

Dignidad humana y libertad de conciencia en la era de la Inteligencia Artificial: la aportación de la Iglesia Católica a la fundamentación de los neuroderechos

Introducción. I. IA y dignidad: criterios éticos desde el pensamiento social cristiano. II. El derecho a formar libremente la conciencia. III. Regulación europea y gobernanza internacional de la neurotecnología. IV. Neuroderechos y tutela de la esfera mental: fundamentos y categorías emergentes. Conclusiones

Resumen: El avance de la inteligencia artificial y de las neurotecnologías plantea riesgos directos para la privacidad mental, la integridad cognitiva y la libertad de conciencia, tradicionalmente protegidas como parte del *forum internum*. La posibilidad de inferir o modular procesos mentales mediante sistemas algorítmicos cuestiona la suficiencia del marco clásico de los derechos humanos. En este contexto, los neuroderechos emergen como propuesta normativa destinada a reforzar la tutela de la autonomía mental. El presente trabajo analiza su fundamento en la dignidad humana y su encaje en el Derecho internacional, interpretándolos como actualización jurídica de la protección del *forum internum* en la era digital.

Palabras clave: dignidad humana, libertad de conciencia, *forum internum*, ética, neuroderechos,

«La dimensión teológica se hace necesaria para interpretar y resolver los actuales problemas de la convivencia humana». (*Centesimus annus*¹, 55)

Introducción

El desarrollo acelerado de la inteligencia artificial (en adelante IA) y de las neurotecnologías² ha inaugurado un escenario de profundas transformaciones en las formas de comprender, intervenir y eventualmente modificar la cognición humana. Estos avances técnicos, que incluyen entre otros, el procesamiento masivo de datos neuronales, la decodificación de estados mentales, las técnicas de neuromodulación y la integración de interfaces cerebro-máquina, han dado lugar a desafíos éticos y jurídicos sin precedentes, al poner en riesgo dimensiones tradicionalmente consideradas inviolables, como la privacidad mental, la autonomía de la voluntad y la integridad de la

¹ Carta Encíclica del Sumo Pontífice Juan Pablo II, a sus hermanos en el episcopado, al clero, a las familias religiosas, a los fieles y a todos los hombres de buena voluntad en el centenario de la *Rerum Novarum*, de 1 de mayo de 1991. https://www.vatican.va/content/john-paul-ii/es/encyclicals/documents/hf_jp-ii_enc_01051991_centesimus-annus.html

² La neurotecnología es un campo en constante evolución que busca desentrañar el cerebro humano y desarrollar tecnologías. Así, varios proyectos de investigación, tanto gubernamentales como privados, están avanzando en este campo con objetivos que van desde el mapeo de la actividad cerebral hasta la creación de interfaces cerebro-máquina. A la luz de las implicaciones éticas, sociales y jurídicas que emergen con el avance de la neurotecnología, surge el concepto de neuroderechos, estrechamente vinculado al Proyecto BRAIN (*Brain Research through Advancing Innovative Neurotechnologies*), una ambiciosa iniciativa científica lanzada en 2013 por el entonces presidente de los Estados Unidos, Barack Obama. Este proyecto tuvo como objetivo principal impulsar el conocimiento sobre el funcionamiento del cerebro humano y fomentar el desarrollo de tecnologías destinadas al tratamiento de enfermedades como el Alzheimer, la epilepsia o las lesiones cerebrales. *Vid.* Peña- Cuellar, D.M., y otros, “Neurotecnología y derechos humanos: Evaluando la necesidad de los neuroderechos”, en *Revista Academia & Derecho*, Año 15, No. 29, 2024, pp. 1-26

identidad personal. La posibilidad real de influir en la conducta, inferir preferencias internas o incluso potenciar capacidades cognitivas plantea interrogantes sobre la suficiencia del actual sistema de derechos humanos para enfrentar estos riesgos crecientes.

En este contexto ha surgido con fuerza la propuesta de los denominados neuroderechos, un campo normativo al alza, orientado a salvaguardar la integridad mental y garantizar la libertad cognitiva en la era de la IA y la neurotecnología. Su finalidad es ofrecer un marco de protección reforzada frente a tecnologías capaces de irrumpir en el *forum internum* —la esfera interna de pensamiento y conciencia— identificado como el núcleo más protegido de la dignidad humana. El debate actual oscila entre quienes sostienen que estos derechos pueden integrarse en el sistema de derechos humanos existentes y quienes, por el contrario, defienden la necesidad de reconocer una nueva categoría jurídica ante la especificidad de los riesgos neurotecnológicos.

Con esta perspectiva, no podemos olvidar la posición de la Iglesia Católica. En las últimas décadas ha desarrollado un cuerpo doctrinal y bioético particularmente atento a los desafíos antropológicos que plantean las nuevas tecnologías cognitivas. Documentos como *Gaudium et Spes* o el *Rome Call for AI Ethics*, así como los mensajes del Papa Francisco y de León XIV sobre la IA, reafirman la centralidad de la persona humana y la inviolabilidad de la conciencia como fundamentos éticos ineludibles para orientar la innovación tecnológica. Desde esta perspectiva, la reflexión eclesial ofrece criterios antropológicos y morales que resultan especialmente valiosos para evaluar las neurotecnologías.

