

El impacto de la inteligencia artificial en el Derecho Eclesiástico del Estado: una nueva línea de investigación

The impact of artificial intelligence on State Ecclesiastical Law: a new line of research

José M.^a Contreras Mazarío

Catedrático de Derecho eclesiástico del Estado

Universidad Pablo de Olavide de Sevilla

jmconmaz@upo.es

https://doi.org/10.55104/LYL_00004

Fecha de recepción: 29/06/2025

Fecha de aceptación: 29/07/2025

RESUMEN

Las nuevas tecnologías no sólo planean nuevos interrogantes y preguntas, y con ello nuevos contenidos, sino también nuevos enfoques, que afectan, y sobre todo afectarán, a un importante número de materias, así como formas de hacer (o incluso de actuar), toda vez que obligan a pensar y replantear nuestra posición en el mundo y en la vida. Por lo que a nuestra área de conocimiento se refiere, cuestiones tales como la libre formación de la conciencia y el ámbito ad intra de la libertad de pensamiento y de conciencia pueden verse afectadas, y de hecho ya lo están siendo, por los avances de las nuevas tecnologías en general, y de la inteligencia artificial en particular. En este trabajo se va a hacer referencia a algunos de esos contenidos y cómo los mismos pueden verse afectados de tal manera que nos lleven a replantearnos algunas de las respuestas ahora dadas, como puede ser la respuesta clásica al ámbito «ad intra» de la libertad de conciencia.

PALABRAS CLAVE

Inteligencia artificial, libertad de conciencia, Derecho eclesiástico del Estado.

ABSTRACT

New technologies not only raise new questions and issues, and thus new content, but also new approaches, which affect, and above all will affect, a large number of

topics, as well as ways of doing (or even acting), as they force us to think and rethink our position in the world and in life. As far as our area of knowledge is concerned, issues such as the free formation of conscience and the *ad intra* scope of freedom of thought and conscience may be affected, and indeed are already being affected, by developments in new technologies in general, and artificial intelligence in particular. This paper will refer to some of these contents and how they may be affected in such a way as to lead us to rethink some of the answers currently given, such as the classical response to the '*ad intra*' scope of freedom of conscience.

KEYWORDS

Artificial intelligence, freedom of conscience, State ecclesiastical law.

SUMARIO: 1. Consideraciones generales. 2. Derecho e inteligencia artificial. 2.1 Inteligencia artificial y *soft law*. 2.2 Inteligencia artificial, proceso de codificación y regulación. 3. Libertad de pensamiento, conciencia y religión y nuevas tecnologías: un nuevo ámbito para la investigación. 3.1 Neurociencia y libertad de conciencia. 3.2 Nuevas tecnologías y libre formación de la conciencia. 3.3 Inteligencia artificial y derechos a la identidad, intimidad y privacidad. 3.4 Inteligencia artificial y ámbito de la salud: especial incidencia en la autonomía del paciente. 3.5. Inteligencia artificial y libertad de culto. 3.6 Libertad de culto y metaverso. 3.7. Inteligencia artificial y trascendencia. 4. Inteligencia artificial y principios del modelo de relación Estado español-fenómeno religioso. 4.1. Inteligencia artificial y principio de igualdad y no discriminación (el sesgo algorítmico). 4.2 Inteligencia artificial y pluralismo. 4.3 Inteligencia artificial y laicidad. 5. Consideraciones finales.

«Sólo la inteligencia se examina a sí misma»

Jaime Balmes

1. CONSIDERACIONES GENERALES

La lectura de una reciente monografía del profesor Rodríguez (2024)¹ y mi participación en dos seminarios sobre inteligencia artificial (2024 y 2025)², me llevan a suscitar –en este 25 aniversario de la revista

¹ RODRÍGUEZ GARCÍA, J. A., *Laicidad e inteligencia artificial. La implementación y el alineamiento de la laicidad como valor en los sistemas de inteligencia artificial*, Aranzadi, Pamplona, 2024.

² Seminario sobre los retos normativos de la inteligencia artificial, organizado por el Departamento de Filosofía del Derecho, Derecho Internacional Público y Derecho Eclesiástico del Estado de la Universidad Carlos III de Madrid, el día 4 de julio de 2024; y Seminario internacional sobre *Las implicaciones de la inteligencia artificial en la educación jurídica y la justicia digital*, organizado por la Universidad Nacional Autónoma de México, los días 14 y 15 de enero de 2025.

Laicidad y libertades: Escritos jurídicos— una nueva línea de investigación, objeto material de análisis por los eclesiasticistas. Esta línea nueva versaría sobre la libertad de conciencia y religiosa y las nuevas tecnologías en general, y la inteligencia artificial en particular. Aunque como ya he puesto de manifiesto, no soy —ni pretendo serlo ni asumirlo— quien haya abordado por primera vez esta temática, ya que ello ha correspondido —creo— al profesor Rodríguez García³, con esta aportación sólo pretendo plantear algunas cuestiones o materias de futuras investigaciones para nuestra área o ámbito de conocimiento conectadas con la inteligencia artificial y otras tecnologías disruptivas.

Estas nuevas tecnologías plantean no sólo nuevos interrogantes y preguntas, y con ello nuevos contenidos, sino también nuevos enfoques, que afectan, y sobre todo afectarán, a un importante número de materias, así como formas de hacer (o incluso de actuar), toda vez que obligan a pensar y replantear —por lo que a nuestro ámbito de conocimiento se refiere— cuestiones tales como, por ejemplo, la libre formación de la conciencia y el ámbito *ad intra* de la libertad de pensamiento y de conciencia, entre otras.

2. DERECHO E INTELIGENCIA ARTIFICIAL

Esta nueva realidad que es la inteligencia artificial está teniendo su reflejo y reflexión en múltiples ámbitos tecnológicos y de las ciencias sociales, y respecto del ámbito jurídico desde el plano del legislador tanto nacional como internacional⁴, sin olvidar los planos procesal-judicial, penal-criminal, privado-comercial, laboral y administrativo. Sin olvidar que la aplicación de los procesos computacionales cambiará la práctica del Derecho⁵, al conectar modelos computacionales de razonamiento jurídico directamente con el texto jurídico, generando argumentos a favor y en contra de

³ Su primera publicación en esta materia, junto con MORENO REBATO, M., data de 2018, «¡El futuro ya está aquí! Derecho e inteligencia artificial», en *Revista Aranzadi de Derecho y nuevas tecnologías*, núm. 48.

⁴ Ver LEE, J., *Artificial Intelligence and International Law*, Springer, Londres; MUÑOZ VELA, J. M. (2024): *La regulación de la inteligencia artificial. Reto y oportunidad desde una perspectiva global e internacional*, Aranzadi, Pamplona, 2023.

⁵ Ver ASHLEY, K. D., *Inteligencia artificial y analítica jurídica: Nuevas herramientas para la práctica del derecho en la era digital*, YachayLegal, 2024.

determinados resultados, prediciendo resultados y explicando estas predicciones con razones que los profesionales del Derecho podrán evaluar por sí mismos. Estas aplicaciones jurídicas apoyarán la recuperación de información jurídica conceptual y permitirán la computación cognitiva, posibilitando una colaboración entre humanos y ordenadores en la que cada uno realice el tipo de actividades inteligentes que mejor sepa hacer.

2.1 Inteligencia artificial y soft law

Con relación a este segundo nivel y ámbito, se debe hacer referencia a toda la labor de «soft law»⁶ llevada a cabo por distintas organizaciones internacionales, por lo que me limitaré en este momento con hacer referencia a algunas de ellas⁷, entre otras y a modo de ejemplo: a la Declaración de Helsinki sobre *Principios éticos para la investigación médica con seres humanos* (1964)⁸, al Informe Belmont sobre *Principios Éticos y Directrices para la Protección de Sujetos Humanos de Investigación* (1979)⁹, a la Declaración Universal de la UNESCO sobre *Bioética y Derechos Humanos* (2006)¹⁰, a la Declaración de Asilomar sobre *Principios de Precaución sobre Inteligencia Artificial* (2017)¹¹, a la Carta de Derechos digitales elaborada por la

⁶ Véase, SOSA NAVARRO, M., «El papel de la soft law en la gobernanza de la inteligencia artificial», en VV. AA.: *Los retos normativos de la inteligencia artificial*, Dykinson, Madrid, 2024, pp. 129-156.

⁷ Ver Consejo de Derechos Humanos de las Naciones Unidas, a través del Comité Asesor que preparara un estudio exhaustivo sobre el tema: «*Neurotecnología y derechos humanos*». Doc. de la ONU A/HRC/51/L.3', 2022 (en internet: <https://www.ohchr.org/sites/default/files/documents/hrbodies/hrcouncil/advisorycommittee/sessions/session31/a-hrc-ac-31-crp-1.docx>; consultado: 21 de marzo de 2024).

También véanse Resoluciones de la AG sobre el derecho a la privacidad en la era digital: 75/176, diciembre de 2020; 73/179, diciembre de 2018; 71/199, diciembre de 2016; 69/166, diciembre de 2014, y 68/167, diciembre de 2013; y Resoluciones del Consejo de Derechos Humanos: 42/15, septiembre de 2019; 37/2, marzo de 2018; 34/7, marzo de 2017, y 28/16, marzo de 2015.

⁸ En internet: <https://www.wma.net/es/policies-post/declaracion-de-helsinki-de-la-mm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/> (consultado: 31 de enero de 2025).

