pensamientos que pueden constituir una injerencia en otros derechos fundamentales –como son el derecho al honor o el derecho al respeto de la vida privada y familiar– pueden ser difundidos como nunca antes se había visto, garantizando a todos su acceso y redifusión a terceros en cuestión de segundos, cualquiera que sea el lugar del mundo en el que se encuentre, con la facultad, incluso, de “permanecer accesibles en la web de forma constante” (STEDH de 16 de junio de 2015, asunto *Delfi AS c. Estonia*, § 133). Ambas ideas: multiplicación y permanencia en el tiempo, muestran por sí mismas la necesidad de proteger de forma eficaz los derechos eventualmente afectados. No puede, por tanto, aceptarse, tal y como propone la mercantil recurrente, la consideración genérica de internet como un contexto comunicativo puramente ocasional o trivial» (STC 83/2023, FJ 5º). En un mismo sentido se pronuncia el Tribunal Europeo de Derechos Humanos considerando que «Internet es una herramienta de información y comunicación particularmente distinta de los medios impresos, especialmente en lo que se refiere a la capacidad de almacenar y transmitir información [...] El riesgo de perjuicio que suponen los contenidos y las comunicaciones en internet para el ejercicio y disfrute de los derechos humanos y las libertades, en particular el derecho al respeto de la vida privada, es ciertamente mayor que el que plantea la prensa [escrita]» (STEDH de 5 de mayo de 2011, asunto *Consejo Editorial de Pravoye Delo y Shtekel c. Ucrania*, apdo. 63).
- 7 «El riesgo para los derechos de la personalidad de las personas afectadas por la información guardada en la hemeroteca digital no radica tanto en que la información sea accesible a través del motor de búsqueda interno del sitio web en que se encuentra alojada, pues se trata de una búsqueda comparable a la que efectuaban quienes acudían a las viejas hemerotecas en papel, como en la multiplicación de la publicidad que generan los motores de búsqueda de Internet, y en la posibilidad de que mediante una simple consulta utilizando los datos personales, cualquier internauta pueda obtener un perfil completo de la persona afectada en el que aparezcan informaciones obsoletas sobre hechos ya remotos en la trayectoria vital del afectado, con un grave potencial dañoso para su honor y su intimidad, que tengan un efecto distorsionador de

la percepción que de esta persona tengan los demás conciudadanos y le estigmatice». (STS 545/2015, 15 de octubre, FJ 7º)

- 8 El conocido como caso Mario Costeja constituye el origen jurisprudencial europeo del denominado «derecho al olvido» en Internet. Mario Costeja González solicitó que, al introducir su nombre en Google, dejaran de aparecer enlaces a antiguas noticias de *La Vanguardia* relativas a un embargo por deudas con la Seguridad Social, ya resuelto años atrás. El TJUE consideró que la actividad de un motor de búsqueda –localizar, indexar, almacenar y mostrar información vinculada a una persona– constituye un tratamiento de datos personales y que el gestor del buscador es responsable de dicho tratamiento. En consecuencia, la sentencia reconoció que una persona puede solicitar la supresión de determinados enlaces de los resultados obtenidos mediante la búsqueda de su nombre, incluso cuando la información publicada originalmente sea lícita, cuando resulte inadecuada, no pertinente o haya dejado de ser pertinente por el transcurso del tiempo. Este derecho debe ponderarse, no obstante, con el interés del público en acceder a la información, especialmente cuando el afectado desempeñe un papel relevante en la vida pública. (STJUE (Gran Sala), de 13 de mayo de 2014, *Google Spain, S.L. y Google Inc. c. Agencia Española de Protección de Datos (AEPD) y Mario Costeja González*, asunto C-131/12, ECLI:EU:C:2014:317)
- 9 «La constatación de las dificultades existentes para distinguir entre libertad de información y libertad de expresión se acentúa cuando se contextualiza el ejercicio de una y de otra en el ámbito de internet y, más concretamente, en el de las redes sociales. (STC 8/2022, FJ 2º)
- 10 El principio de aislamiento, que es cuando en cualquier familia, grupo de amigos, o incluso pareja comiendo juntos, todo mundo está mirando su teléfono móvil, quedando bastante claro que el móvil es un objeto unipersonal que aísla de los demás. El segundo principio es el infinito, que es la cantidad de contenido al que se tiene acceso, es decir, películas infinitas, las series nunca se acaban, titulares infinitos, zapatos infinitos en la tienda mientras intentas comprarte unos, es decir, tú te acabas, tú te duermes, te cansas, tienes cosas que hacer, pero el contenido es una cascada completamente interminable que te genera esta sensación de que en el momento en el que no estés mirando la pantalla están pasando cosas y tú no te enteras; esto que llaman FOMO (del inglés *fear of missing out*), el miedo de quedarse fuera. Y el tercer principio, muy importante, es la velocidad; hay una expresión dentro del mundo de la industria Google, la frecuencia de acontecimientos o *event frequency*. Es como una fórmula que dice que cuantos más juegos se es capaz de hacer en menos tiempo, más rápido se engancha la o el jugador, en este caso, la persona usuaria. (Peirano, M. (2022). *Democracia y redes sociales*. Instituto Nacional Electoral, pp. 16-18)

- 11 Diversos estudios neurocientíficos han mostrado que determinados mecanismos característicos de las redes sociales –como los *likes*, la aprobación social o la auto-revelación– activan regiones integradas en el sistema mesolímbico de recompensa, particularmente el *nucleus accumbens* y el área tegmental ventral, estrechamente vinculadas con la señalización dopaminérgica. Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., y Dapretto, M. (2016). The power of the Like in adolescence: Effects of peer influence on neural and behavioral responses to social media. *Psychological Science*, 27(7), 1027–1035. <https://doi.org/10.1177/0956797616645673>; Tamir, D. I., y Mitchell, J. P. (2012). Disclosing information about the self is intrinsically rewarding. *Proceedings of the National Academy of Sciences*, 109(21), 8038–8043. <https://doi.org/10.1073/pnas.1202129109>; Meshi, D., Morawetz, C., y Heekeren, H. R. (2013). Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use. *Frontiers in Human Neuroscience*, 7, 439. <https://doi.org/10.3389/fnhum.2013.00439>; Lindström, B., Bellander, M., Schultner, D. T., Chang, A., Tobler, P. N., y Amodio, D. M. (2021). A computational reward learning account of social media engagement. *Nature Communications*, 12, 1311. <https://doi.org/10.1038/s41467-020-19607-x>
- 12 «No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.» 230(c)(1) de la Communications Decency Act, codificada en 47 U.S.C. § 230(c)(1).
- 13 «El prestador de servicios no podrá ser considerado responsable de la información almacenada a petición del destinatario [...] siempre que no tenga conocimiento efectivo de una actividad o contenido ilícito o, cuando adquiera dicho conocimiento, actúe con prontitud para retirar el contenido ilícito o bloquear el acceso a este». 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), DOUE L 277, de 27 de octubre de 2022
- 14 Muchas veces tras un curso de sólo una semana, como me relató una amiga que se trasladó a vivir Dublín y encontró trabajo como moderadora en Meta.
- 15 Gillespie, T. (2018). Platforms are not intermediaries. *Georgetown Law Technology Review*, 2(2), 198–216.
- 16 Chandler, J. A. (2007). A right to reach an audience: An approach to intermediary bias on the Internet. *Hofstra Law Review*, 35(3), 1095–1137; Goldman, E. (2006). Search engine bias and the demise of search engine utopianism. *Yale Journal of Law y Technology*, 8, 188–200.