I. IA y dignidad: criterios éticos desde el pensamiento social cristiano

La dignidad aparece como la razón por la cual la persona posee un ámbito inviolable de interioridad. De este modo, la Declaración *Dignitatis Humanae*³ sostiene explícitamente que la libertad religiosa “se funda en la dignidad misma de la persona humana”, apareciendo así, como la razón por la que una persona posee un ámbito de

³ Declaración sobre la libertad religiosa de 7 de diciembre de 1965. https://www.vatican.va/archive/hist_councils/ii_vatican_council/documents/vat-ii_decl_19651207_dignitatis-humanae_sp.html.

autodeterminación inviolable. La Iglesia sostiene una visión integral de la persona, fundada en su dignidad como ser creado a imagen de Dios.

La denominada “Doctrina Social de la Iglesia”⁴ (en adelante DSI) constituye un patrimonio surgido del encuentro del mensaje evangélico y de sus exigencias éticas con los problemas que surgen en la vida de la sociedad. Desde esta perspectiva, ofrece un marco ético para discernir sobre los desafíos contemporáneos, entre los que se encuentra la IA y todo ello lo hace a la luz de la dignidad humana y el bien común. Desde sus principios fundamentales enunciados en el Capítulo cuarto -dignidad de la persona, bien común, solidaridad y subsidiariedad-, la Iglesia sostiene que la IA debe estar al servicio del ser humano, respetar sus derechos, promover la justicia y evitar nuevas formas de exclusión o control. En este sentido, los principios del pensamiento social católico, derivados del Evangelio y el magisterio, se aplican a la IA para equilibrar innovación con ética humana.

De forma recíproca a los principios enunciados, estable los valores sociales que “deben ser atribuidos a determinados aspectos del bien moral que aquéllos pretenden conseguir”, “los valores requieren tanto la práctica de los principios fundamentales de la vida social como el ejercicio de las actitudes morales correspondientes a los valores mismos” (nº197). De este modo, nacerán como guía para el uso ético de la IA, basados en la verdad, la libertad y la justicia. La “verdad” exige que la IA se programe para verificar datos y promover la veracidad, “Las personas y los grupos sociales cuanto más se esfuerzan por resolver los problemas sociales según la verdad, tanto más se alejan del arbitrio y se adecúan a las exigencias objetivas de la moralidad” (nº198). Por su parte, la “libertad” siendo un símbolo máximo de la dignidad humana (nº199), defiende la garantía y protección frente a los algoritmos que efectúan formas de control tecnológico invadiendo la autonomía humana. Finalmente, la “justicia”, dentro de la denominada “justicia social”, se recoge para establecer una dimensión concerniente a “los aspectos sociales, políticos y económicos y, sobre todo, a la dimensión estructural de los problemas y las soluciones correspondientes” (nº201), todo ello a la luz de formas no invasivas por parte de la IA. En general, podemos advertir como los valores subrayan la centralidad de la persona, la responsabilidad moral de quienes diseñan y usan estas

⁴ Documento elaborado a petición del Santo Juan Pablo II para recoger la enseñanza social de la Iglesia. https://www.vatican.va/roman_curia/pontifical_councils/justpeace/documents/rc_pc_justpeace_doc_2006_0526_compendio-dott-soc_sp.html#PRESENTACI%C3%93N

tecnologías, la transparencia, y la necesidad de una supervisión humana que garantice que el progreso tecnológico contribuya a una sociedad más humana, justa y solidaria.

Es por ello necesario destacar el deseo de la Iglesia de participar en los distintos debates surgidos sobre la IA y que como podemos comprobar, afectan directamente al presente y futuro de la familia humana. Como estima la Santa Sede, “junto con su extraordinario potencial para beneficiar a la familia humana, el rápido desarrollo de la inteligencia artificial también plantea cuestiones más profundas sobre el uso correcto de esta tecnología para generar una sociedad global más auténticamente justa y humana, la inteligencia artificial es ante todo un instrumento”⁵.

La Pontificia Academia para la Vida promovió la organización de la conferencia “*RenAIssance: por una inteligencia artificial humanista*”, de la que el 28 de febrero de 2020 nació la firma del documento “*Rome Call for AI Ethics*”⁶ (Llamamiento de Roma por la Ética de la Inteligencia Artificial). Bajo los principios de transparencia, inclusión, responsabilidad y seguridad, la intención de la Santa Sede era propugnar una perspectiva de desarrollo de la IA no centrada en su base tecnológica sino más bien, tomando en consideración la humanidad y el medio ambiente. De este modo, “se trata de promover la responsabilidad compartida entre organizaciones internacionales, gobiernos, instituciones y el sector privado para crear un futuro en el que la innovación digital y el progreso tecnológico otorguen a la humanidad su centralidad”. Se circunscribe a tres áreas de impacto, la ética, la educación y el derecho, todo ello, por medio del establecimiento de seis principios básicos de actuación; Transparencia: los sistemas de IA deben ser comprensibles para todos; Inclusión: los sistemas no deben discriminar a nadie, todo ser humano tiene la misma dignidad; Responsabilidad: detrás de cada sistema de IA debe haber un responsable de lo que hace la máquina; Imparcialidad: los sistemas no deben seguir pautas ni crear sesgos, debiendo mantener la equidad y la dignidad humana; Fiabilidad: los sistemas de IA deben poder funcionar de forma fiable; Seguridad y privacidad: los sistemas de IA deben funcionar de forma segura y respetar la privacidad de los usuarios.