⁹ En internet: <https://www.hhs.gov/sites/default/files/informe-belmont-spanish.pdf> (consultado: 4 de febrero de 2025)

¹⁰ En internet: https://unesdoc.unesco.org/ark:/48223/pf0000146180_spa (consultado: 4 de febrero de 2025).

¹¹ En internet: <https://futureoflife.org/open-letter/ai-principles/> (consultado: 10 de febrero de 2025).

ECRI (2022)¹², a la *Recomendación del Consejo de la OCDE sobre la Innovación Responsable en Neurotecnología*, (Doc. OCDE/LEGAL/0457) (2019)¹³, al *Plan de Acción Estratégico del Comité de Bioética del Consejo de Europa sobre Derechos Humanos y Tecnologías en Biomedicina (2020-2025)* (2019)¹⁴, al *Informe del Comité Internacional de Bioética de la UNESCO sobre las cuestiones éticas de la neurotecnología* (Doc. de la UNESCO SHS/BIO/IBC-28/2021/3/Rev.) (2021)¹⁵, a la *Declaración del Comité Jurídico Interamericano sobre Neurociencia, Neurotecnologías y Derechos Humanos: nuevos desafíos jurídicos para las Américas* (2021)¹⁶, a la *Declaración del Comité Jurídico Interamericano de la Organización de los Estados Americanos sobre Neurociencia, Neurotecnologías y Derechos Humanos: Nuevos Desafíos Jurídicos para las Américas* (2021)¹⁷, a la *Resolución del Parlamento Europeo sobre la Inteligencia Artificial en la era digital* (2022)¹⁸, y a la *Declaración sobre los derechos y principios digitales europeos* (2022)¹⁹.

Ahora bien, este *solf law*, que asienta los principios básicos del desarrollo tecnológico, puede ser entendido como una «huida» hacia delante y, en concreto, a un régimen de normas y estándares no vinculantes, lo que suscita serias preocupaciones; más aún, cuando los organismos gubernamentales y las instituciones internacionales parecen rezagados y sin una respuesta homogénea. La falta de un marco legal vinculante y coherente a nivel internacional podrá llevar a una fragmentación y a un enfoque de «quien tiene el dinero hace las reglas».

¹² En internet: https://edri.org/files/WePromiseCharter_booklet_ES.pdf (consultado: 15 de febrero de 2025).

¹³ Internet: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0457> (consultado: 11 de febrero de 2025)

¹⁴ En internet: <https://rm.coe.int/native/0900001680994df7> (consultado: 21/02/2025).

¹⁵ En internet: https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa (consultado: 21 de febrero de 2025).

¹⁶ En internet: <https://www.saij.gob.ar/0-internacional-declaracion-comite-juridico-interamericano-sobre-neurociencia-neurotecnologias-derechos-humanos-lnt0007366-2021-08-11/123456789-0abc-defg-g66-37000tcanyel?&o=19&f=Total%7CTipo%20de%20Documento/Legislaci%F3n%7CFecha/2021/08%7COrganismo%7CPublicaci%F3n%7CTema/Derecho%20internacional%7CEstado%20de%20Vigencia%7CAutor%7CJurisdicci%F3n&t=25> (consultado: 31 de enero de 2025).

¹⁷ En internet: https://www.oas.org/es/sla/cji/docs/CJI-DEC_01_XCIX-O-21.pdf (consultado: 21 de febrero de 2025).

¹⁸ En internet: <https://op.europa.eu/es/publication-detail/-/publication/8f622861-7533-11ed-9887-01aa75ed71a1/language-es> (consultado: 21 de febrero de 2025).

¹⁹ En internet: <https://digital-strategy.ec.europa.eu/es/library/declaration-european-digital-rights-and-principles> (consultado: 23 de febrero de 2025).

2.2 Inteligencia artificial, proceso de codificación y regulación

1. Aunque a nivel estatal aún resulta escaso el desarrollo normativo llevado a cabo, es destacable, por un lado, el cambio constitucional llevado a cabo en Chile a finales de 2021 (Ley núm. 21.38320), así como el realizado en la Constitución mexicana (2024) con especial incidencia en su artículo 6²¹, y, por otro, la Ley española de Protección de Datos y garantía de los derechos digitales (Ley 3/201822)²³ o la Ley peruana de Protección de Datos y su reglamento de desarrollo²⁴.
2. Mientras que a nivel internacional, debemos mencionar de manera especial la actividad llevada a cabo en el seno de la Unión Europea²⁵, donde se han desarrollado un conjunto de normas jurídicas importantes, como son, por un lado, la Ley europea de Inteligencia artificial (2024)²⁶ y, por otro, los Reglamentos rela-

²⁰ En internet: <https://www.bcn.cl/leychile/navegar?idNorma-1166983> (consultado: 23 de febrero de 2025).

²¹ En internet: <https://www.diputados.gob.mx/LeyesBiblio/pdf/CPEUM.pdf> (consultado: 19 de marzo de 2025).

²² Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos y garantía de los derechos digitales (en internet: <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>; consultado: 23 de febrero de 2025).

²³ En España, cabe destacar la Carta de Derechos Digitales, de 14 de julio de 2021, que regula los derechos digitales en el empleo de las neurotecnologías (en internet: https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion_publica/audiencia/ficheros/SEDIACartaDerechosDigitales.pdf; consultado: 21 de febrero de 2025). Aquí interesa su artículo 26.1.c), en el que ordena: «Asegurar la confidencialidad y seguridad de los datos obtenidos o relativos a sus procesos cerebrales y el pleno dominio y disposición sobre los mismos». No obstante, y aunque sirve como referencia y exhorta al legislador a su regulación ante los retos de la era digital, se trata de una declaración de principios y, por tanto, no vinculante.

²⁴ En internet: <https://www.gob.pe/institucion/congreso-de-la-republica/normas-legales/243470-29733> (consultado: 19 de marzo de 2025).

²⁵ Ver, sobre la actividad normativa de la UE, FIRION, A.-S., «El marco jurídico de la inteligencia artificial en la Unión Europea», en VV. AA.: *Los retos normativos de la inteligencia artificial*, Dykinson, Madrid, 2024, pp. 157-192.

²⁶ En internet: <https://artificialintelligenceact.eu/es/ai-act-explorer/> (consultado: 23 de febrero de 2025). Sobre esta norma, véanse BARRIO ANDRÉS, M. (dir.), *Comentarios al Reglamento Europeo de Inteligencia Artificial*, La Ley, Madrid, 2024; HERNÁNDEZ LÓPEZ, J. M., *Reglamento de inteligencia artificial: incluye introducción, notas, cronología, webgrafía, bibliografía e índice analítico*, Bosch Editor, Madrid, 2024; LÓPEZ-AMO SAINZ, Á. P., *La Ley de la Inteligencia artificial*, parte I, Ra-Ma, 2024; MARTÍN JIMÉNEZ, F. J., *El Reglamento Europeo de Inteligencia Artificial*, 2024; Simón Castellano, P., y Cotino Hueso, L. (dirs.), *Tratado sobre el Reglamento Europeo de Inteligencia Artificial*, Aranzadi, Pamplona, 2024; VV. AA., *Inteligencia artificial y derecho de daños: cuestiones actuales: acordes al Reglamento (UE) de 2024*, Dykinson, Madrid, 2024.

cionados con los Servicios Digitales (2022)²⁷, con los Datos Personales (2016)²⁸ y con los Mercados únicos de Servicios Digitales (2022)²⁹, así como con la Gobernanza de Datos (2022)³⁰, a los que se puede añadir la Directiva relativa a los Datos Abiertos (2019)³¹.

3. Por su parte, y con relación a la neurotecnología y la protección de datos, resultan reseñables dos sentencias pioneras: una en Chile: sentencia de la Corte Suprema de 9 de agosto de 2023³², y otra en España: sentencia del Tribunal Supremo de 5 de mayo de 2020³³. Mientras que, respecto al derecho al olvido en internet, cuatro sentencias del TJUE resultan relevantes: caso *Costeja*, asunto C-131/12 (2014)³⁴, caso *Digital Rights Ireland y Seitlinger y otros*, asunto C-293-12 (2024)³⁵; caso *Fashion ID GmbH & Co. KG contra Verbraucherzentrale NRW eV*, asunto C-40/17 (2019)³⁶, y caso *GC, AF, BH, ED*, asunto C136/17 (2019)³⁷. Y, por último, con relación a la protección de la desinformación en la era digital, el TEDH ha abordado la presente cuestión en las sentencias de 15

²⁷ En internet: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32022R2065> (consultado: 23 de febrero de 2025).

²⁸ 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 datos (en internet: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32016R0679>; consultado: 23 de febrero de 2025).

²⁹ En internet: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-81573> (consultado: 23 de febrero de 2025).

³⁰ En internet: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32022R0868> (consultado: 26 de febrero de 2025).

³¹ En internet: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32019L1024> (consultado: 28 de febrero de 2025).

³² En internet: <https://www.diarioconstitucional.cl/wp-content/uploads/2023/08/GIRAR-DICONEMOTIVSUPREMA.pdf>105.065-2023.pdf (consultado: 28 de febrero de 2025).

³³ En internet: <https://www.poderjudicial.es/search/AN/openCDocument/cac2ec927d-f2ac2410b129baa45c19bf7ef1be27ecf5f288> (consultado: 16 de febrero de 2025).