- 17 Sander, B. (2020). Freedom of expression in the age of online platforms: The promise and pitfalls of a human-rights based approach to content moderation. *Fordham International Law Journal*, 43(4), 939–1006.
- 18 De Gregorio, G. (2020). Democratising online content moderation: A constitutional framework. *Computer Law y Security Review*, 36, 105374. <https://doi.org/10.1016/j.clsr.2019.105374>.
- 19 Zuboff, S. (2020). *La era del capitalismo de la vigilancia: La lucha por un futuro humano frente a las nuevas fronteras del poder* (A. Santos Mosquera, Trad.). Paidós.
- 20 Balkin, J. M. (2018). Free speech in the algorithmic society: Big data, private governance, and new school speech regulation. *UC Davis Law Review*, 51(3), 1149–1210
- 21 Sunstein, C. R. (2009). *On rumors: How falsehoods spread, why we believe them, what can be done*. Farrar, Straus and Giroux.
- 22 Pariser, E. (2017). *El filtro burbuja: Cómo la red decide lo que leemos y lo que pensamos*. Taurus.
- 23 Bakshy, E., Messing, S., y Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132. <https://doi.org/10.1126/science.aaa1160>.
- 24 Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., y Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118. <https://doi.org/10.1073/pnas.2023301118>.
- 25 Bail, C. A., Argyle, L. P., Brown, T. W., et al. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37), 9216–9221. <https://doi.org/10.1073/pnas.1804840115>
- 26 Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., y Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3), 554–559. <https://doi.org/10.1073/pnas.1517441113>
- 27 Weismueller, J., Harrigan, P., Coussement, K., y Tessitore, T. (2022). What makes people share political content on social media? The role of emotion, authority and ideology. *Computers in Human Behavior*, 129, 107150. <https://doi.org/10.1016/j.chb.2021.107150>

- 28 Rathje, S., Van Bavel, J. J., y Van der Linden, S. (2021). Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences*, 118(26), e2024292118. <https://doi.org/10.1073/pnas.2024292118>
- 29 De León, E., y Trilling, D. (2021). A sadness bias in political news sharing? The role of discrete emotions in the engagement and dissemination of political news on Facebook. *Social Media + Society*, 7(4). <https://doi.org/10.1177/20563051211059710>
- 30 Arias Maldonado, M. (2016). *La democracia sentimental: Política y emociones en el siglo XXI*. Página Indómita.
- 31 <https://reutersinstitute.politics.ox.ac.uk/es/digital-news-report/2025>
- 32 Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- 33 *Ibidem*, 202.
- 34 Romero, A., y Nelson, J. (2026). Cómo sobrevivir a la guerra cognitiva. *Política Exterior*, (232), p. 123.
- 35 La *Alliance for Security Democracy* ha empezado a denominar a estos actores como Manipulador Persistente Avanzado (*Advanced Persistent Manipulator -APM-*), operadores que llevan a cabo operaciones coordinadas y sostenidas a lo largo de múltiples sucesos, con la misma infraestructura *proxy*, blanqueo y amplificación reactivada en torno al suceso que el adversario intenta moldear. Los APT atacan sistemas, atacan las condiciones bajo las cuales los sistemas, y los humanos que los gestionan, permitirían comprenderlos, atribuirlos y responder ante ellos. (Romero y Nelson, *Op. Cit.*, p. 126.)
- 36 Alyukov, M., Makhortykh, M., Voronovici, A., y Sydorova, M. (2025). LLMs grooming or data voids? LLM-powered chatbot references to Kremlin disinformation reflect information gaps, not manipulation. *Harvard Kennedy School Misinformation Review*, 6(5). <https://doi.org/10.37016/mr-2020-1>
- 37 Romero y Nelson, *Op. Cit.*, p. 127.
- 38 La denominada Ley de Moore describe la evolución exponencial de la capacidad computacional, derivada del aumento periódico del número de transistores integrables en los microprocesadores, fenómeno que durante décadas ha impulsado el rápido desarrollo de las tecnologías digitales (Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, 38(8), 114–117.) Moore acertó mucho más de lo que pretendía, pues una predicción formulada en 1965 para apenas una década terminó describiendo durante más de medio siglo el crecimiento exponencial de la capacidad tecnológica.

- 39 Collingridge, D. (1980). *The social control of technology*. Frances Pinter, especialmente. pp. 13-21.
- 40 Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press, pág. 9.
- 41 Bratton, B. H. (2016). *The stack: On software and sovereignty*. The MIT Press. <https://doi.org/10.7551/mitpress/9780262029575.001.0001>
- 42 Van Dijck, J. (2021). Seeing the forest for the trees: Visualizing platformization and its governance. *New Media & Society*, 23(9), 2801–2819. <https://doi.org/10.1177/1461444820940293>
- 43 Ibidem
- 44 Lynskey, O. (2017). *Regulating 'platform power'* (LSE Legal Studies Working Paper No. 1/2017). London School of Economics and Political Science.
- 45 Van Dijck, J. (2021). *Op.cit.*, p. 2809.
- 46 Owen, T. (2019). Introduction: Why platform governance? En T. Owen (Ed.), *Models for platform governance*. Centre for International Governance Innovation.
- 47 Poell, T., Nieborg, D., & Van Dijck, J. (2019). Platformisation. *Internet Policy Review*, 8(4). <https://doi.org/10.14763/2019.4.1425>, p. 3.
- 48 En el contexto asiático, WeChat reúne 1.410 millones de usuarios, mientras que la red social TikTok de la empresa china ByteDance se sitúa en torno a los 1.990 millones.
- 49 Como indica Harari, «supongamos que una empresa local emergente desarrolla su propio motor de búsqueda e intenta competir con Google. No tendrá nada que hacer. Puesto que Google lo utilizan ya millones de personas, dispone de muchísimos más datos con los que adiestrar y mejorar sus algoritmos, lo que servirá para atraer más tráfico, que se usará para adiestrar a la nueva generación de algoritmos, y así sucesivamente». Harari, Y. N. (2024). *Nexus: Una breve historia de las redes de información desde la Edad de Piedra hasta la IA* (J. Ros, Trad.). Debate, p. 407.
- 50 Como nos indica Garton Ash, «la lucha es más complicada en la Red. Una plétora de organizaciones internacionales, gobiernos nacionales, parlamentos, empresas, ingenieros, grupos de medios, tuiteros célebres y campañas masivas –tanto físicas como virtuales– a través de las redes sociales compite ahora en un juego de múltiples niveles y dimensiones. A menudo el resultado depende de intrincadas intersecciones entre negocios, política, derecho, regulaciones y tecnologías de la comunicación que se