⁵ Discurso del papa Francisco en la Sesión del G7 sobre Inteligencia Artificial, el 14 de junio de 2024. <https://www.vatican.va/content/francesco/es/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html>

⁶ <https://www.romecall.org/the-call/>

En 2021 se creaba la Fundación *RenAIssance*⁷, con el fin de promover la reflexión antropológica y ética de las nuevas tecnologías en la vida humana. Como establecen, su ideal determina el compromiso científico, eclesial, cultural, empresarial y profesional en la sociedad, asimismo, fomentar las iniciativas científicas y la colaboración con organismos internacionales, estados soberanos, universidades, centros de investigación, empresas privadas y públicas que desarrollen actividades, servicios y estudios en el campo de la IA para difundir la Convocatoria de Roma por la Ética de la IA y promover la recaudación de fondos para apoyar estas actividades.

Por su parte, el Dicasterio para la Cultura y la Educación en colaboración con el Instituto de Tecnología, Ética y Cultura⁸ (ITEC), ubicado en Santa Clara University, California, dieron a conocer el documento “*Ethics in the Age of Disruptive Technologies: An Operational Roadmap*”⁹ (La ética en la era de las tecnologías disruptivas: un mapa de ruta operativo) en el que se convocaba a empresas, sociedad civil, academia y gobiernos a promover una reflexión profunda sobre el impacto de la tecnología en la humanidad.

Sobre esta base y de forma directa, las encíclicas “*Laudato Si*”¹⁰ y “*Fratelli Tutti*”¹¹, promulgadas por el papa Francisco, constituyen una profundización significativa de los principios de la DSI anteriormente desarrollados, frente a los desafíos éticos planteados por el desarrollo tecnológico contemporáneo. En ambos textos, se ofrece una lectura crítica del progreso técnico cuando este se desvincula de la dignidad humana, del bien común y de la fraternidad social, principios fundamentales del pensamiento social cristiano.

⁷ Establecida el 12 de abril de 2021 por el Papa Francisco con personalidad jurídica canónica pública consta inscrita en el Registro de Organizaciones sin Fines de Lucro de la Gobernación del Estado de la Ciudad del Vaticano. <https://www.romecall.org/renaissance-foundation/>

⁸ El ITEC es una iniciativa del Centro Markkula de Ética Aplicada de la Universidad de Santa Clara, con el apoyo y la colaboración del Centro de Cultura Digital del Vaticano, perteneciente al Dicasterio para la Cultura y la Educación. El Instituto convoca a líderes del mundo empresarial, la sociedad civil, el mundo académico, los gobiernos y todas las tradiciones religiosas para promover una reflexión más profunda sobre el impacto de la tecnología en la humanidad. <https://www.scu.edu/institute-for-technology-ethics-and-culture/itec-principles/>

⁹ <https://www.scu.edu/institute-for-technology-ethics-and-culture/itec-handbook/>

¹⁰ Carta Encíclica. 24 de mayo de 2015. https://www.vatican.va/content/francesco/es/encyclicals/documents/papa-francesco_20150524_enciclica-laudato-si.html

¹¹ Carta Encíclica. 3 de octubre de 2020. https://www.vatican.va/content/francesco/es/encyclicals/documents/papa-francesco_20201003_enciclica-fratelli-tutti.html

En la primera de ellas, señala el papa como “hay un modo de entender la vida y la acción humana que se ha desviado y que contradice la realidad hasta dañarla”, para su reflexión se centra en “el paradigma tecnocrático dominante y en el lugar del ser humano y de su acción en el mundo”. Ensalza el poder de la humanidad y la escasa garantía de que vaya a utilizarlo bien¹² (nº.104), y recoge “el problema fundamental en el modo en el que la humanidad ha asumido la tecnología” (nº.106), pues estos elementos junto con la metodología y los objetivos señalados en la tecnociencia condicionan la vida de las personas y el funcionamiento de la sociedad (nº.107), concluyendo que «el desarrollo humano integral supone el desarrollo responsable de la tecnología» (nº 131).

Por su parte, en la encíclica *Fratelli Tutti* profundiza la dimensión social del análisis, centrando su atención en el impacto de las tecnologías digitales sobre los vínculos humanos y la cohesión social. “En el mundo de hoy persisten numerosas formas de injusticia” (nº.22), en este orden, “se constata un deterioro de la ética, que condiciona la acción internacional, y un debilitamiento de los valores espirituales y del sentido de responsabilidad” (nº.29).

En ambas nos encontramos con un ejercicio reflexivo sobre la relación entre la tecnología, la dignidad humana y la fraternidad social. Desde esta perspectiva, cada vez más, los avances alcanzados en la IA han ido provocando un impacto más profundo en la actividad humana, en la vida personal y social y en la política y la economía. La preocupación por estos hechos infirió en el mensaje del papa Francisco dedicado en la Jornada Mundial de la Paz¹³ del 1 de enero de 2024 bajo el título “Inteligencia Artificial y Paz”. En su contenido destaca la necesidad de orientar de manera responsable la concepción y el uso de la IA para que esté al servicio de la humanidad y para ello, exige que la reflexión ética se extienda al ámbito de la educación y del derecho. La protección de la dignidad de la persona es indispensable para que el desarrollo tecnológico contribuya a promover la justicia y la paz en el mundo, entiendo “la IA como una galaxia de realidades distintas y no podemos presumir *a priori* que su desarrollo aporte

¹² En este sentido se manifestó posteriormente en la Exhortación Apostólica “*Laudate Deum*” de 4 de octubre de 2023, afirmando que “Dentro del “paradigma tecnocrático”, la inteligencia artificial y las últimas novedades tecnológicas parten de la idea de un ser humano sin límite alguno, cuyas capacidades y posibilidades podrían ser ampliadas hasta el infinito gracias a la tecnología. Así, el paradigma tecnocrático se retroalimenta monstruosamente”.