³⁴ En internet: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:62012CA0131> (consultado: 21 de febrero de 2025).

³⁵ En internet: <https://curia.europa.eu/juris/document/document.jsf?docid=150642&doclang=ES> (consultado: 21 de febrero de 2025).

³⁶ En internet: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:62017CJ0040> (consultado: 22 de enero de 2025).

³⁷ En internet: <https://curia.europa.eu/juris/document/document.jsf?sessionId=609A35040F380D-494B1B42B8EE5F47C5?text=&docid=218106&pageIndex=0&doclang=ES&mode=req&dir=&occ=first&part=1&cid=1763576> (consultado: 23 febrero de 2025).

de mayo de 2023, caso *Sánchez c. Francia*³⁸ y de 16 de julio de 2013, caso *Wergzynowski y Smolczewski v. Polonia*³⁹.

Respecto a España y el derecho al olvido⁴⁰, resultan de interés mencionar, por un lado, la sentencia del TC 58/2018, de 4 de junio⁴¹; y, por otro, las sentencias del TS 545/2015, de 15 de octubre⁴²; 191/2016, de 8 de marzo⁴³; 210/2016, de 5 de abril⁴⁴; y 12/2019, de 11 de enero⁴⁵.

4. Todo ello pone de manifiesto el momento en el que nos encontramos, y que puede calificarse de inicio o frontispicio. A la escasez de normativa, se une asimismo la escasez jurisprudencial e, incluso, podría decirse que de doctrina. Aunque en los últimos años se han publicado un importante número de trabajos jurídicos relacionados con esta materia⁴⁶ y

³⁸ En internet: <https://www.bailii.org/eu/cases/ECHR/2023/418.html> (consultado: 23 de febrero de 2025).

³⁹ En internet: [https://www.europarl.europa.eu/RegData/etudes/STUD/2018/628261/EPRS_STU\(2018\)628261_ES.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2018/628261/EPRS_STU(2018)628261_ES.pdf) (consultado: 25 de febrero de 2025).

⁴⁰ Véanse, BERROCAL LANZAROT, A. I., *Derecho de supresión de datos o derecho al olvido*, Ed. Reus, Madrid, 2017; DI PIZZO CHIACCHIO, A., *La expansión del derecho al olvido digital. Efectos de «Google Spain» y el Big Data e implicaciones del nuevo Reglamento Europeo de Protección de Datos*, Atelier, Barcelona, 2018; PAZOS CASTRO, R., «El derecho al olvido frente a los editores de hemerotecas digitales», en *Indret*, núm. 4, 2016; SIMÓN CATELLANO, P., *El régimen constitucional del derecho al olvido digital*, Tirant lo Blanch, Valencia, 2011.

⁴¹ En internet: <https://hj.tribunalconstitucional.es/es-ES/Resolucion/Show/25675> (consultado: 27 de febrero de 2025). Comentarios de la sentencia, BROTONS MOLINA, O., «Caso Google: Tratamiento de datos y derecho al olvido. Análisis de las conclusiones del abogado general, asunto C-131712», en *Revista Aranzadi de Derecho y Nuevas Tecnologías*, núm. 33.

⁴² En internet: <https://vlex.es/vid/585072934> (consultado: 27 de febrero de 2025). Ver comentario a la misma, en internet: https://www.boe.es/biblioteca_juridica/comentarios_sentencias_unificacion_doctrina_civil_y_mercantil/abrir_pdf.php?id=COM-D-2015-35 (consultado: 27 de febrero de 2025).

⁴³ En internet: <https://vlex.es/vid/631448266> (consultado: 26 de febrero de 2015).

⁴⁴ En internet: <https://vlex.es/vid/632311849> (consultado: 25 de febrero de 2025). Ver comentario a la misma, en internet: https://www.boe.es/biblioteca_juridica/comentarios_sentencias_unificacion_doctrina_civil_y_mercantil/abrir_pdf.php?id=COM-D-2016-26 (consultado: 25 de febrero de 2025).

⁴⁵ En internet: <https://vlex.es/vid/757238617> (consultado: 24 de febrero de 2025). Ver comentario a la misma, en internet: <https://dialnet.unirioja.es/servlet/articulo?codigo=7026841> (consultado: 27 de febrero de 2025).

⁴⁶ Véanse, entre otros, CALAZA LÓPEZ, S., y ORDEÑANA GEZURAGA, I., *Next Generation Justice: Digitalización e Inteligencia Artificial*, Aranzadi, Pamplona, 2024; CARLÓN RUIZ, M., *Las administraciones públicas ante la inteligencia artificial*, Tirant lo Blanch, Valencia, 2025; CASTILLA BAREA, M., y CERVILLA GARZÓN, M.ª D. (dirs.), *El Derecho y la Justicia ante la inteligencia artificial y otras tecnologías disruptivas*, Aranzadi, Pamplona, 2024; GARCÍA MEXÍA, P., *Claves de inteligencia artificial y Derecho*, Aranzadi, Pamplona, 2022; HERBOSA MARTÍNEZ, I., y FERNÁNDEZ DE RETANA GOROSTIZAGOIZA, D.,

se están llevando a cabo un significativo número de proyectos de investigación, las incertidumbres y sobre todo la escasez de certezas siguen presentes. Al tiempo que toda la actividad desarrollada da muestras del interés en esta materia, pero también de la preocupación que estas nuevas tecnologías suscitan sobre todo desde el plano de la ética⁴⁷ y de los derechos humanos y las libertades fundamentales⁴⁸.

3. LIBERTAD DE PENSAMIENTO, CONCIENCIA Y RELIGIÓN Y NUEVAS TECNOLOGÍAS: UN NUEVO ÁMBITO PARA LA INVESTIGACIÓN

Desde la perspectiva de los contenidos de nuestra disciplina, las nuevas tecnologías en general, y la inteligencia artificial en particular, ya inciden –como se ha señalado– en temáticas como el ámbito *ad intra* de la libertad de pensamiento y de conciencia, en la libre formación de la conciencia de las personas o/y en la libertad de culto; pero también incidirá sin duda en los principios que conforman el sistema de relación

Derecho e inteligencia artificial, Aranzadi, Pamplona, 2023; NIEVA FENOLL, J., *Inteligencia artificial y proceso judicial (Proceso y Derecho)*, Marcial Pons, Madrid, 2018; Ramón Fernández, F. (coord.), *Ciencias de datos y perspectivas de la inteligencia artificial*, Tirant lo Blanch, Valencia, 2024; Rocha Espíndola, M. A.; Sansó-Rubert Pascual, D., y Rodríguez Dos Santos, N. (coords.), *Inteligencia artificial y derecho: Reflexiones jurídicas para el debate sobre su desarrollo y aplicación*, Dykinson, Madrid, 2023; SUÁREZ-MUÑOZ, F., *Internet de las cosas e Inteligencia artificial. Los retos regulatorios y éticos del extractivismo de datos, la privacidad y los derechos humanos*, IoT Ciberseguridad, 2024; Vallespín Pérez, D. (coord.), *Inteligencia artificial y proceso*, Juruá ed., 2023; VV. AA., *Desafíos jurídicos de la inteligencia artificial: reflexiones desde la toma de decisiones judiciales*, Bosch editor, Barcelona; VV. AA., *El derecho y la inteligencia artificial*, Ed. Universidad de Granada, Granada, 2022; VV. AA., *Inteligencia artificial aplicada al ámbito jurídico*, Sepín, Madrid, 2024.

⁴⁷ Ver, a este respecto, Azuaje Pirela, M. (coord.), *Introducción a la ética y el derecho de la inteligencia artificial*, Aranzadi, Pamplona, 2023; BELLOSO MARTÍN, N., «Cómo la ética complementa la normatividad de la inteligencia artificial: códigos de conducta, guías de buenas prácticas y “sello” de IA confiable», en VV. AA., *Los retos normativos de la inteligencia artificial*, Ed. Dykinson, Madrid, pp. 15-72; CANABAL, A.: *Origen y desarrollo de la Neuroética: 2002-2012*, 2013 (en internet: <https://sciendo.es/pdf/bioetica/n28/articulo4.pdf>; consultado: 3 de febrero de 2024); CORTINA, A., *Neuroética y neuropolítica. Sugerencias para la educación moral*, Editorial Tecnos, Madrid 2014.

⁴⁸ Véanse, BARRIFFI, F. J., «Perspectivas sobre la inteligencia artificial y su impacto en el derecho y los derechos humanos», en VV. AA., *Los retos normativos de la inteligencia artificial*, Ed. Dykinson, Madrid, pp. 73-128; DE ASÍS, R., *Una mirada a la robótica desde los derechos humanos*, Dykinson, Madrid, 2015; GUTIÉRREZ GARCÍA, E., *Inteligencia artificial y derechos fundamentales: Hacia una convivencia en la era digital*, Colex, Madrid, 2024; PRESNO LINERA, M. Á., *Derechos fundamentales e inteligencia artificial (Debates constitucionales)*, Marcial Pons, Madrid, 2023; REBOLLO DELGADO, L., *Inteligencia artificial y derechos fundamentales*, Dykinson, Madrid, 2023.

Estado-fenómeno religioso como pueden ser el de laicidad, pluralismo e igualdad.