- desarrollan velozmente». Garton Ash, T. (2017). *Libertad de palabra: Diez principios para un mundo conectado*. Tusquets, p.49.
- 51 Schmidt, E., y Cohen, J. (2014). *El futuro digital*. Anaya Multimedia.
- 52 Tomo una parte del contenido de este epígrafe de mi artículo «Los imperios de la palabra. El poder global de las plataformas digitales», *Política Exterior Vol. XL. n° 232 (Ejemplar dedicado a Tecnología y Poder)*, pp.132-143.
- 53 Robles Carrillo, M. (2023). La articulación de la soberanía digital en el marco de la Unión Europea. *Revista de Derecho Comunitario Europeo*, (75), 133–171. <https://doi.org/10.18042/cepc/rdce.75.05>
- 54 Pérez Martínez, J., y Rodríguez Pita, P. (2023). La fragmentación de Internet: Ciberguerra geopolítica entre países orientales y occidentales. *Telos: Cuadernos de Comunicación e Innovación*, (122), 102–109.
- 55 <https://unctad.org/es/publication/informe-sobre-las-inversiones-en-el-mundo-2025>
- 56 Sassen, S. (2006). *Territory, authority, rights: From medieval to global assemblages*. Princeton University Press.
- 57 En lo Garton Ash (*Op. Cit.*, p. 314) denomina los *lobbies* de la identidad y su subasta del victimismo.
- 58 Comisión Interamericana de Derechos Humanos, Relatoría Especial para la Libertad de Expresión, *Libertad de expresión e Internet*, OEA/Ser.L/V/II, CIDH/RELE/INE 11/13, 31 de diciembre de 2013, pp. 47-48.
- 59 Taylor, E. (2016). *The privatization of human rights: Illusions of consent, automation and neutrality* (Global Commission on Internet Governance Paper Series, Paper No. 24). Centre for International Governance Innovation.
- 60 Morozov, E. (2011). *The net delusion: The dark side of Internet freedom*. PublicAffairs, p. 223.
- 61 McDonald, A. M., y Cranor, L. F. (2008). The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society*, 4(3), 543–568, p. 563.
- 62 Brown, I., y Marsden, C. T. (2013). *Regulating code: Good governance and better regulation in the information age*. MIT Press, p. 54.
- 63 Taylor, *Op. Cit.* pp. 15-16.

64 El artículo 14 regula las Condiciones Generales, convertidas en el instrumento de vinculación legal del usuario a las plataformas, estableciendo que los prestadores de servicios intermediarios incluirán en sus condiciones generales información sobre cualesquiera políticas, procedimientos, medidas y herramientas empleadas para moderar los contenidos, incluidas la toma de decisiones mediante algoritmos y la revisión humana, así como sobre las normas de procedimiento de su sistema interno de gestión de reclamaciones. Se expondrá en lenguaje claro, sencillo, inteligible, accesible al usuario e inequívoco, y se hará pública en un formato fácilmente accesible y legible por máquina. El control del uso de dichos algoritmos se realizará por la propia Comisión Europea (arts.69.5 y 72 DSA). Sin mencionar de forma directa el uso de sistemas de IA, la DSA aborda los peligros de las decisiones automatizadas (algunas pueden incluir sistemas de IA y otras no). Entre las obligaciones de transparencia informativa de los prestadores de servicios intermediarios encontramos la obligación de informar sobre cualquier actividad de moderación de contenidos en que se haya hecho «uso de medios automatizados con fines de moderación de contenidos, incluyendo una descripción cualitativa, una especificación de los fines precisos, los indicadores de la precisión y la posible tasa de error de los medios automatizados empleados para cumplir dichos fines, y las salvaguardias aplicadas» (art. 15.e). También los prestadores de servicios de alojamiento están obligados a declarar los motivos para catalogar la información suministrada por un usuario como contenido ilegal o incompatible con sus condiciones generales, y «en su caso, información sobre el uso de medios automatizados para adoptar la decisión, incluida información sobre si la decisión se ha adoptado respecto de contenidos detectados o identificados utilizando medios automatizados» (art. 17.c).

65 Lukes, S. (1974). *Power: A radical view*. Macmillan.

66 Nye, J. S. (2011). *The future of power*. PublicAffairs

67 Han, B.-C. (2016). *Sobre el poder*. Herder.

68 Han, B.-C. (2014). *Psicopolítica: Neoliberalismo y nuevas técnicas de poder*. Herder.

69 Zuboff, *Op.cit.*, p. 21, p. 181.

70 Hasselbalch, G. (2025). *Human power: Seven traits for the politics of the AI machine age*. CRC Press. <https://doi.org/10.1201/9781003527855>. Las siete cualidades específicamente humanas que Hasselbalch contrapone a las capacidades de las máquinas son: creatividad, intuición, emoción, vida, rebeldía (*defiance*), amor y sabiduría.

Parte tercera

- 1 Puede consultarse este documento en <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>
 - 2 Sun, T. Q., y Medaglia, R. (2019). Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), 368–383. <https://doi.org/10.1016/j.giq.2018.09.008>
 - 3 Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., y Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4).
 - 4 Siguiendo a Barrio, podemos iniciar la sistematización de las diferentes IA con los denominados *Machine Learning* (ML) o sistemas de IA de aprendizaje automático, que han posibilitado la automatización de procesos cognitivos a partir de un entrenamiento inicial. Estos modelos se basan en un enfoque de aprendizaje supervisado en que se introducen inicialmente conjuntos de datos previamente resueltos –es decir, datos de entrada acompañados de sus respectivas salidas esperadas–. A través de este proceso, el algoritmo es entrenado para operar de manera autónoma, permitiéndole generar resultados sobre la base de nuevos conjuntos de datos. En lugar de proporcionar explicaciones causales, estos sistemas realizan inferencias de tipo probabilístico fundamentadas en los datos de entrenamiento. Dentro del ámbito del aprendizaje automático (*machine learning*), destaca el aprendizaje profundo (*deep learning*, DL), una subdisciplina que se fundamenta en el uso de redes neuronales artificiales de múltiples capas. Este enfoque se basa en el aprendizaje jerárquico de representaciones o características a partir de ejemplos, permitiendo que el modelo adquiera conocimiento a través de distintos niveles de abstracción. En este contexto, se inscribe también la inteligencia artificial generativa, un tipo de sistema capaz de producir texto, imágenes u otros formatos de contenido en respuesta a instrucciones o prompts. Ejemplos representativos de esta tecnología son ChatGPT y su integración en herramientas como Microsoft Copilot.
- La siguiente tipología de sistemas de IA son aquellos que procesan lenguaje natural (*natural language processing*). El procesamiento del lenguaje natural (PLN) constituye un campo fundamental dentro de las tecnologías de inteligencia artificial, al posibilitar la interacción entre humanos y máquinas mediante el uso del lenguaje natural. A diferencia de los lenguajes de programación tradicionales, que se caracterizan por su precisión, univocidad y estructura rigurosa, el PLN permite a los sistemas computacionales analizar, interpretar y comprender el lenguaje humano en toda su complejidad y ambigüedad. Los algoritmos empleados en el PLN se sustentan, en su mayoría, en técnicas de aprendizaje automático. En lugar de depender de la codificación manual de

extensas reglas, como ocurría en los sistemas expertos clásicos, estos sistemas infieren patrones y regularidades a partir del análisis estadístico de grandes volúmenes de datos lingüísticos (*corpus*), generando modelos capaces de extrapolar comportamientos lingüísticos sobre nuevos textos no vistos durante el entrenamiento.