¹³ https://www.vatican.va/content/francesco/es/apost_exhortations/documents/20231004-laudate-deum.html
<https://www.vatican.va/content/francesco/es/messages/peace/documents/20231208-messaggio-57giornatamondiale-pace2024.html>

una contribución benéfica al futuro de la humanidad y a la paz entre los pueblos. Tal resultado positivo sólo será posible si somos capaces de actuar de forma responsable y de respetar los valores humanos fundamentales como «la inclusión, la transparencia, la seguridad, la equidad, la privacidad y la responsabilidad».

En enero de 2025 el Vaticano publicó el documento titulado “*Antiqua et nova*”¹⁴, abordando de nuevo las cuestiones filosófico-antropológicas y éticas planteadas por la IA. El texto, firmado por los Dicasterios para la Doctrina de la Fe y para la Cultura y la Educación, destaca las distinciones entre la inteligencia humana y la artificial. Este documento fue la base para que el nuevo papa, León XIV, en su primer discurso al Colegio Cardenalicio el 10 de mayo de 2025¹⁵, advirtiese sobre los avances de la IA y los nuevos desafíos de la misma, llamando a un uso responsable basado en la ética que integre la fe católica con la innovación tecnológica como criterio de la dignidad humana¹⁶.

II. El derecho a formar libremente la conciencia.

Una de las afirmaciones iniciales que debemos realizar es que “la primera dimensión de la libertad de conciencia es la libre formación de nuestra conciencia”¹⁷. En lo que respecta a su dimensión interna, este derecho se configura como la facultad inherente a toda persona para poseer, elegir, mantener, ejercer y modificar libremente sus ideas, principios, valores, creencias y convicciones personales¹⁸. Dicha esfera

¹⁴https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_dcf_doc_20250128_antiqua-et-nova_sp.html

¹⁵ <https://www.vatican.va/content/leo-xiv/es/speeches/2025/may/documents/20250510-collegio-cardinalizio.html>

¹⁶ Asimismo, en su mensaje dirigido a los participantes de la Segunda Conferencia Anual sobre Inteligencia Artificial, celebrada el 17 de junio de 2025, subrayó la urgencia de una reflexión profunda y sostenida sobre las implicaciones éticas de la IA y su impacto en el porvenir de la humanidad. Recalcó la presencia de los asistentes evidenciaba la necesidad de un diálogo continuo en torno a la dimensión moral inherente a estas tecnologías emergentes. Desde esta perspectiva, enfatizó que el desarrollo y la gobernanza de la IA deben orientarse por principios éticos claros, promoviendo un progreso inclusivo que respete la dignidad intrínseca de toda persona. Estas manifestaciones fueron utilizadas posteriormente en un mensaje dirigido a líderes corporativos del sector tecnológico el 20 de junio de 2025. Posicionado en contra de una concepción reduccionista de la inteligencia basada únicamente en la acumulación de datos, entendiendo que, la verdadera inteligencia no consiste en disponer de grandes volúmenes de información, sino en la búsqueda del sentido auténtico de la vida, el pontífice cuestionó un enfoque tecnocrático que corre el riesgo de marginar las dimensiones espirituales y humanistas de la existencia. <https://www.vatican.va/content/leo-xiv/es/messages/pont-messages/2025/documents/20250617-messaggio-ia.html>

¹⁷ Rodríguez García, J.A., *Laicidad e inteligencia artificial: la implementación y el alineamiento de la laicidad como valor en los sistemas de inteligencia artificial*, Dykinson, 2024, pág. 172.

¹⁸ En palabras de Llamazares, es *una vivencia íntima, la más íntima de las vivencias, fuente de la moral personal, y consiguientemente de la integridad moral de la persona*”, Llamazares Fernández, D., “Libertad de conciencia, una libertad implícita en la Constitución”, en *Cuestiones de Pluralismo*, Vol. 2, n°1, 2022.

interna constituye un ámbito reservado de autodeterminación y autonomía personal, estrechamente vinculado al desarrollo de la personalidad y a la dignidad humana. Su protección no solo garantiza la libertad religiosa, sino que tutela la esfera interna del pensamiento, las convicciones morales y la autodeterminación espiritual, es decir, el espacio donde se forma la identidad ética de la persona¹⁹.

Por su propia naturaleza y como indica el Comité de Derechos Humanos de Naciones Unidas (en adelante CDNU) en su Observación General nº 22 (1993)²⁰, este ámbito interno reviste un carácter inviolable y absoluto, en tanto que se sitúa fuera de toda posibilidad de injerencia o control externo. En consecuencia, no resulta jurídicamente admisible intervención alguna en este espacio íntimo de la conciencia, ni por parte de los poderes públicos ni de terceros, lo que determina su configuración como un derecho ilimitado en su dimensión interna, inmune a restricciones o condicionamientos. Desde esta perspectiva, la libre conformación de la conciencia no es solo una libertad negativa (ausencia de coacción), sino también una garantía estructural contra cualquier forma de manipulación directa o indirecta de los procesos mentales.