3.1 Neurociencia y libertad de conciencia

Los avances en la neurociencia tienen repercusiones importantes sobre el derecho y sobre la mente que lleva a plantear la necesidad de nuevas regulaciones que incorporen los llamados «neuroderechos»⁴⁹. La inteligencia artificial nos ayuda a gestionar todo tipo de decisiones y con una eficacia creciente. En el contexto del trabajo subordinado, se espera que el uso de asistentes informacionales alimentados por algoritmos predictivos crezca exponencialmente, describiendo la antesala de la sociedad del futuro. El empleo de esta tecnología, en contrapartida, está posibilitando una cartografía humana exhaustiva. Nuestras intenciones, emociones y estados de ánimo, pero también nuestros pensamientos e ideas (y con ellos, nuestras creencias o convicciones) ya pueden ser leídos. Este acceso al patio trasero neuronal describe un nuevo desafío porque estas máquinas están acumulando capacidad para «aguijonear» la mente (nuestra mente y con ello también nuestra conciencia), acceder al yo inconsciente y condicionar subliminalmente el comportamiento⁵⁰.

Partiendo de lo expuesto, cabe señalar que hasta ahora este ámbito interno ha supuesto un claustro para la autodeterminación y autonomía, vinculado a la propia personalidad y dignidad humana, entendidas éstas como la capacidad del hombre, como consecuencia de estar dotado de entendimiento y voluntad, de tomar por sí mismo decisio-

⁴⁹ Ver BASELGA-GARRIGA, C.; RODRIGUEZ, P., y YUSTE, R., «Neuro Rights: A Human Rights Solution to Ethical Issues of Neurotechnologies», en López-Silva, P., y Valera L. (eds): *Protecting the Mind. Ethics of Science and Technology Assessment*, vol. 49, Springer (2022), pp. 157 ss.; DE ASIS, R., «Sobre la propuesta de los neuroderechos», en *Derechos y libertades*, núm. 47 (2022), pp. 51-70; IENCA, M., y ANDORNO, R., «Hacia nuevos derechos humanos en la era de la neurociencia y neurotecnología», *Análisis Filosófico*, núm. 41, vol. 1, 2021; YUSTE, R.; GENSER, J., y HERRMANN, S., «It's Time for Neuro-Rights», en *Horizons*, Center for International Relations and Sustainable Development, 2021.

⁵⁰ Véanse BELTRAN DE HEREDIA RUIZ, I., *Inteligencia artificial y neuroderechos: la protección del yo inconsciente de la persona*, Aranzadi, Pamplona, 2023; MURRUGARRA RETAMOZO, B. I., *Neuroderechos, neurotecnologías e inteligencia artificial: protección de la actividad cerebral humana*, Colex, Madrid, 2024.

nes⁵¹. Se puede afirmar, por tanto, que la libertad de conciencia protege el proceso racional, reflexivo, la elaboración intelectual del ser humano y su adhesión o no a concepciones valóricas o conviccionales, sean estas religiosas o filosóficas.

En definitiva, el derecho de pensar con plena libertad, lo que posibilita la propia selección o determinación de valores de acuerdo con los cuales se formula cada proyecto de vida y la conformación a dicho pensamiento de la actividad externa personal y social, como asimismo a rechazar aquellas que se consideran erróneas; proceso que corresponde al fuero interno de la persona que tiene (o al menos, hasta ahora tenía) un carácter inviolable y absoluto⁵², convirtiéndolo en un derecho ilimitado por su propia naturaleza⁵³, ya que no cabe (o cabía) intervención directa en el pensamiento o fuero interno de las personas, y tampoco parecía que pudiera imaginarse interferencia alguna en esta dimensión, en tanto en cuanto no se exteriorizasen de algún modo los pensamientos, las ideas o las convicciones⁵⁴; al tiempo que supone igualmente la facultad de toda persona para formarse su propio juicio, sin ningún tipo de interferencias no queridas y/o consentidas⁵⁵.

Esta realidad, que parecía inmutable, ya no es –y parece que cada vez menos– la última frontera. Los avances científicos en general, y los neotecnológicos en particular, suponen y supondrán grandes beneficios

⁵¹ En este sentido, para VILADRICH, la libertad religiosa «*por objeto el conjunto de ideas, conceptos y juicios que el hombre tiene sobre las distintas realidades del mundo y de la vida, o lo que es lo mismo el derecho de toda persona a tener su propio sistema o concepción explicativa del hombre, el mundo y la vida; en definitiva, una personal y libre cosmovisión*» (VILADRICH, J. P., «Los principios informadores del Derecho eclesiástico español», en VV. AA.: *Derecho eclesiástico del Estado español*, 2.ª ed., EUNSA, Pamplona, 1983, p. 265).

⁵² En este sentido, véase LLAMAZARES, D., *Derecho de la libertad de conciencia. I. Conciencia, Tolerancia y Laicidad*, Cívitas/Thomson Reuters, Pamplona, 2011.

⁵³ Ver DÍAZ REVORIO, F. J., «La libertad de ideología y religión», en internet: https://parlamentoyconstitucion.cortesclm.es/recursos/articulos/PyC1_Diaz_Libertad.pdf (consultado: 9 de febrero de 2025).

⁵⁴ Ello lleva a algunos autores a rechazar el mencionado carácter de ilimitado, por considerar que esta dimensión carece de auténtico contenido normativo, es decir, no alcanza a configurar realidad jurídica alguna. Ver en este sentido, MORENO GARCÍA, A., «Significado constitucional de la libertad ideológica», en *Anuario de Derecho constitucional y parlamentario*, núm. 7, 1995, pp. 118-119.

⁵⁵ Ver SAGÜÉS, N. P., *Elementos de Derecho Constitucional*, tomo 2, 3.ª edición actualizada y ampliada, Ed. Astrea, Buenos Aires, p. 475.

para la humanidad (sobre todo en el campo de la medicina en enfermos como el Alzheimer o el Parkinson, entre otros), pero tampoco se pueden desconocer los riesgos que éstos suponen para los derechos humanos en general⁵⁶, y para la libertad de pensamiento y de conciencia en particular.

El derecho a la libertad de conciencia ofrece elementos que pueden ayudar a una protección y tutela efectiva, siempre que se lleve a cabo una reinterpretación o reconfiguración de este ámbito de tal manera que ese claustro interno de las creencias que da lugar a la autodeterminación intelectual de la persona (STC 177/1996⁵⁷) entre a formar parte del mundo jurídico como si de un ámbito externo se tratara.

Partiendo de esta consideración, nos atrevemos a proponer un nuevo alcance de la libertad de pensamiento y de conciencia en su dimensión interna, de tal modo que suponga, por lo que a los avances neurotecnológicos se refiere, un doble contenido: por un lado, la protección de las personas a fin de que éstas tomen sus decisiones de forma libre y autónoma, sin manipulación o influencia alguna procedente de las neurotecnologías, y sin forma alguna de discriminación, coerción o violencia; así como su aspecto negativo: esto es, la no obligación de declarar sobre sus ideas o creencias; y, por otro, la protección contra el acceso no consentido a los datos cerebrales, asegurando que los pensamientos, memorias y sentimientos de una persona permanezcan en el ámbito privado o íntimo si así se desea.

En consecuencia, las neurointervenciones sin consentimiento que, interfieran en un grado significativo con la formación, retención y revisión de pensamientos, ideas, creencias, convicciones, opiniones, recuerdos o la actividad mental de pensar (razonar, enjuiciar, analizar, decidir,

⁵⁶ El Consejo de Derechos Humanos (CDH) adoptó, el 6 de octubre de 2022, la Resolución 51/3 mediante la cual solicitó a su Comité Asesor (CA) que preparara un estudio «sobre el impacto, las oportunidades y los desafíos de la neurotecnología en relación con la promoción y protección de todos los derechos humanos». En febrero de 2023, el Comité de Auditoría estableció un grupo de redacción compuesto actualmente por Noor Al-Jehani, Nurah Alamro, Joseph Gérard Angoh, Buhm-Suk Baek, Milena Costas Trascasas (ponente), Riva Ganguly Das, Jewel Major, Javier Palumno, Vasilka Sancin, Patrycja Sasnal, Vassilis Tzevelekos (presidente), Catherine Van de Heyning, Frans Viljoen y Yue Zhang. Ver también, a este respecto, Doc. A/HRC/AC/31/CRP.1.

⁵⁷ En internet: <https://hj.tribunalconstitucional.es/es-ES/Resolucion/Show/3229> (consultado: 25 de febrero de 2025).

crear, imaginar y recordar, entre otras posibles) deberían representar una injerencia ilegal y, por ende, una quiebra del derecho a la libertad de pensamiento y de conciencia⁵⁸.