Entre las principales tareas asociadas al procesamiento del lenguaje natural se encuentran el análisis sintáctico y el etiquetado gramatical (*part-of-speech tagging*), el análisis semántico, el reconocimiento óptico de caracteres (OCR), la extracción de relaciones semánticas, la generación de resúmenes automáticos, el análisis discursivo y el reconocimiento del habla. Estas capacidades permiten una comprensión más profunda del contenido textual y oral, facilitando así aplicaciones avanzadas en diversos ámbitos, desde los asistentes virtuales hasta los sistemas de gestión documental y análisis de opinión. Otros sistemas de IA son el Diseño de sistemas expertos, es decir, sistemas que emulan la capacidad de tomar decisiones de un profesional humano; la Visión artificial (*computer vision*), que incluye la visión automática y el reconocimiento de imágenes; Reconocimiento del habla, que permite la comunicación hablada entre máquinas y personas, haciendo posible la traducción de lenguaje hablado a texto y viceversa; o la Planificación automática, que tiene por objeto la elaboración de programas de planificación, generalmente para su ejecución por robots. (Barrio Andrés, M. (2024). *Manual de Derecho digital* (3.ª ed.). Tirant lo Blanch, p. 79ss)

- 5 Samoil, S., et al. (2020). *AI WATCH: Defining artificial intelligence* (JRC118163). Publications Office of the European Union. <https://publications.jrc.ec.europa.eu/repository/handle/JRC118163>
- 6 Muñoz Vela J.M. (2024). *La regulación de la inteligencia artificial. Reto y oportunidad desde una perspectiva global e internacional*. Tirant lo Blanch.
- 7 Gallego Sánchez, E. (2019, junio 1). La patentabilidad de la inteligencia artificial: La compatibilidad con otros sistemas de protección. *La Ley Mercantil*, (59), p. 4.
- 8 OCDE define los sistemas de IA como todo «sistema basado en máquinas que, para objetivos explícitos o implícitos, infiere, a partir de la información que recibe, cómo generar resultados como predicciones, contenido, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales. Los diferentes sistemas de IA varían en sus niveles de autonomía y adaptabilidad después de su implementación».(Consejo de la Organización para la Cooperación y el Desarrollo Económicos. (2024). *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449), p. 7). La UNESCO caracteriza los sistemas de IA como «tecnologías de procesamiento de la información que integran modelos y algoritmos que producen una capacidad para aprender y realizar tareas cognitivas, dando lugar a resultados como la predicción y la adopción de

decisiones en entornos materiales y virtuales. Los sistemas de IA están diseñados para funcionar con diferentes grados de autonomía, mediante la modelización y representación del conocimiento y la explotación de datos y el cálculo de correlaciones. Pueden incluir varios métodos, como, por ejemplo, aunque no exclusivamente: i) el aprendizaje automático, incluido el aprendizaje profundo y el aprendizaje de refuerzo; ii) el razonamiento automático, incluidas la planificación, la programación, la representación del conocimiento y el razonamiento, la búsqueda y la optimización. Los sistemas de IA pueden utilizarse en los sistemas ciberfísicos, incluidos la Internet de las cosas, los sistemas robóticos, la robótica social y las interfaces entre seres humanos y ordenadores, que comportan el control, la percepción, el procesamiento de los datos recogidos por sensores y el funcionamiento de los actuadores en el entorno en que operan los sistemas de IA». (UNESCO (2021). Recomendación sobre la Ética de la Inteligencia Artificial, 2021. p. 66. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa)

- 9 Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) nº 300/2008, (UE) nº 167/2013, (UE) nº 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial).
- 10 Es crítico con esta Plaza Penadés, J. (2025). Descifrando el concepto jurídico de inteligencia artificial contenido en el Reglamento IA. *Revista Aranzadi de Derecho y Nuevas Tecnologías*, (67).
- 11 Sobre la definición es necesario referirnos a la Comunicación sobre la «Aprobación del contenido del proyecto de Comunicación de la Comisión - Directrices de la Comisión sobre la definición de sistema de inteligencia artificial establecido por el Reglamento (UE) 2024/1689 (Ley AI)» <https://ec.europa.eu/newsroom/dae/redirection/document/112455>. Tapia establece como la Comunicación descompone la noción legal en siete factores principales que son: «(1) un sistema basado en máquinas; (2) que está diseñado para operar con distintos niveles de autonomía; (3) que puede mostrar capacidad de adaptación tras su despliegue; (4) y que, para objetivos explícitos o implícitos; (5) infiere, a partir de la entrada que recibe, cómo generar salidas (6) como predicciones, contenidos, recomendaciones o decisiones (7) que pueden influir en entornos físicos o virtuales» (aptdo.9). (Tapia Hermida, A. J. (2023). El sistema europeo digital: la declaración europea de 23 de enero de 2023 sobre los derechos y principios digitales. *La Ley Unión Europea*, (112), p. 3).
- 12 Nos remitimos a Muñoz Vela, quien lo ha hecho de forma detallada en la obra anteriormente citada.

- 13 Sobre esta cuestión, valga la reflexión de Rubí Puig, A. (2024). Una lectura del reglamento de inteligencia artificial desde el derecho privado. *Indret: Revista para el Análisis del Derecho*, (4).
- 14 <https://www.boe.es/doue/2024/2853/L00001-00022.pdf>. También estamos a la espera de la promulgación de la paralizada Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA) [COM(2022) 496 final].
- 15 Valga como ejemplo, el software Predpol empleado en algunos Estados de Estados Unidos, para recopilando y analizando datos sobre crímenes pasados, predecir los futuros, y que ha suscitado dudas por su falta de acierto y por el empleo de sesgos raciales en los datos de entrada sobre los que basar las predicciones. Vid al respecto <https://hrdag.org/pressroom/predpol-amplifies-racially-biased-policing/>
- 16 Vid. al respecto Bellomo Stefano; Mezzacapo, Domenico; Ferraro, Fabrizio; y Calderara, Dario. (2023). *Improving working conditions in platform work in the light of the recent proposal for a directive*, Convegno (Diritto, Política, Economía).
- 17 «Las administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente, y que priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos».
- 18 Valgan como ejemplo, entre otros, Cerrillo i Martínez, A. (2019). El impacto de la inteligencia artificial en el Derecho administrativo. ¿Nuevos conceptos para nuevas realidades técnicas? *Revista General de Derecho Administrativo*, (50); Gutiérrez David, M. E. (2022). Acceso al código fuente y a los algoritmos de las Administraciones inteligentes: Lecciones a partir de experiencias comparadas. En L. Cotino Hueso y J. Castellanos Claramunt (Eds.), *Transparencia y explicabilidad de la inteligencia artificial* (pp. 135–168). Tirant lo Blanch, que disecciona las diferentes posiciones de la doctrina; o Gamero Casado, E. (Dir.), y Pérez Guerrero, F. L. (Coord.). (2023). *Inteligencia artificial y sector público: Retos, límites y medios*. Tirant lo Blanch, con interesantes aportaciones de diferentes académicos.
- 19 Nos remitimos a obras que lo han abordado con brillantez y profundidad, como Cerrillo i Martínez, A., Di Lascio, F., Martín Delgado, I., y Velasco Rico, C. (Dirs.). (2024). *Inteligencia artificial y administraciones públicas: Una triple visión en clave comparada*. Iustel; Cotino Hueso, L. (2023). Los usos de la inteligencia artificial en el sector público,