Ahora bien, los avances tecnológicos, los sistemas de IA, entre otros, pueden llegar a condicionar y manipular la propia personalidad²¹. Este hecho nos permite avanzar como desde la libertad de conciencia se pueden derivar tres principios que están íntimamente relacionados con el funcionamiento de los sistemas de IA²²: el derecho a no revelar las propias ideas, creencias, convicciones, pensamientos y opiniones²³; el derecho a que no se manipulen nuestras ideas, creencias, convicciones, pensamientos y

¹⁹ La libre conformación de la conciencia se sitúa en el núcleo del sistema internacional de derechos humanos como dimensión intangible del *forum internum*. La Declaración Universal de los Derechos Humanos establece en su artículo 18 que, “Toda persona tiene derecho a la libertad de pensamiento, de conciencia y de religión”. Este reconocimiento fue desarrollado jurídicamente con fuerza vinculante en el art. 18 del Pacto Internacional de Derechos Civiles y Políticos. Por su parte, en el ámbito europeo, el artículo 9 del Convenio Europeo de Derechos Humanos reconoce igualmente la libertad de pensamiento, conciencia y religión. Asimismo, la jurisprudencia del Tribunal Europeo de Derechos Humanos ha insistido en que: La libertad de pensamiento es una de las bases de la “sociedad democrática”; El Estado debe mantener una posición de neutralidad; El núcleo interno de convicciones es inmune a intervención estatal. En sentencias como *Kokkinakis c. Grecia* (1993), el Tribunal reafirma que la libertad de conciencia constituye uno de los fundamentos de la identidad personal y colectiva.

²⁰<https://hrlibrary.umn.edu/hrcommittee/Sgencom22.html#:~:text=No%20se%20puede%20restringir%20a,la%20educaci%C3%B3n%20religiosa%20y%20moral>.

²¹ Rodríguez García, J.A., “Laicidad e inteligencia artificial...”, *Op. Cit.*, pág. 173.

²² *Idem*. Págs. 173-180.

²³ Pese a que el texto constitucional reconoce que “nadie podrá ser obligado a declarar sobre su ideología, religión, creencias” por ejemplo, las redes sociales analizan los datos de nuestros perfiles, incluido el ideológico y el religioso, analizan nuestro estado emocional, lo que implica una vulneración del antedicho derecho.

opiniones²⁴; y el derecho a no ser penalizados por nuestras propias ideas, creencias, convicciones, pensamientos y opiniones²⁵. Esto nos permite entender dos formas de actuar sobre el cerebro; leyendo el cerebro e interviniendo en él para cambiar su actividad “escribiendo” nuestro cerebro²⁶. En el primero de los casos, mediante las “interfaces” cerebro-ordenador (BCIs) con los que se puede explorar o visualizar médicamente, diagnosticar, tratar y curar el cerebro, pero también a través de estas “interfaces” se podría llegar a controlar la actividad del cerebro, con lo que se abre la posibilidad de monitorizar, investigar, evaluar, manipular, emular, obtener y/o extraer información (pensamientos, ideas, datos...) del cerebro. En segundo lugar ¿cabría la posibilidad de escribir en el cerebro humano? ¿cabría, por tanto, introducir pensamientos o ideas en el cerebro de una persona? Esta situación supondría abstraer de la persona el tener unos pensamientos, ideas u opiniones propios, su capacidad de decisión y de acción, así como la plena autonomía de la voluntad.

Desde esta perspectiva, la cuestión jurídica emergente es si la manipulación algorítmica predictiva o persuasiva puede considerarse una forma indirecta de afectación del *forum internum*. Esto nos aboca a situar un puente entre la tradición clásica de los derechos fundamentales y los desafíos contemporáneos derivados de la IA y las neurotecnologías. La continuidad entre dignidad humana, conciencia y neuroderechos resulta así jurídicamente coherente.

III. Regulación europea y gobernanza internacional de la neurotecnología

La IA permite inferir patrones cognitivos a partir de datos conductuales aparentemente neutros, como movimientos del cursor, voz, ritmo cardíaco o actividad digital. Esta capacidad genera riesgos para el derecho a la privacidad y a la libertad de

²⁴ Ni los poderes públicos ni particulares pueden interferir ni injerir en la dimensión interna, de otro modo, se cercenaría la autonomía de la voluntad. La vigilancia emocional incide en la manipulación de nuestras conciencias, se define así, como “sistema de reconocimiento de emociones” al sistema de IA destinado a distinguir o inferir las emociones o las intenciones de las personas a través de sus datos biométricos.

²⁵ Reconducido al ámbito interno, la posibilidad de “leer” nuestras ideas, pensamientos y emociones, puede conllevar que se penalicen o sancionen por ello.

²⁶ “El cerebro que, para algunos no es un órgano más, “sino el que genera toda nuestra actividad mental y cognitiva. Todos nuestros pensamientos, percepciones, imaginación, recuerdos, decisiones y emociones son generados por el disparo orquestado de circuitos neuronales en nuestros cerebros” en Yuste, R., Goering, S., y otros, “*Four ethical priorities for neurotechnologies and AI*”, en *Nature*, vol. 551, 2017, p. 159.

conciencia y pensamiento, tradicionalmente entendido como inviolable por pertenecer a la esfera interna y más íntima del individuo. Las tecnologías desarrolladas a través de la IA aplicadas entre otros, a sistemas de vigilancia, publicidad conductual o análisis predictivo pueden afectar la autonomía decisional.