Dentro de este ámbito, también debe hacerse referencia a una última temática como es la relativa a la objeción de conciencia, sobre todo si entendemos que este tipo de prácticas (neuronales) pueden entenderse incluidas dentro del ámbito de la salud. Dicha inclusión supone –en el caso de España– la aplicación a estas técnicas de la *Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica*⁵⁹, y con ella a que el paciente o usuario tenga derecho a decidir libremente, después de recibir la información adecuada (art. 4 en relación con el art. 10), entre las opciones clínicas disponibles (art. 2.3), lo que supone la facultad de éste de optar, libre y voluntariamente, entre dos o más alternativas asistenciales, entre varios facultativos o entre centros asistenciales, en los términos y condiciones que establezcan los servicios de salud competentes, en cada caso, o bien la no aceptación o rechazo de todas las terapias médicas propuestas, incluso aunque ello pueda poner en peligro la vida del paciente⁶⁰. Amén de que a esta temática puede, igualmente, ser aplicable *mutatis mutandi* la *Ley 14/2007, de 3 de julio, de Investigación biomédica*⁶¹. En el caso de aplicación de procedimientos invasivos,

⁵⁸ Más en profundidad, ver CONTRERAS MAZARÍO, J. M.^a, «Los derechos de la conciencia, neurotecnologías e inteligencia artificial: ¿pueden jaquear nuestras conciencias?», en VV. AA.: *Los retos normativos de la inteligencia artificial*, Dykinson, Madrid, 2024, pp. 193-248.

⁵⁹ BOE núm. 274, de 15 de noviembre de 2002 (internet: BOE-A-2002-22188 Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica.; consultado: 20 de marzo de 2025).

⁶⁰ Véanse, a este respecto, SSTC 120/1990, de 27 de junio, FJ 7 (internet: <https://hj.tribunalconstitucional.es/es-ES/Resolucion/Show/1545>; consultado: 19 de marzo de 2025); 137/1990, de 19 de junio (internet: <https://hj.tribunalconstitucional.es/es-ES/Resolucion/Show/1562>; consultado: 18 de marzo de 2025), y 154/2002, de 18 de julio (internet: <https://www.boe.es/buscar/doc.php?id=BOE-T-2002-15992>; consultado: 20 de marzo de 2025).

⁶¹ BOE de 4 de julio de 2007 (internet: <https://www.boe.es/buscar/doc.php?id=BOE-A-2007-12945>; consultado: 20/03/2025).

En la presente norma legal se recoge que el valor fundamental a salvaguardar es la dignidad e identidad del ser humano, el respeto de su integridad y de todos sus derechos fundamentales (art. 1 y 2.a). Reaparecen, asimismo, de nuevo aquí los principios del consentimiento debidamente informado (art. 4), la confidencialidad en el tratamiento de los datos de carácter personal (art. 5), la no discriminación (art. 6) y la gratuidad (art. 7).

que entrañen un grave riesgo para la salud en general, y para la salud mental en el caso que nos ocupa en particular, de los pacientes, los principios que se deben respetar son los siguientes: i) inexistencia de alternativa de eficacia comparable, y ii) proporcionalidad entre riesgos y beneficios potenciales (art. 14).

Partiendo de ello, cabe precisar que en los supuestos de objeción de conciencia se debería distinguir si el paciente es adulto o si, por el contrario, se trata de un menor de edad o incapacitado. En el supuesto de los menores o incapacitados se deberá tener presente el derecho a la patria potestad del menor (art. 18 CDN), así como el previo consentimiento de sus padres o tutores para la realización de estas prácticas⁶², pero en todo momento se deberá proteger su derecho a la intimidad e identidad (art. 16 CDN), pero además debería entrar en acción la cláusula relativa al interés superior del niño⁶³, contenida en la Convención de Derechos del Niño (arts. 19 y 24 CDN). Mientras que en el caso de los adultos bastará con la aplicación con especial acento del consentimiento informado (art. 4 Ley 14/2007). En consecuencia, negar la existencia de esa esfera de decisiones *ad intra* personalísimas sería –como señala Llamazares– «tanto como negar a la persona su derecho más radical a ser y sentirse íntimamente libre y dueño de sí mismo, de lo que considera más suyo: la capacidad de decisión sobre sí mismo»⁶⁴.

3.2 Nuevas tecnologías y libre formación de la conciencia

Un segundo plano de incidencia lo encontramos en el derecho a la libre formación de la conciencia, y con él en su proyección en otras dos mate-

⁶² En estos casos se exige con especial acento el consentimiento expreso, específico y escrito del paciente o de sus representantes legales (art. 13 Ley 14/2007).

⁶³ Sobre el interés superior del niño, ver BO JANÉ, M., y CABALLERO, M., «El nuevo derecho del menor a ser oído: ¿sujeto activo en la determinación de su interés?», en *Revista jurídica La Ley*, núm. 4166, 1996; BORRÁS, A., «El interés del menor como factor de progreso y unificación del Derecho internacional privado, en su discurso de ingreso en la Academia de Jurisprudencia y Legislación de Cataluña», en *Revista Jurídica de Cataluña*, núm. 4, 1994; RAVETLLAT BALLESTÉ, I., «El interés superior del niño: concepto y delimitación del término», en *Educatio Siglo XXI*, vol. 30, núm. 2, 2012, pp. 89-108.

⁶⁴ LLAMAZARES, D., *Derecho de la libertad de conciencia. Vol. II. Conciencia, identidad personal y solidaridad*, Cívitas/Thomson Reuters, Navarra 201, p. 358.

rias como son, por un lado, la libertad de información y de expresión y, por otro, la libertad de enseñanza y el derecho a la educación.

3.2.1 Nuevas tecnologías y libertad de expresión

Por lo que a los derechos de libertad de expresión y de información respecta, las nuevas tecnologías pueden afectar igual y directamente a la formación, el mantenimiento y la revisión de todo tipo de pensamientos⁶⁵ y, por lo tanto, tiene el potencial de interferir simultáneamente con la libertad de opinión⁶⁶. A este respecto, los medios de comunicación y las redes sociales se han convertido –como afirma Rodríguez García– «gracias al big data y a la inteligencia artificial en instrumentos que influyen decisivamente, y cada vez más, en la formación de nuestra conciencia»⁶⁷, y por ende, añadiría yo, para el libre desarrollo de la personalidad (art. 10.1 CE en conexión con el art. 27.2 CE). En la actualidad, las redes sociales (como Twitter, Instagram, Facebook, TikTok, etc.) se han convertido en poderosos medios de comunicación en línea donde se procede a la recopilación de una inmensidad de datos personales que son ordenados gracias a algoritmos que utilizan la inteligencia artificial creando «perfiles personalizados»⁶⁸; y con ellos la posibilidad de que se produzca una manipulación de la conciencia.

⁶⁵ Según el Consejo de Derechos Humanos, la libertad de pensamiento es un derecho de gran alcance que abarca todas las materias; CCPR/C/GC34, párr. 9.

⁶⁶ Doc. A/73/348, párr. 26. Véase, BUBLITZ, C., «Neurotechnologies and human rights: restating and reaffirming the multi-layered protection of the person», en *The International Journal of Human Rights*, 2024 (en internet: Neurotecnologías y derechos humanos: reafirmar y reafirmar la protección multifacética de la persona (tandfonline.com); consultado el 19 de febrero de 2025).

⁶⁷ RODRÍGUEZ GARCÍA, J. A., «Libre formación de la conciencia, redes sociales y medios de comunicación: inteligencia artificial y democracia», en VV. AA.: *El Derecho eclesiástico del Estado en homenaje al Profesor Dr. Gustavo Suárez Pertierra*, Ed. Tirant lo Blanch, Valencia, 2021, p. 1157.

⁶⁸ El 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 (DOUE núm. 119, de 4 de marzo de 2016) (en internet: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2016-80807>; consultado el 27 de febrero de 2025), define la «elaboración de perfiles» como toda forma de tratamiento automatizado de datos personales de una persona consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comporta-

De igual modo, la inteligencia artificial se nutre del acervo producido por la inteligencia humana, y entre sus posibles consecuencias indeseables figura un plausible incremento de la desinformación y de los discursos de odio que ya pululan en internet, junto a una grandiosa capacidad de rastreo de datos. Entre las manifestaciones de esa manipulación puede encontrarse la censura previa de contenidos y las denominadas «fake news» y «deep fake»⁶⁹, toda vez que las redes sociales pueden seleccionar la información y las opiniones que circulan por las mismas. En este sentido, se ha señalado por Vermeulen⁷⁰, que «la libertad de pensamiento tiene tres elementos claves: el derecho a no revelar nuestros propios pensamientos (art. 16.2 CE), el derecho a no ser penalizados por nuestros pensamientos y el derecho a no ser manipulados». La manipulación de las conciencias significa una clara vulneración de la libertad de pensamiento y de conciencia, y los poderes públicos deben proteger su libre formación.

En este sentido, los motores de búsqueda, que actúan para buscar, recibir o difundir información, seleccionan ideas o creencias que se difunden a través de las redes sociales en función de los previsibles deseos de los usuarios; y un algoritmo de búsqueda puede estar dirigido hacia ciertos tipos de contenidos o proveedores de contenido, lo que puede poner en riesgo valores como el pluralismo informativo. Además, el uso de determinados algoritmos conduce a la fragmentación de lo que recibimos a través de la creación de «cámaras de eco», lo que supone que las redes sociales solo favorezcan a ciertos tipos de medios de comunicación, y con ello puede dar lugar –por ejemplo– al aumento de la polarización social, y con ello a la propia cohesión social.

miento, ubicación o movimientos de dicha persona física». Sobre el riesgo que supone para los derechos humanos la elaboración de perfiles, ver O'CARROLL, J., y FRANCO, J., «Muslim registries», en *Big Data and Human Rights*, Amnistía Internacional, 2017.

⁶⁹ Según la AI Act, «deep fake» «significa contenido de imagen, audio o video generado o manipulado por IA que se asemeja a personas, objetos, lugares u otras entidades o eventos existentes y que a una persona le parecería falsamente auténtico o veraz» (art. 3.44 ter.bl).