- su variable impacto y categorización jurídica, *Revista Canaria De Administración Pública*, (1), pp. 211–242; Guichot Reina, E. (2024). La transparencia de los algoritmos desde la normativa sobre transparencia y acceso a la información pública. *Revista Andaluza de Administración Pública*, (117), 77–115; Boix Palop, A. (2022). Transparencia en la utilización de inteligencia artificial por parte de la Administración. *El Cronista del Estado Social y Democrático de Derecho*, (100), 90–105; Gutiérrez David, M. E. (2021). Administraciones inteligentes y acceso al código fuente y los algoritmos públicos: Conjurando riesgos de cajas negras decisionales. *Derecom*, (31), 19–105; Vestri, G. (2022). *La disrupción tecnológica en la Administración Pública: Retos y desafíos de la inteligencia artificial*. Thomson Reuters Aranzadi.
- 20 Abad Alcalá, L. (2023). IA, Administraciones Públicas y participación ciudadana. En W. Arellano Toledo (Ed.), *Derecho, ética e inteligencia artificial* (pp. 295–334). Tirant lo Blanch.
 - 21 <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>
 - 22 Presno Linera, M. Á., y Meuwese, A. C. M. (2024). La regulación de la inteligencia artificial en Europa. *Teoría y Realidad Constitucional*, (54), p. 139.
 - 23 Harari, *Op.cit.*, p. 23; p. 265.
 - 24 Véliz, C. (2026). Predicción y vigilancia: los pecados originales de la tecnología digital. *Política Exterior* (232), p. 51.
 - 25 Fernández de Lis, P. (2026, 23 de agosto). El desafío coordinado de las IA a sus creadores. *El País*, p. 34.
 - 26 Finn, E. (2018). *La búsqueda del algoritmo: Imaginación en la era de la informática*. Alpha Decay, p. 41.
 - 27 Garton Ash nos explica que «algo parecido al lenguaje humano surgió probablemente hace no más de cien mil millones de años como resultado de los desarrollos evolutivos del cerebro, el torax y el aparato de fonación. Hablar, en este sentido muy elemental, consiste en modular el chorro de aire que emiten los pulmones mediante movimientos del torax, la cavidad bucal, la lengua y los labios, produciendo secuencias de diferentes sonidos con significados reconocibles. Cuando decimos de una niña pequeña *acaba de empezar a hablar*, es eso lo que ha aprendido a hacer. Una competencia comunicativa altamente desarrollada supone el empleo del lenguaje y el pensamiento abstracto, y es lo que distingue a los seres humanos de nuestros parientes más cercanos como

el chimpancé o el bonono. [...] El lenguaje no es uno más de los atributos humanos. Es el rasgo definitorio del hombre. (Garton Ash, *Op.cit.*, p.26)

- 28 Las reflexiones sobre si una IA puede crear arte se han intensificado en los últimos tiempos. Una imagen generada algorítmicamente puede poseer cualidades estéticas, ser expuesta en un museo, alcanzar valor económico y provocar una experiencia estética en quien la contempla. Pero, debemos cuestionarnos intencionalidad, conciencia, experiencia, libertad creadora, interpretación y atribución de significado. La literatura científica más actual parece abrir una tercera opción planteando conceptos como co-creación y autoría distribuida que permiten abordar la autoría de ciertas obras como resultado de una interacción entre intención humana, *prompts*, selección, entrenamiento algorítmico y generación computacional. Winter, D. (2023). Aesthetic aspects of digital humanism: An aesthetic-philosophical analysis of whether AI can create art. En H. Werthner et al. (Eds.), *Introduction to digital humanism*. Springer; Moore, J. G. (2024). Creativity, credit, and copyright in the age of artificial art. *The Journal of Aesthetics and Art Criticism*, 82(3), 265–277. <https://doi.org/10.1093/jaac/kpae034>; Cascales, R. (2023). Interpreting AI-generated art: Arthur Danto's perspective on intention, authorship, and creative traditions in the age of artificial intelligence. *The Polish Journal of Aesthetics*, 71(4), 17–29. <https://doi.org/10.19205/71.23.2>
- 29 El 5 de mayo pasado se estrenó en la Opera Real del Palacio de Versalles la obra «Molière ex *Machina*», una obra de teatro auspiciada por la Sorbona en colaboración con Théâtre Molière Sorbonne, Obvious Collective y Mistral IA, escrita con ayuda de la IA, que juega con el concepto de que la habría podido firmar Molière si hubiese vivido un año más. Lo curioso, es que un gran número de espectadores afirmaron que podría haber sido escrita enteramente por humanos.
- 30 Harari, *Op.cit.*, p. 254-255.
- 31 Criddle, C. (2026). Artificial intimacy: A teenager's last conversation. *Financial Times*, *Tech Tonic*; *The Guardian*. (2025, 30 de octubre). Teenage boys using 'personalised' AI for therapy and romance, survey finds. *The Guardian*; Pascual, M. G. (2026, 28 de julio). *Amigos digitales con IA, la nueva amenaza para la salud mental de los adolescentes. El País*. Un estudio de Common Sense Media (2025), realizado sobre una muestra representativa de 1.060 adolescentes estadounidenses de entre 13 y 17 años, revela que el 72 % ha utilizado compañeros de IA y que un tercio los emplea para algún tipo de interacción social o relacional, incluyendo apoyo emocional, amistad e incluso relaciones románticas. (Common Sense Media. (2025). *Talk, trust, and trade-offs: How and why teens use AI companions*. Common Sense Media.)