Desde esta perspectiva, la Unión Europea (en adelante UE) ha avanzado en la regulación de estas prácticas²⁷ con la aprobación del Reglamento de Inteligencia Artificial²⁸ (*AI Act*, 2024) que restringe el desarrollo y aplicabilidad de los sistemas de manipulación subliminal y usos gubernamentales de reconocimiento emocional. Previamente, el Reglamento General de Protección de Datos²⁹ ya recogía el derecho a no ser objeto de decisiones basadas exclusivamente en tratamientos automatizados sin garantías adicionales.

La *AI Act*. establece una prohibición absoluta sancionando la manipulación cognitiva como práctica incompatible con la dignidad humana y la autonomía mental (art. 5, apartado 1, letra a). Directamente relacionada con la protección de la esfera interna, la imposibilidad de que algoritmos infieran elementos centrales de la identidad ideológica, religiosa y personal, se encuentra la limitación en el uso de “Sistemas de IA destinados a inferir emociones de una persona en el lugar de trabajo o en instituciones educativas, así como sistemas destinados a inferir características sensibles de las personas físicas, basándose en datos biométricos, tales como origen étnico, opiniones políticas, convicciones religiosas o filosóficas, orientación sexual o estado de salud” (art. 5, apartado 1, letra c).

A raíz de una resolución del CDNU, la Oficina del Alto Comisionado para los Derechos Humanos iniciaron también en 2023 la preparación de un estudio sobre neurotecnología y derechos humanos³⁰.

²⁷ El Consejo de Europa, mediante el Convenio de Oviedo (1997), establece la primacía de la persona sobre los avances biomédicos, formulación que sirve de base para incluir la neuroprotección.

²⁸ 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) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 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). https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=OJ:L_202401689

²⁹ Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). <https://eur-lex.europa.eu/eli/reg/2016/679/oj/spa>

³⁰ Resolución de la ONU adoptada por el Consejo de Derechos Humanos el 6 de octubre de 2022. <https://docs.un.org/es/A/hrc/RES/51/3>.

En 2025 nos encontramos ante el primer instrumento a nivel internacional, la Recomendación sobre la Ética de la Neurotecnología³¹, definiendo las reglas globales para que las tecnologías puedan leer, registrar o modular la actividad cerebral. En su contenido, se reconoce la existencia de los denominados “datos neuronales”, que conlleva identificar la actividad cerebral con las emociones, los rasgos de la personalidad, los estados mentales, patrones cognitivos, etc. además de entre otros, señala como objetivos “la garantía y protección, el respeto y la promoción de los derechos humanos y las libertades fundamentales, la dignidad y la igualdad (...)”, indicando que la dignidad es la base de los derechos humanos y por tanto, el respeto, protección y promoción de esa dignidad humana es esencial.

Dentro de los Valores y Principios recogidos en la Recomendación, debemos destacar el pronunciamiento sobre la “autonomía y libertad de pensamiento” y la “protección de los datos neuronales y datos que permiten inferencias mentales”. En lo que respecta al primero, estima “la protección de la autonomía y la libertad de pensamiento a lo largo del ciclo de vida de la neurotecnología, así como la integridad mental y física contra interferencias no deseadas o dañinas”. La autonomía es reconducida a su aspecto individual y relacional, de modo que podemos hablar de un margen apreciativo de “autodeterminación” entendida ésta como la toma de decisiones libres, informadas y voluntarias respecto del uso de la neurotecnología, sin coerción explícita ni implícita. Los procedimientos de consentimiento deben ser claros, comprensibles, transparentes, etc., evitando en todo caso, el uso reconducido a una influencia indebidamente, manipulada o coactiva de pensamientos o decisiones internas o externas. Por otro lado, y en lo referente a los riesgos para la “protección de los datos neuronales y de datos que permiten interferencias mentales” provienen de las tecnologías que recolectan datos biométricos que permiten inferir estados mentales con los riesgos que conllevan para la privacidad. La recopilación, procesamiento y compartición de datos neuronales solo deben realizarse con consentimiento previo, libre e informado (salvo emergencias médicas). En este orden, otras muchas organizaciones intergubernamentales también han considerado los neuroderechos como instrumentos útiles de gobernanza³².

³¹ Aprobada por la UNESCO el 11 de noviembre de 2025. <https://www.unesco.org/es/legal-affairs/recommendation-ethics-neurotechnologies>

³² El Consejo de Europa lanzó el Plan de acción estratégico (2020-2025) para explorar las implicaciones de la neurotecnología para los derechos humanos, renovado ahora para los años 2026-2030), cuyo

IV. Neuroderechos y tutela de la esfera mental: fundamentos y categorías emergentes

Sin poder advertir un estudio más complejo en el que se indique el momento de la aparición de los neuroderechos como concepto, alumbramos una necesaria aproximación a su definición.