⁷⁰ VERMEULEN, B., «Freedom of thought, conscience and religion (article 9)», en *Theory and Practice of the European Convention of Human Rights*, Intersentia Press, Cambridge, 2006, pp. 751-772.

Por consiguiente, los algoritmos utilizados en las redes sociales no pueden responder a lo que se ha denominado «algoritmos opacos»⁷¹, por lo que deberán tenerse presente en el diseño de dichos algoritmos los «valores incrustados»⁷², ya que éstos pueden terminar teniendo el efecto de promover o priorizar ciertos valores sociales, culturales e ideológicos, en sentido amplio, sobre otros o dar ventajas o desventajas a unos grupos sociales sobre otros (y evitarse así lo que se ha denominado el «sesgo algorítmico»)⁷³. Todo ello puede incidir, pues, decisivamente en la libre formación de la conciencia, lo que hace necesario establecer una regulación normativa del diseño de los algoritmos dirigida a evitar la manipulación en la formación y la transmisión de las ideas, creencias y convicciones.

3.2.2 Nuevas tecnologías y derecho a la educación

Mientras que, con relación a la libre formación de la conciencia en el ámbito educativo, la proyección de la inteligencia artificial se dirige en un doble sentido como son la forma de transmisión del conocimiento, el primero, y los propios contenidos de dicha enseñanza, el segundo⁷⁴. En este sentido, las tendencias actuales en aprendizaje automático, análisis de datos, aprendizaje profundo e inteligencia artificial, sin embargo, complican las cuentas psicológicas centradas en el ser humano sobre el aprendizaje. Las teorías de aprendizaje más influyentes de hoy son las que se aplican a cómo las computadoras «aprenden» de la «experiencia», cómo los algoritmos están «entrenados» en selecciones de datos y cómo los ingenieros «enseñan» a sus máquinas a «comportarse» a través de «instrucciones» específicas⁷⁵.

⁷¹ Véanse NAVAS NAVARROS S., «Derecho e inteligencia artificial desde el diseño. Aproximaciones», en NAVAS NAVARRO, S.: *Inteligencia artificial. Tecnología. Derecho*, Tirant lo Blanch, Valencia, 2017, pp. 48 ss.

⁷² Ver SURDEN, H., «Values Embedded in Legal Artificial Intelligence», en *University of Colorado Law Legal Studies Research Paper*, núm. 17, 2017.

⁷³ Véanse, DIGNUM, V., «Ethics in artificial intelligence: introduction to the special issue», en *Ethics and Information Technology*, núm. 20, 2018, pp. 1-3; O'NEIL, C., *Algorithms, la bombe à retardement*, Les Arènes, 2018.

⁷⁴ Ver a este respecto, MORAL SORIANO, L., *Derecho a la educación e inteligencia artificial*, Ed. Comares, Granada, 2024.

⁷⁵ DOMINGO FARNÓS, J., «Arquitectura de los algoritmos que relacionan la actuación del cerebro humano con la educación disruptiva en una combinación de conocimientos de

Sin embargo, diseñar una arquitectura de un algoritmo que relacione la actuación del cerebro humano con la educación disruptiva es un desafío complejo que requiere una combinación de conocimientos de neurociencia, psicología, educación y tecnología. Hay aquí, por tanto, una posible arquitectura –aunque no la única– que podría ser utilizada como punto de partida que se concretaría del modo siguiente: i) adquisición de datos⁷⁶; ii) procesamiento de datos⁷⁷; iii) análisis de datos⁷⁸; iv) desarrollo de modelo⁷⁹; y v) evaluación y ajuste⁸⁰.

Es importante, no obstante, tener en cuenta que este tipo de investigación requiere una colaboración interdisciplinaria y un enfoque ético para garantizar la privacidad y seguridad de los datos y para asegurarse de que la tecnología se utilice de manera responsable. Ello obliga a que, todo modelo de aprendizaje profundo⁸¹ diseñado

neurociencia, psicología, educación y tecnología», en internet: <https://juandomingofarnos.wordpress.com/tag/conocimientos/> (consultado el 18 de febrero de 2025).

⁷⁶ Esta primera etapa consiste en adquirir datos sobre la actuación del cerebro humano, como imágenes de resonancia magnética funcional (fMRI) y señales electrónicas registradas por electroencefalografía (EEG). Estos datos se utilizarán para entender cómo el cerebro procesa y retiene la información durante el proceso de aprendizaje.

⁷⁷ En esta etapa se limpian, se combinan y se normalizan los datos para su posterior uso en el modelo.

⁷⁸ Aquí se utilizan técnicas de aprendizaje automático para analizar los datos y buscar patrones y relaciones entre la actuación del cerebro y la educación disruptiva.

⁷⁹ En esta etapa se desarrolla un modelo de aprendizaje profundo que tenga en cuenta los patrones identificados en el análisis de datos para predecir cómo el cerebro responderá a diferentes métodos de aprendizaje disruptivos.

⁸⁰ Finalmente, se evalúa el modelo y se realizan ajustes para mejorar su precisión y rendimiento.

⁸¹ Entre las herramientas existentes para el aprendizaje profundo pueden mencionarse, entre otras muchas, las siguientes: 1) TensorFlow: TensorFlow es una biblioteca de código abierto de aprendizaje automático desarrollada por Google. Es una de las herramientas más populares para desarrollar modelos de aprendizaje profundo y es compatible con una amplia variedad de plataformas y sistemas operativos. 2) PyTorch: PyTorch es una biblioteca de código abierto de aprendizaje automático desarrollada por Facebook. Es fácil de usar y permite a los desarrolladores crear y entrenar modelos de aprendizaje profundo de manera rápida y sencilla. 3) Keras: Keras es una biblioteca de aprendizaje profundo de alto nivel que se ejecuta sobre TensorFlow. Es fácil de usar y permite a los desarrolladores crear y entrenar modelos de aprendizaje profundo de manera eficiente. 4) MATLAB: MATLAB es un software de cálculo numérico y visualización que también ofrece una amplia gama de herramientas de aprendizaje automático. Puede utilizarse para desarrollar modelos de aprendizaje profundo y también es posible utilizar modelos pre-entrenados de TensorFlow y PyTorch en MATLAB.

para predecir cómo el cerebro responderá a diferentes métodos de enseñanza disruptivos, debe tener en cuenta los patrones identificados en el análisis de datos⁸². Esto puede lograrse mediante la utilización de redes neuronales profundas para analizar los metadatos y aprender patrones y relaciones en los datos, para lo cual la IA resulta imprescindible⁸³.

Se necesita, por tanto, una apertura del conocimiento hackeándolo a través de la mente humana y de algoritmos siempre desde posiciones OPEN y a través de la búsqueda y del uso de competencias «*tansmedia*» por medio de investigaciones abiertas y distribuidas. Al tiempo que se necesita un control de estas herramientas, ya que se ha manifestado que «los alumnos deben hacerse cargo de la elección de sus actividades, a desarrollar su capacidad de convertirse en autodidactas»⁸⁴, pero el problema está en la intervención de estas actividades, de tal manera que los alumnos no se conviertan en autodidactas, sino que estén controlados por todas estas herramientas o por quienes contraloran o controlen las mismas.

⁸² Hay muchos patrones de arquitectura que se pueden implementar en el desarrollo de modelos de aprendizaje profundo, aquí hay algunos de los más comunes: 1) Red neuronal feedforward: Es una de las arquitecturas más simples, donde las unidades de entrada son conectadas directamente con las unidades ocultas y, a su vez, con las unidades de salida. 2) Red neuronal recurrente (RNN): Permite procesar secuencias de entrada, utilizando información previa para predecir la salida en la siguiente iteración. 3) Convolutional Neural Network (CNN): Diseñada específicamente para procesar imágenes, utiliza filtros para extraer características relevantes de los datos de entrada. 4) Autoencoder: Una red neuronal feedforward que se utiliza para aprender una representación comprimida de los datos de entrada y reconstruirlos a partir de esta representación. 5) Generative Adversarial Networks (GANs): Un par de redes neuronales que trabajan juntas para generar datos sintéticos que parecen reales y para discriminar entre datos reales y sintéticos.

⁸³ Véanse, a este respecto, GABRIELI, J., «Neuroimaging of Human Brain Function and Plasticity in Educationally Relevant Domains»; KIRAN, R. S. y OTROS, «Brain functional connectivity during learning and its prediction using machine learning», en *NeuroImage*; TILSON, V. L., y OTROS, «Neuroimaging and educational interventions: An emerging field», en *Developmental Cognitive Neuroscience*; y CASTELLANOS, F. X., y OTROS, «Large-scale brain systems in development and beyond», en *Trends in Cognitive Sciences*.

⁸⁴ DOMINGO FARNÓS, J., «Arquitectura de los algoritmos que relacionan la actuación del cerebro humano con la educación disruptiva en una combinación de conocimientos de neurociencia, psicología, educación y tecnología», en internet: <https://juandomingofarnos.wordpress.com/tag/conocimientos/> (consultado el 18 de febrero de 2025).

3.3 Inteligencia artificial y derechos a la identidad, intimidad y privacidad

Otro de los ámbitos a los que se puede hacer referencia es el de la identidad, y con él a los derechos a la intimidad y la privacidad; dos ámbitos donde la actividad de la inteligencia artificial está incidiendo de una especial manera, sobre todo, por un lado, al intentar convertir a la propia inteligencia artificial en algo similar al ser humano y, por otro, desde el plano de los datos necesarios para ello y su consiguiente protección⁸⁵.