- 32 El proceso seguido en Estados Unidos contra Meta por los posibles efectos perjudiciales de Facebook e Instagram sobre los menores concluyó en agosto de 2026 mediante un acuerdo que evitó una sentencia definitiva sobre el fondo. La compañía aceptó pagar hasta aproximadamente 16.700 millones de dólares para resolver las reclamaciones de numerosos Estados que la acusaban de haber diseñado deliberadamente sus plataformas para generar comportamientos adictivos entre niños y adolescentes, de haber minimizado públicamente sus riesgos y de haber recopilado indebidamente datos de menores. El acuerdo contempla, además, modificaciones relevantes en el diseño de Facebook e Instagram, entre ellas límites de uso para adolescentes, restricciones nocturnas y mayores medidas de protección y control parental. Meta, sin embargo, no reconoció haber cometido ninguna infracción, por lo que el acuerdo impidió que el tribunal llegara a pronunciarse sobre una cuestión de enorme trascendencia jurídica, como era hasta qué punto una plataforma puede ser considerada responsable no ya por los contenidos publicados por terceros, sino por el propio diseño de sus sistemas de recomendación y de captación de la atención. (Reuters. (2026, 26 de agosto). *Meta reaches \$16.68 billion settlement over social media harms to children*. Reuters.)
- 33 En este contexto, les recomiendo la serie DEVS , creada por Alex Garland, que confronta determinismo y libre albedrío. La serie plantea que, si el universo está regido completamente por relaciones de causa y efecto, cada decisión humana sería consecuencia inevitable de estados anteriores y, por tanto, nuestra sensación de elegir libremente sería una ilusión. La computadora cuántica de Devs lleva esta hipótesis al extremo de que si conociéramos con absoluta precisión el estado del universo, podríamos reconstruir el pasado y predecir el futuro. Sin entrar en *spoilers*, la serie introduce, sin embargo, una cuestión tan decisiva como alarmante: ¿qué sucede si un ser humano conoce de antemano lo que supuestamente va a hacer? ¿Puede decidir hacer otra cosa?, convirtiendo un problema físico en uno filosófico y moral: si nuestros actos están determinados, ¿hasta qué punto somos responsables de ellos? La serie plantea como aceptar dicho determinismo modificaría profundamente nuestra concepción de la culpa, el castigo y la responsabilidad moral.
- 34 En 1991, Mark Weiser ya preconizó lo que el denominó como *computación ubicua* explicando que las tecnologías que calan más hondo son las que se pierden de vista; su imbricación es la vida diaria es tan íntima que terminan por pasar inadvertidas. Las máquinas que consigan ajustarse al entorno humano en lugar de obligar a las personas a entrar en el mundo de la computación harán que su uso resulte tan agradable como un paseo por el parque. (Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3), 94–104.. Hay una traducción al castellano publicada por el nº 182 de la revista *Investigación y Ciencia* en noviembre de ese año.)

- 35 GPS/GNSS, que registra ubicación, desplazamientos y rutas; acelerómetro y giroscopio, que detectan movimientos, orientación y actividad física; magnetómetro o brújula; cámaras y micrófonos, capaces de captar imágenes, rostros, voz y sonidos; sensores de proximidad y luz ambiental; barómetro, presente en algunos dispositivos y capaz de detectar variaciones de presión y altitud; y sensores biométricos, como la huella dactilar o el reconocimiento facial. A ello se añaden tecnologías de conectividad que también generan información, como Bluetooth, Wi-Fi, NFC y UWB, capaces de revelar proximidad a otros dispositivos, redes utilizadas, interacciones, pagos o incluso posiciones con gran precisión.
- 36 Datos obtenidos de Moody, R. (2025, 25 de junio). *Surveillance camera statistics: Which are the most surveilled cities?* Comparitech.
- 37 Según Norris, C., y Armstrong, G. (1999). *The maximum surveillance society: The rise of CCTV*. Berg
- 38 Tenemos un ejemplo muy reciente y bastante impactante. En el Carnaval de Notting Hill de 2026, la Policía Metropolitana utilizó reconocimiento facial en directo y atribuyó a esa tecnología 121 detenciones, cerca del 19 % de todas las realizadas durante los dos días. *The Times*. (2026, September 1). *Notting Hill Carnival: 121 arrests came from face scanning*. *The Times*.
- 39 Reuters. (2023, 8 de abril). Iran installs cameras in public places to identify, penalise unveiled women—Police statement. *Reuters*.
- 40 Sobre esta cuestión, Véliz nos explica que lo que muchas veces nos ofrece la IA son predicciones, que pueden parecer descripciones del mundo, como hechos, pero no son ni lo uno ni lo otro. Son lo que el filósofo J.L. Austin denominó «actos de habla», un lenguaje que hace algo en lugar de describir. Esto pone en duda programas como Veripol en España o COMPAS en Estados Unidos, que precisamente han dejado de emplearse. Y concluye que «si castigamos o negamos oportunidades basándonos en predicciones, no hay forma de impugnarlas. Dado que se refieren al futuro, no se puede demostrar que son falsas» (Véliz, *Op.cit.*, p. 54-55)
- 41 Abad Alcalá, L. (2026). Transparencia algorítmica: Enfoques y desafíos para el Estado democrático. *Revista Española de la Transparencia*, (23, Extra), 91–138. <https://doi.org/10.51915/ret.443>
- 42 STS 1119/2025, de 11 de septiembre, Sala de lo Contencioso-Administrativo, Sección 3.ª, rec. 7878/2024, ECLI:ES:TS:2025:3826. Puede consultarse un análisis que hemos realizado de dicha sentencia en las Actas del XXIII Congreso de la Asociación de Constitucionalistas de España (de próxima publicación).

- 43 Aragón Reyes, M. (2020). Estado democrático, En Pendás García, B. (ed. lit.), *Enciclopedia de las Ciencias Morales y Políticas para el siglo XXI. Ciencias Políticas y Jurídicas (con especial referencia a la sociedad poscovid 19)*. Real Academia de Ciencias Morales y Políticas-Boletín Oficial del Estado, p. 778.
- 44 Garrorena Morales, A. (2011). «Estado Democrático», En Aragón Reyes, M. (Dtor.) y Aguado Renedo, C. (Codtor.), *Constitución, Estado Constitucional, Partidos y Elecciones y Fuentes del Derecho. Temas básicos de Derecho Constitucional*. Tomo I. Civitas-Thomson Reuters, pp. 127.
- 45 Habermas, J. 2018. Verdad y justificación. Madrid: Trotta. Seguimos en esta parte a Acosta Yparraguirre, E. (2023). La inteligencia artificial y el fin del Estado Constitucional y la Democracia... ¿o no?». *Eunomía. Revista en Cultura de la Legalidad* 27, pp. 137-154.
- 46 Sartori, G. (2003). *¿Qué es la democracia?*. Taurus, p. 93.
- 47 Arendt, H. (1996). *La condición humana*. Paidós, p.66.
- 48 Dworkin, G. (2015). «The nature of autonomy». *Nordic Journal of Studies in Educational Policy*, 2015(2): 7-14. DOI: <https://doi.org/10.3402/nstep.v1.28479>
- 49 Häberle, P. (2017). *Tiempo y Constitución*. Palestra, p. 69.
- 50 Dahl, R. A. (2022). *La democracia*. Ariel, pp. 100-101.
- 51 Considera que «esta acumulación de poder en manos de unos pocos debe ser observada de modo abarcativo: incluye el poder del dinero, el poder sobre la infraestructura democrática y el discurso público, el poder sobre los individuos a través de la elaboración de perfiles precisos y la supremacía en el campo de la inteligencia artificial. Los gigantes de internet han tenido logros sin precedentes como empresas: buena parte de su producción no está regulada, son dominantes en su mercado y son las empresas más rentables en la bolsa de valores. Tienen una inmensa influencia en la formación de opinión y en la política, al tiempo que han conservado su popularidad entre el público» (Nemitz, Paul (2021). La democracia en la era de la inteligencia artificial. *Nueva Sociedad* 294 / julio – agosto. 130-140, p. 131).
- 52 Innerarity, D. (2025). *Una teoría crítica de la inteligencia artificial*. Galaxia Gutenberg.
- 53 Han, B.-C. (2022). *Infocracia: La digitalización y la crisis de la democracia* (J. Chamorro Mielke, Trad.). Taurus, p. 66.
- 54 En las elecciones presidenciales estadounidenses de 2020, distintos estudios detectaron decenas de miles de cuentas automatizadas participando activamente en la

conversación electoral. En una muestra longitudinal de casi un millón de cuentas, cerca del 20 % fueron clasificadas como bots y llegaron a generar más del 43 % de los casi 100 millones de mensajes analizados.» (Ferrara, E., Chang, H., Chen, E., Muric, G., y Patel, J. (2020). Characterizing social media manipulation in the 2020 U.S. presidential election. *First Monday*, 25(11). <https://doi.org/10.5210/fm.v25i11.11431>.)