Como hemos venido advirtiendo, la neurotecnología es un campo en constante evolución que busca desentrañar el cerebro humano y desarrollar tecnologías, de ahí que haya habido un debate constante en las últimas fechas sobre la necesaria creación de una nueva categoría dentro de los derechos humanos, los denominados “neuroderechos”. “Las nuevas tecnologías plantean un peligro de manipulación e intervención de nuestro cerebro y sistema nervioso, lo que implica una amenaza a la integridad mental”³³, de ahí que, podamos entender por estos, los principios éticos, legales, sociales o naturales de libertad o derecho relacionados con el ámbito cerebral y mental de una persona. En

objetivo principal es la garantía de que los derechos humanos, la dignidad, la equidad y la democracia sigan en el centro de los sistemas de biomedicina y salud. <https://www.coe.int/en/web/human-rights-and-biomedicine/strategic-action-plan#:~:text=Designed%20to%20protect%20human%20dignity.Artificial%20Intelligence%20in%20healthcare>

Por su parte, la OCDE aprobó en 2019 una Recomendación sobre innovación responsable en neurotecnología, orientada a guiar a gobiernos e innovadores en la gestión de los desafíos éticos, legales y sociales que estas tecnologías generan, al tiempo que fomenta su desarrollo. Esta Recomendación reconoce la importancia de proteger la privacidad mental y la libertad cognitiva, y establece nueve principios, entre ellos la prioridad de la seguridad, la protección de los datos cerebrales y la prevención de usos indebidos. En España, la Carta de Derechos Digitales (2021) se configuró como uno de los marcos programáticos más avanzados en Europa para la protección de derechos fundamentales en entornos digitales. Aunque no tiene carácter vinculante, sirve como guía para futuras regulaciones en ámbitos como la inteligencia artificial, los algoritmos, la protección de datos y las nuevas formas de interacción cognitiva entre humanos y sistemas digitales. La Carta adopta un enfoque de humanismo digital alineado con la tradición europea de protección de la dignidad humana. Sus objetivos principales son proteger los derechos fundamentales frente a riesgos tecnológicos, orientar las políticas públicas en materia digital, promover un desarrollo tecnológico centrado en la persona y anticipar derechos emergentes relacionados con la integridad psicológica y la privacidad mental. Aunque no menciona expresamente los “neuroderechos”, sí incorpora referencias a la privacidad mental, la integridad cognitiva y la protección frente a tecnologías persuasivas.

³³ De Asís, R. *Derechos y tecnologías*, ed., España: Dykinson, 2022. El mismo autor en “Sobre la propuesta de los Neuroderechos”, en *Derechos y Libertades: Revista De Filosofía Del Derecho Y Derechos Humanos*, 47, págs. 51-70. El autor considera que este tipo de propuestas —reconocer nuevas categorías de derechos humanos— está condicionado por el argumento de la inflación, el cual establece que cuanto más se extiende la lista de derechos, menores son las posibilidades de otorgarles una protección especial.

otras palabras, son las reglas normativas fundamentales para la protección y preservación del cerebro y la mente humanos³⁴.

Desde esta perspectiva de definición abstracta, los estudios se han centrado en tratar de concretar la categoría de neuroderechos que podrían ser recogidos para proteger la esfera mental humana frente a riesgos derivados de la lectura, alteración o ampliación tecnológica de la actividad cerebral. Inicialmente podríamos estar ante tres derechos³⁵: el derecho a la privacidad mental, el derecho a la integridad mental y el derecho a un acceso equitativo al aumento mental³⁶. Ahora bien, nuestro estudio nos obliga a sintetizarlos tomando como referencia la libertad de pensamiento y de conciencia, el *fórum internum* de la persona³⁷. La “mente constituye el último reducto inviolable de la persona”, por ello, cualquier tecnología capaz de acceder o modificar su contenido activa automáticamente la necesidad de nuevos estándares de protección³⁸.

1. Privacidad mental y libertad de conciencia interna

En primer lugar, el derecho a la privacidad mental vendría a impedir la obtención, decodificación o tratamiento no consentido de datos cerebrales, protege el espacio interior donde se generan pensamientos y creencias. Los avances en algoritmos de decodificación neuronal permiten, con un nivel de precisión creciente, inferir contenidos internos de la vida mental, incluyendo recuerdos, estados emocionales,

³⁴ Se utilizará la versión traducida al español: Ienca, M., Andorno, R., “Hacia nuevos derechos humanos en la era de la neurociencia y neurotecnología” (Wajnerman Paz, A., trad.), en *Análisis Filosófico*, 2021, 41: 1, p. 143. (trabajo original publicado en 2017).

³⁵ Tomando lo dispuesto por Ienca, su clasificación diferencia cinco neuroderechos. En primer lugar, los derechos derivados de la libertad de pensamiento, donde incluye a la libertad cognitiva (el derecho a la autodeterminación mental), el derecho a la agencia (el derecho al reconocimiento de la capacidad de elección o al libre albedrío), la libertad mental (el derecho al control consciente sobre la propia mente) y la libertad de pensamiento. En segundo lugar, se refiere a los derechos derivados de la privacidad, donde se situaría el derecho a la privacidad mental (el derecho de las personas contra la intrusión no consentida de terceros en sus datos cerebrales, contra la recopilación no autorizada de esos datos y contra su divulgación) y el derecho a la neuroprivacidad (el derecho a la protección de los datos neuronales o cerebrales). En tercer lugar, los derechos derivados del bien integridad, donde se encuentra el derecho a la integridad mental (el derecho de las personas a ser protegidas de manipulaciones ilícitas y nocivas de su actividad mental). En cuarto lugar, los derechos derivados de la identidad, donde está el derecho a la identidad personal, que posee dos proyecciones: (i) como derecho a preservar “la identidad personal de las personas y la continuidad de su mentalidad”; y, (ii) como derecho de controlar la integridad física y la integridad mental. En quinto lugar, habla de dos derechos relacionados con la promoción de requisitos necesarios para la realización de los otros derechos: el derecho a un acceso igual a la mejora mental y el derecho a la protección contra el sesgo algorítmico. *Vid.* Ienca, M., Andorno, R., “Hacia nuevos derechos humanos en la era de la neurociencia...”, *Op. Cit.*