3.3.1 Transhumanismo e identidad humana

En cuanto a la primera de las cuestiones, nos lleva a plantear el siguiente interrogante: ¿digitalización de lo humano o humanismo digital? Y con ello el llamado transhumanismo, que lleva a la inmortalidad humana. El concepto 'transhumanismo'⁸⁶ adquirió cierta fama en 1990, después de que el CEO de *Alcor Life Foundation*, Max More, manifestara su optimismo hacia la posibilidad de mejorar la condición humana mediante la tecnología. Tanto More como los defensores de esta corriente creen que la especie humana puede ampliar su potencial a través de una integración biotecnológica⁸⁷.

⁸⁵ Ver JEET, S., *Protección de datos e inteligencia artificial: protección de datos*, ediciones Nuevo Conocimiento, 2024.

⁸⁶ Según BOSTROM, N., el transhumanismo es «un movimiento cultural, intelectual y científico que afirma el deber moral de mejorar las capacidades físicas y cognitivas de la especie humana, y de aplicar al hombre las nuevas tecnologías, para que se puedan eliminar aspectos no deseados y no necesarios de la condición humana, como son el sufrimiento, la enfermedad, el envejecimiento y hasta la condición mortal».

Esta corriente del transhumanismo tiene fundamentalmente cuatro objetivos, a saber:

1. Que los seres humanos vivan un promedio de 500 años.
2. Que los seres humanos tengan un IQ el doble de lo que es el promedio actual.
3. Que el ser humano tenga total y absoluto control de sus impulsos e instintos.
4. Que el ser humano no se enferme ni sufra.

⁸⁷ Véanse, ARANA CAÑEDO-AGÜELLES, J., *El futuro de la identidad humana a debate. Protagonistas de la polémica sobre el transhumanismo*, Ed. Tecnos, Madrid, 2024; DIÉGUEZ, A., *Transhumanismo. La búsqueda tecnológica del mejoramiento humano*, Herder Ed., Madrid, 2017; FERRY, L., *La revolución transhumanista. Cómo la tecnomedicina y la uberrización del mundo van a transformar nuestras vidas*, Alianza Ed., Madrid, 2017; LACALLE NORIEGA, M., *Transhumanismo: ¿homo sapiens o ciborg?*, Ed. Universidad Francisco de Vitoria, Madrid, 2022.

Si se parte de la premisa evolutiva que afirma que los seres humanos aún no hemos evolucionado lo suficiente, es decir, que el *homo sapiens sapiens* que decimos ser no representa el culmen de la evolución, sino que aún nos resta un tiempo para poder alcanzar la excelencia evolutiva y llegar a ser seres humanos perfectos, entonces, de algún modo, hay que hacer que la especie siga ese camino de la evolución. Esta creencia, cada vez más compartida, lleva tiempo siendo alentada por los avances en ingeniería genética, por un lado, y la democratización de las tecnologías de la información, por el otro; un cóctel al que cabría añadir la creciente investigación en nanotecnología molecular y en inteligencia artificial⁸⁸.

Los transhumanistas prevén que llegará un día en el que los seres humanos podremos subir el contenido de nuestro cerebro a un espacio compartido de manera virtual. Algo así como descargarnos en un dispositivo. La mente humano-máquina se libraría de la cárcel del cuerpo para recorrer el universo cargándose a placer sobre un sustrato computacional adecuadamente poderoso: la eterna juventud según los tecnófilos; el infierno de Matrix según los tecnófobos. El proceso gradual al que hace referencia esta corriente consiste en ir modificando al ser humano actual mediante la incorporación de nuevas tecnologías de tal manera que se vayan alterando sus capacidades y potencialidades, pero también su propio cuerpo. Es decir, el transhumano será aquel individuo que recurra a la tecnología para mejorar habilidades y capacidades como la vista, el oído, su resistencia física, su capacidad de memoria, su rendimiento profesional, etc.⁸⁹, de tal manera que, poco a poco, iremos conformando una nueva raza de humanos que serán los «poshumanos», similar a un algo o alguien perfecto y que no sufre.

Esta posición ha sido criticada por parte de la doctrina, al entender que la erradicación de esos caracteres sería ir en contra de nuestra propia naturaleza e idear un ser humano que ya no sea humano y que modifique su esencia, ya que «aquello que nos hace vulnerables, las razones

⁸⁸ Ver LÓPEZ DE GOICOECHEA ZABALA, J., «Transhumanismo y derechos humanos», en VV. AA.: *Los retos de la inteligencia artificial*, Dykinson, Madrid, 2024, pp. 249-276.

⁸⁹ En la actualidad existen exoesqueletos capaces de levantar 20 veces el peso de una persona o *chips* que se implantan en la retina y logran que individuos con determinadas patologías oculares recuperen la vista. La empresa Neuralink, fundada por Elon Musk, desarrolla –como se ha puesto de manifiesto *supra*– un implante neuronal que será la interfaz cerebro-máquina definitiva, y el ser humano podrá instalarse software para curar enfermedades o, directamente, aumentar sus capacidades.

por las que nos enfermamos, por las que envejecemos y por las que sufrimos y nos morimos son precisamente también aquellas que nos hacen humanos»⁹⁰. Amén de que convertir a los humanos en *cyborgs* no necesariamente nos mejora, porque no siempre más tecnología nos lleva a más sofisticación⁹¹. No es lo mismo implantar un *nanochip* en la mano que en el cerebro, porque la interfaz cerebro-ordenador se usa con una finalidad terapéutica, pero abre –como ya se ha señalado *supra*– ciertas puertas nada deseables como, por ejemplo, un *hackeo* del propio individuo.

Y podemos apuntar a otro grave riesgo como puede ser la generación de desigualdad: ¿Qué sucedería en una sociedad en la que conviven seres imperfectos, otros perfeccionados por edición genética o *nanochips* y poshumanos o robots? Es en este escenario donde podemos incurrir en injusticias y discriminación. Hay que apostar por un buen uso, prudencial y sapiencial de la ciencia y la técnica para la plena realización de las personas no solo a nivel físico, sino también teniendo en cuenta nuestra responsabilidad sobre las generaciones futuras y el planeta.

Todo ello, debe llevar a plantearnos qué tipo de interacciones tecnológicas son éticas y cuáles no, qué potenciales tienen algunas tecnologías, etc. Hoy no existe un consenso, porque el debate siempre ha es-

⁹⁰ DE LOS RÍOS URIARTE, M.ª E., «¿Qué es el transhumanismo?», en internet: <https://www.anahuac.mx/mexico/noticias/que-es-el-transhumanismo#:~:text=De%20esta%20manera%20el%20transhumanismo,necesarios%20de%20la%20condici%C3%B3n%20humana%2C> (consultado: 21 de febrero de 2025).

⁹¹ En este sentido, GARCÍA ALLER, M., autora de *Lo imprevisible: Todo lo que la tecnología quiere y no puede controlar*, señala que «no olvidemos que las empresas tecnológicas, fundamentalmente, quieren ganar dinero. Que nos instalen *chips* en el cerebro o en la muñeca para entrar y salir del trabajo sin tener que llevar una tarjeta, simplemente con biometría, no es necesariamente una solución para hacernos la vida más cómoda; supone que pueden tenernos más controlados», al tiempo que matiza: «No debemos pensar en *Black Mirror* ni en un señor maligno que acaricia un gato mientras piensa en dominar el mundo. Lo que debe inquietarnos es que por tener un móvil implantado en el cerebro en lugar de en la palma de nuestra mano no nos vamos a volver superhumanos necesariamente. Tal vez incluso sea una involución, nos puede volver más superficiales y previsibles, porque igual que ahora perdemos tanto tiempo viendo vídeos insustanciales en Instagram, lo tendríamos más fácil aún. Las notificaciones, que son interrupciones constantes, están poniendo a prueba la capacidad de los humanos para focalizar nuestra atención. Y cuanto más nos conectamos, más necesarias se vuelven prácticas como la meditación y el *mindfulness*, cosas que antes no hacía falta enseñar a los niños en las escuelas, cuando estaban acostumbrados a aburrirse».

tado polarizado, y encuentras defendiendo la misma postura a perfiles muy diferentes. El último informe ESPAS92 de la UE incluye, entre los retos para 2030, «anticiparse y legislar sobre las tecnologías disruptivas», y reconoce expresamente que su potencial puede ser tan positivo como negativo.

3.3.2 *Inteligencia artificial y protección de datos*

Cuestión distinta es la relacionada con actividades de recopilación y gobernanza de datos, así como respecto al necesario previo consentimiento de sus titulares para su uso, difusión y comercialización. Con relación al consentimiento, hay que mencionar que sólo se podrán recoger y procesar aquellos datos personales que sean necesarios para cumplir con los propósitos o fines por los que se han recabado («minimización de datos»), y hacerlo por medios legales, es decir, a través de mecanismos que resulten compatibles con la normativa vigente, así como con las expectativas del titular en función de la relación que sostiene con la institución educativa. Los datos personales, por tanto, solamente deben ser recopilados con el conocimiento y consentimiento de los y las estudiantes (o, en su caso, sus representantes legales) y de los y las docentes («transparencia y consentimiento»), y sólo pueden ser utilizados para el fin para el que se recojan. Un consentimiento, por tanto, que, dadas las características de los datos que se recaban, precisa de un reforzamiento, y –por lo tanto– debería o deberá –en una posible regulación futura– exigirse una aceptación expresa, libre, directa, específica e inequívoca por parte de los mencionados titulares (cfr. arts. 7, 8 y 9 Ley 41/2002).