- 55 Es el fenómeno denominado «shadow banning». Sanciones de moderación de contenido imperceptibles para los afectados. Esta técnica conecta las preocupaciones actuales sobre nuevas técnicas de gestión de la visibilidad en la moderación de contenido, como la exclusión de listas y la degradación. Las técnicas de moderación convencionales, como la eliminación directa de contenido o la suspensión de cuentas, pueden ser observadas por los afectados, pero esta nueva actividad a menudo no. Heldt, A. (2023). An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation. *Computer Law y Security Review*, 48, 105790. <https://doi.org/10.1016/j.clsr.2023.105790>
- 56 Una versión más desarrollada de lo expuesto en este epígrafe puede encontrarse en Abad Alcalá, L. (2026). IA y Estado democrático de Derecho. En M.^a R. Fernández Riveira y Á. Sánchez Navarro (Coords.), *Democracia y tecnología (Participar mediante tecnología)* (pp. 239–268). Centro de Estudios Políticos y Constitucionales. [En imprenta].
- 57 Cotino Hueso, L. (2023). Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida. *Revista Española de la Transparencia*, 16, p. 28.
- 58 Ciertas obligaciones del REIA han sido modificadas por el *Digital Omnibus on AI*, aprobado como Reglamento (UE) 2026/1744, de 8 de julio de 2026, publicado el 24 de julio y en vigor desde el 27 de julio de 2026. Probablemente el cambio más relevante en la práctica está en el nuevo art. 111.4. El Omnibus introduce un período transitorio de cuatro meses para los proveedores de sistemas de IA –incluidos sistemas de IA de propósito general– que generen audio, imágenes, vídeo o texto sintéticos y que hubieran sido introducidos en el mercado antes del 2 de agosto de 2026. Estos sistemas tienen hasta el 2 de diciembre de 2026 para cumplir específicamente con el art. 50.2, es decir, con la obligación de marcado legible por máquina y detectabilidad del contenido generado o manipulado artificialmente.
- 59 https://commission.europa.eu/documents_en?prefLang=es
- 60 <https://digital-strategy.ec.europa.eu/en/library/code-conduct-disinformation>
- 61 Parlamento Europeo y Consejo de la Unión Europea. (2022). Reglamento (UE) 2022/2065, 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). *Diario Oficial de la Unión Europea*, L 277, 1-102.
- 62 Pauner Chulvi, C. (2023). Transparencia algorítmica en los medios de comunicación y las plataformas digitales. *Revista Española de la Transparencia*, 17, pp. 120-121.
 - 63 Kaiser, B. (2019). *La dictadura de los datos*. Harper Collins Ibérica.
 - 64 Palacio, A. (2026, 11 de julio). Soberanía artificial. *El Mundo*.
 - 65 León XIV. (2026). *Magnifica Humanitas. Carta encíclica sobre la custodia de la persona humana en el tiempo de la inteligencia artificial*. Santa Sede. (https://www.vatican.va/content/leo-xiv/es/encyclicals/documents/20260515-magnifica-humanitas.html?utm_source=chatgpt.com)

Epílogo

- 1 Taylor, P. M. (2003). *Munitions of the mind: A history of propaganda from the ancient world to the present era* (3rd ed.). Manchester University Press.
- 2 La cita completa es «Uno de los fenómenos más importantes de la vida jurídica y espiritual de la humanidad es que quien posee el verdadero poder puede determinar por sí mismo los conceptos y las palabras. *Caesar dominus et supra grammaticam*: el César es también señor de la gramática. Schmitt, C. (2001). El imperialismo moderno en el derecho internacional público. En H. Orestes Aguilar (Comp.), *Carl Schmitt, teólogo de la política* (p. 112). Fondo de Cultura Económica, p. 112.
- 3 Klemperer, V. (2001). *LTI. La lengua del Tercer Reich. Apuntes de un filólogo* (A. Kovacsics, Trad.). Minúscula.
- 4 Según el Tribunal Constitucional el uso de la vía penal al juzgar la vulneración del derecho, puede «implicar un sacrificio desproporcionado e innecesario de los derechos fundamentales en juego, que podría tener un efecto disuasorio o delentador de su ejercicio» (STC 190/2020, FJº5º)
- 5 Por su parte, el Tribunal Europeo de Derechos Humanos, caso *Feldek vs. Slovakia*, de 12 julio de 2001, apdo. 54, define la censura como «una violación radical tanto del derecho de cada persona a expresarse como del derecho de todos a estar bien informados, de modo que se afecta una de las condiciones básicas de una sociedad democrática»
- 6 Expresión que procede de Tácito –aparece en las *Historias* (I, 4) y en los *Anales* (II, 36)–, aunque será la teoría de la razón de Estado de la Edad Moderna la que la transforme en una verdadera categoría política, al entenderlos en una definición clásica como

«*intimae et occultae rationes sive consilia eorum qui in republica principatum obtinent*», esto es, las razones y decisiones íntimas y ocultas de quienes ocupan la posición de poder en el Estado. (Clapmar, A. (1605). *De arcanis rerumpublicarum libri sex*. Wessel).

- 7 https://www.abc.es/espana/abci-gobierno-insiste-tratar-secreto-oficial-coste-real-viaje-sanchez-falcon-201901021335_noticia.html
- 8 Con las revelaciones de Edward Snowden y el caso Wikileaks como principales ejemplos.
- 9 Valgan como ejemplos la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno; la Directiva (UE) 2019/1937 del Parlamento Europeo y del Consejo, de 23 de octubre de 2019, relativa a la protección de las personas que informen sobre infracciones del Derecho de la Unión (transpuesta a la legislación nacional mediante la Ley 2/2023, la cual regula los sistemas internos de información y la lucha contra la corrupción); o la Directiva del Parlamento Europeo y Consejo de la Unión Europea (UE) 2024/1069, de 11 de abril de 2024, relativa a la protección de las personas que se implican en la participación pública frente a pretensiones manifiestamente infundadas o acciones judiciales abusivas («demandas estratégicas contra la participación pública»). *Diario Oficial de la Unión Europea*, L, 16 de abril de 2024; entre otras.
- 10 Peirano lo expresa diciendo que «Esa asimetría de conocimiento, de información por la cual hay un poder que lo sabe todo de todas y todos los ciudadanos, quiénes son, dónde están, qué les preocupa, con quién se relacionan, qué comen cuando nadie les ve, toda esa información que infiere sobre esos datos es la que los algoritmos generan a partir de material estadístico y de las cosas que saben: todo lo que hace todo el mundo en tiempo real» (Peirano, *Op.Cit.*, p. 46).
- 11 Zuboff, *Op.cit.*, p. 25.
- 12 Zuckerberg, M. (2019, 17 de octubre). *Standing for voice and free expression*. Georgetown University / Meta.: «Dar voz a todo el mundo empodera a quienes carecen de poder y contribuye, con el tiempo, a mejorar la sociedad». Dorsey, J. (2018, 5 de septiembre). Dorsey, J. (2018, 5 de septiembre). *Testimony of Jack Dorsey, Chief Executive Officer, Twitter, Inc.* U.S. Senate Select Committee on Intelligence, S. Hrg. 115-460, p. 21: «Queremos ser una plaza pública global, donde personas de todo el mundo se reúnan en un intercambio abierto y libre de ideas. Debemos ser un espacio fiable y saludable que favorezca un debate democrático libre y abierto». Musk, E. (2022, 25 de abril). Declaración reproducida en *Twitter to be acquired by Elon Musk*. Twitter, Inc. / U.S. Securities and Exchange Commission: «La libertad de expresión es el fundamento de una democracia que funcione, y Twitter es la plaza pública digital donde se debaten asuntos esenciales para el futuro de la humanidad». Google / Brin, S., y Page, L. (2004).