³⁶ Yuste, R., Goering, S., “Four ethical priorities for neurotechnologies and AI”, en *Nature*, 551, 2017, págs. 159–163. <https://www.nature.com/articles/551159a>

³⁷ Ienca, M., Andorno, R., “Hacia nuevos derechos humanos en la era de la neurociencia y...”, *Op. Cit.*

³⁸ *Vid.* Yuste, R., *Neuroderechos. Un viaje hacia la protección de los que nos hace humanos*, ed. Paidós, 2025.

preferencias profundas o patrones cognitivos inconscientes³⁹. Tales datos no pueden considerarse meramente una categoría más de información personal, sino datos cuya protección debe ser absoluta por pertenecer a la rama más íntima y privada del ser. Son la garantía de la libertad de conciencia en su aspecto interno, pues esta libertad implica el derecho a mantener pensamientos, creencias, dudas, preferencias o estados afectivos frente a injerencias no consentidas por parte del sujeto. Cualquier forma de acceso no consentido a ellos constituiría una violación directa del *forum internum*, que —a diferencia de la libertad religiosa o ideológica externa— no admite limitaciones. Así, el neuroderecho a la privacidad mental garantizaría que “lo que ocurre en la mente” permanezca exclusivamente bajo control del sujeto, preservando el espacio donde se forman las creencias, juicios morales y procesos de deliberación personal, todos ellos esenciales para la autonomía espiritual y cognitiva.

2. Integridad mental y protección del *forum internum*

En segundo lugar, el denominado derecho a la integridad mental. En este caso, se propone salvaguardar la unidad, estabilidad y autenticidad de la experiencia mental, donde se haya la identidad personal, la personalidad y cualquier proceso decisorio. Se busca impedir cualquier forma de manipulación, interferencia o alteración de la actividad neuronal que comprometa la autonomía cognitiva y volitiva. Este derecho tiene una relación directa con la libertad de conciencia interna, porque protege la unidad, autenticidad y coherencia de la experiencia mental, componentes esenciales para que la persona pueda pensar, creer y decidir libremente, por ello, se incluiría aquí el derecho a proteger el libre albedrío y la libertad cognitiva, impidiendo que factores externos puedan interferir de forma coercitiva en la actividad cerebral de los sujetos. En definitiva, su necesidad es garantizar de forma indispensable la autodeterminación moral, que es el fundamento último de la libertad de conciencia.

3. Acceso equitativo al aumento cognitivo

Finalmente, y, en tercer lugar, el derecho al acceso equitativo al aumento cognitivo evita desigualdades estructurales en la capacidad de deliberar y ejercer la autonomía. Si bien estas tecnologías podrían acelerar el aprendizaje, favorecer la innovación y mejorar el bienestar general, también acarrear riesgos significativos para

³⁹ *Ídem*.

la justicia social. En este caso, su relación con la libertad de conciencia se encuentra en un nivel diferente; mientras los dos derechos anteriores protegen el núcleo interno de la mente frente a amenazas, este derecho garantiza las condiciones materiales y sociales necesarias para el ejercicio pleno de la autonomía cognitiva. De ahí que, una distribución desigual de estas mejoras podría derivar en brechas cognitivas sin precedentes, generando sociedades estratificadas entre quienes tienen acceso a capacidades ampliadas y quienes quedan excluidos, así la libertad de conciencia dejaría de ser un derecho ejercido en condiciones de igualdad, pues algunos individuos tendrían un mayor control sobre su mente y procesos deliberativos simplemente por su acceso privilegiado a mejoras cognitivas.

Conclusiones

La convergencia entre la IA y la neurotecnología proyecta riesgos sobre el *forum internum*, núcleo intangible de la libertad de conciencia. La capacidad de inferencia, perfilado y eventual modulación de procesos cognitivos introduce supuestos que afectan directamente a la autonomía mental como presupuesto del libre desarrollo de la personalidad.

Los denominados neuroderechos pueden ser interpretados, desde una perspectiva dogmática, no como un catálogo autónomo, sino como concreción evolutiva de la dignidad humana y de la libertad de pensamiento en el entorno digital. La privacidad mental, la integridad cognitiva y la autodeterminación frente a intervenciones neuronales constituyen manifestaciones contemporáneas de la tutela del *forum internum* reconocida en el Derecho internacional. En este contexto, la aportación antropológica y bioética de la Doctrina Social de la Iglesia ofrece un criterio hermenéutico relevante para la interpretación y desarrollo de estas garantías. La centralidad de la dignidad humana —reafirmada por la tradición jurídica y por la reflexión antropológica cristiana— opera como criterio imprescindible. La protección reforzada de la autonomía mental debe configurarse como principio estructural del constitucionalismo contemporáneo, garantizando que la innovación tecnológica permanezca jurídicamente subordinada a la inviolabilidad de la conciencia y a la primacía de la persona.