- 2004 *founders' letter*. «organizar la información del mundo y hacerla universalmente accesible y útil». Chavez, P. (2008, 20 de mayo). *Promoting free expression on the Internet*. Google Public Policy Blog.: «el compromiso de Google con la libertad de expresión está en el núcleo de todo lo que hacemos».
- 13 Vid al respecto Panagopoulou, F., Parpoula, C., y Karpouzis, K. (2025). Legal perspectives on AI and the right to digital literacy in education. *Frontiers in Computer Science*, 7, 1692268. <https://doi.org/10.3389/fcomp.2025.1692268>; Reviglio, U., y Agosti, C. (2020). Thinking outside the black-box: The case for «algorithmic sovereignty» in social media. *Social Media + Society*, 6(2), 1–12. <https://doi.org/10.1177/2056305120915613>; o König, P. D. (2024). Challenges in enabling user control over algorithm-based services. *AI y Society*, 39, 195–205. <https://doi.org/10.1007/s00146-022-01395-1>
 - 14 Rosanvallon, P. (2015). *El buen gobierno*. Manantial, p. 214.
 - 15 Abordo este aspecto de las sociedad actuales en Abad Alcalá, L. (2024). Democracias frágiles, tiempo de oportunistas. *Revista de Occidente*, (512), 53–65
 - 16 Rubio Núñez, R. (2025). El uso de la inteligencia artificial en las campañas electorales y sus efectos democráticos. *Revista de Derecho Político*, (122), p. 76.
 - 17 Una reflexión sobre esta norma la realizamos en el artículo «La reciente regulación comunitaria sobre medios de comunicación: Buenas intenciones y dudas razonables». *Cuadernos de Periodistas*, (48), 19–32.
 - 18 Cadwalladr, C. y Graham-Harrison, E. (17 de marzo de 2018). «Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach». *The Guardian*
 - 19 Lanier, J. (2011). *Contra el rebaño digital*. Debate, p. 31.
 - 20 Lassalle, J.M. (2019). *Ciberleviatán. El colapso de la democracia liberal frente a la revolución digital*. Arpa, p. 29.
 - 21 Bradford, A. (2020). *The Brussels effect: How the European Union rules the world*. Oxford University Press.
 - 22 Palacio, *Op. Cit.*
 - 23 Las previsiones de los propios responsables de las principales empresas de inteligencia artificial apuntan a que este escenario podría estar más próximo de lo que pensamos. Anthropic, por ejemplo, ha señalado que «los sistemas de inteligencia artificial potentes podrían aparecer ya a finales de 2026 o en 2027», entendiendo por tales sistemas aquellos capaces de igualar o superar las capacidades de un premio Nobel

en numerosas disciplinas, operar autónomamente sobre interfaces digitales y desarrollar tareas complejas durante periodos prolongados (Anthropic (2025). Anthropic's recommendations to OSTP for the U.S. AI Action Plan.). Sam Altman, por su parte, ha afirmado que «dentro de diez años, creo que casi con certeza habremos construido una superinteligencia» mientras que Google DeepMind sostiene que nos encontramos ya en el camino hacia una AGI que podría constituir «la tecnología más profunda jamás inventada (Hassabis, D. (2026). *From games to biology and beyond: 10 years of AlphaGo's impact*. Google DeepMind.).

- 24 OECD. (2022). *OECD framework for the classification of AI systems* (OECD Digital Economy Papers No. 323). OECD Publishing. <https://doi.org/10.1787/cb6d9eca-en>
- 25 Hand, L. (1952). *The spirit of liberty: Papers and addresses of Learned Hand* (I. Dilliard, Ed.). Alfred A. Knopf, P. 190.

Leopoldo Abad Alcalá

Catedrático de Derecho Constitucional y responsable del Área de Derecho Constitucional de la Universidad CEU San Pablo, institución en la que desarrolla su actividad docente desde 1996.

Licenciado en Derecho por la UNED y en Periodismo por el CEU-Universidad Complutense de Madrid (Premio Nacional de Terminación de Estudios Universitarios). Doctor en Derecho por la Universidad CEU San Pablo (sobresaliente *cum laude*) y Doctor en Ciencias de la Información por la Universidad Complutense de Madrid (Apto *cum laude* por unanimidad).

Investigador principal de dos proyectos financiados en el marco del Plan Nacional de I+D+i (CSO2012-36872 y CSO2015-66746-R), de un proyecto PROFIT de la Secretaría de Estado de Política Científica y Tecnológica (FIT-020100-2003-2002) y de un proyecto financiado en la convocatoria de Proyectos de Colaboración Público-Privada (CPP2023-010455); y como investigador en otros ocho proyectos competitivos. Esta actividad investigadora se ha concretado en más de sesenta publicaciones académicas, entre monografías, capítulos de libro y artículos científicos. Su producción científica supera las 1.000 citas en Google Scholar. Ha dirigido doce tesis doctorales.

Tiene reconocidos tres sexenios de investigación y un sexenio de transferencia del conocimiento por la Comisión Nacional Evaluadora de la Actividad Investigadora (CNEAI).

Profesor invitado en universidades europeas, estadounidenses e iberoamericanas, entre ellas Columbia University, Duke University, Université Paris II Panthéon-Assas, Université Montesquieu-Bordeaux IV, Universidad de Guadalajara, Universidad Rafael Landívar, Universidad del Desarrollo y Universidad Central de Chile, entre otras.

Ha desempeñado asimismo diversas responsabilidades de gestión universitaria. Entre otras, las de Vicerrector de Alumnos, Director de Programas Internacionales y Secretario Académico del Real Instituto Universitario de Estudios Europeos, todas ellas en la Universidad CEU San Pablo.

Es Académico correspondiente de la Real Academia de Jurisprudencia y Legislación, presidente del Foro Internacional de Ética y Derecho de la Información (FIEDI) y miembro de la Asociación de Constitucionalistas de España y de la International Society of Public Law (ICON-S).

Universidad CEU San Pablo

Vicerrectorado de Relaciones Institucionales y Secretaría General

Isaac Peral 58, 28040 Madrid

Teléfono: 91 456 63 00

www.uspcceu.es