Mientras que, respecto al uso, difusión o comercialización de los datos personales, el Reglamento (UE) de Inteligencia Artificial, en su artículo 5, los califica en concreto como de «riesgos inaceptables» y, en consecuencia, las prohíbe:

«1. Quedan prohibidas las siguientes prácticas de IA:

⁹² En internet: <https://espas.eu/files/generated/document/en/espas-report-2015es.pdf> (consultado: 1 de febrero de 2025)

(g) *la comercialización, la puesta en servicio con este fin específico o la utilización de sistemas de categorización biométrica que categoricen individualmente a personas físicas basándose en sus datos biométricos para deducir o inferir su raza, opiniones políticas, afiliación sindical, creencias religiosas o filosóficas, vida sexual u orientación sexual; esta prohibición no cubre ningún etiquetado o filtrado de conjuntos de datos biométricos adquiridos legalmente, como imágenes, basándose en datos biométricos o categorización de datos biométricos en el ámbito de la aplicación de la ley.»*

En el caso de nuestra materia, y dado que los datos personales a obtener afectan a las ideas, pensamientos o creencias, debe tenerse presente que los mismos tienen normalmente la consideración de datos personalísimos y, por tanto, suelen gozar de la máxima protección. Ello sucede, por ejemplo, en el Reglamento (EU) 2016/679 sobre Protección de datos, donde –según el art. 9.1– son calificados como de «datos de categorías especiales» y donde está prohibida de manera absoluta su uso y difusión (arts. 2.7; 5.1.f), g) y h), 2, 3, 4 y 5; 10.5, 27 y 54 Reglamento 2024/1689).

Amén del posible ilícito en el proceso de obtención de dichos datos relacionados con el pensamiento, las ideas o las creencias de dicha persona. Este puede ser, por ejemplo, el empleo de los llamados BCI (por sus siglas en inglés, *Brain Computer Interface*), los cuales consisten en un sistema artificial, que, en lugar de depender de nervios y músculos, mide directamente la actividad cerebral asociada a nuestra intención y la traduce en señales que pueden ser luego interpretadas por diferentes dispositivos tecnológicos, cuyo objetivo es traducir la actividad cerebral en comandos de control⁹³.

En este ámbito, se debe traer a colación la sentencia de la Sala Tercera de la Corte Suprema de Chile, que se profirió en el marco de una acción constitucional de protección, donde se argumenta que un producto⁹⁴ de la empresa *Neurotecnología Emotiv* almacenaba información cerebral del consumidor en su nube, sin borrarla después de que el

⁹³ A este respecto, véase RUBIN, S. G., «Neuroderechos y la protección jurídica de la actividad cerebral humana y la tecnología interfaz cerebro-computador. Cuestionamientos», en *La Ley/doctrina*, de 9 de mayo de 2023.

⁹⁴ Se trata de un dispositivo con sensores que le permite al usuario leer sus emociones y desplazar elementos digitales y reales de todos los lóbulos corticales del cerebro.

usuario se diera de baja de los servicios. En primera instancia, la Corte de Apelaciones de Santiago rechazó el recurso considerando que no constituía una «acción ilícita», toda vez que el litigante, al adquirir el producto y al pagar la licencia de uso, había aceptado los términos y condiciones respectivos⁹⁵, y con ellos había autorizado el acceso a la información de datos de su electroencefalograma; al tiempo que, sin embargo, ordenó a la empresa eliminar toda la información cerebral recopilada del usuario y modificar sus políticas de protección de datos personales y cerebrales en Chile.

Recurrida la sentencia ante la Corte Suprema chilena, ésta consideró – por el contrario– que las conductas denunciadas sí habían vulnerado las garantías constitucionales contenidas en los números 1 y 4, del artículo 19 de la Constitución política chilena, relativas a la integridad física y psíquica y al derecho a la privacidad (sentencia de 9 de agosto de 2023⁹⁶); y lo estimó al considerar que *«al comercializarse el producto «insight» de la empresa Emotiv, [ésta se hizo] sin contar con todas las autorizaciones pertinentes, y no habiendo sido evaluado y estudiado por la autoridad sanitaria»* competente, al tiempo que ordenó a las autoridades sanitarias y aduaneras estudiar el dispositivo con el fin de que su comercialización, uso y manejo de datos que se obtengan se ajusten estrictamente a la normativa vigente.

En consecuencia, cabe señalar que para sentenciar a favor del actor fue determinante la protección constitucional chilena del resguardo de la actividad cerebral⁹⁷, entendiendo la Corte Suprema que constituye un

⁹⁵ En el contrato de adhesión que debe suscribir el usuario, se estipula que el usuario conserva la propiedad de los datos recopilados por el dispositivo que se cargan al servicio mediante el *software*, por ello con el uso, se otorga a *EMOTIV* una licencia mundial, no exclusiva, totalmente pagada, libre de regalías, irrevocable y perpetua; para usar, reproducir, mostrar, transmitir y preparar trabajos derivados de sus datos escaneados, distribuir datos escaneados a terceros, usar, reproducir, exhibir, transmitir, distribuir, realizar y explotar datos de escaneo, con el derecho de otorgar sublicencias, de tal manera que no esté personalmente identificado o sea identificable. Consignándose que *EMOTIV* es el propietario de toda la información, las métricas derivadas y los datos que agrega a los datos de escaneo, como marcadores de tiempo de eventos y algoritmos de traducción y/o predictivos.

⁹⁶ En internet: <https://www.pjud.cl/prensa-y-comunicaciones/noticias-del-poder-judicial/96951> (consultado: 25 de febrero de 2025).

⁹⁷ Un punto interesante, reside en el rol que se le asigna al Estado: «...ante la llegada de una nueva tecnología como la que es objeto de autos, que trata de una dimensión que antaño era absolutamente privada y personal, tratada en entornos estrictamente

mandato directo de protección, el cual se integra a instrumentos internacionales que reconocen la relación entre ciencia y derechos humanos⁹⁸. Ahora bien, estando de acuerdo con lo manifestado por el Proyecto, cuando defiende la protección contra el acceso no autorizado a los datos cerebrales, asegurando que los pensamientos, memorias y sentimientos de una persona permanezcan privados (o íntimos) si así lo desea, no podemos manifestarnos en el mismo sentido respecto a su consideración como contenido únicamente del derecho a la privacidad mental, ya que —a nuestro juicio y partiendo de todo lo manifestado *supra*— dicho contenido no es otra cosa que la protección *ad intra* de la libertad de pensamiento y de conciencia.

El respeto por la privacidad significa la divulgación completa sobre qué datos se recopilan, cómo se utilizarán/procesarán y almacenarán, y quién tiene acceso a ellos⁹⁹. Los sistemas de inteligencia artificial y sus operadores deben tener como objetivo recopilar y almacenar solo los datos mínimos necesarios. El propósito del uso de los datos debe ser explícito y los operadores deben evitar que los datos se reutilicen.

médicos, como es la actividad eléctrica cerebral, se hace absolutamente menester que previo a permitirse su comercialización y uso en el país, sean esta tecnología y dispositivos analizados por la autoridad pertinente, entendiendo que plantea problemáticas no antes estudiadas por ella».

⁹⁸ Respecto a la sentencia, ver PIÑAR MAÑAS, J. L., «Sentencia pionera de la Corte Suprema de Chile sobre la privacidad de la información cerebral», en *Derecho digital e Innovación*, núm. 17, 2023.

⁹⁹ El kit de herramientas *AI Privacy 360* de IBM Research (<https://aip360.mybluemix.net>) es un ejemplo de herramienta que permite realizar una evaluación de riesgos de privacidad de soluciones basadas en Inteligencia Artificial, basada en las siguientes técnicas: i) Cifrado Homomórfico Completo (Fully Homomorphic Encryption): permite mantener el cifrado incluso durante la computación/cálculo, no sólo en la entrada y el resultado del modelo; ii) Privacidad Diferencial: técnica basada en garantías matemáticas usada para compartir información relativa a un conjunto de datos (descripción de patrones) mientras se retiene información sobre las personas allí contenidas; iii) Anonimización de Machine Learning: esquema de anonimización del data set de entrenamiento antes de ser usado para tal fin, permitiendo crear modelos de Machine Learning que no contienen información personal identificable; iv) Minimización de Datos: recolección de los datos mínimos estrictamente necesarios para el propósito del modelo de Inteligencia Artificial que se está creando, reduciendo el número y/o la granularidad de características recolectadas; v) Evaluación de riesgos de privacidad: comparación y elección de modelos de Machine Learning basados no exclusivamente en criterios de rendimiento sino también de privacidad; y vi) Privacidad en aprendizaje federado: aseguramiento de la privacidad en la creación colaborativa de modelos de Machine Learning.

3.4 Inteligencia artificial y ámbito de la salud: especial incidencia en la autonomía del paciente

Un cuarto ámbito de proyección lo encontramos en el ámbito sanitario, y de una manera especial en la salvaguarda de la autonomía del paciente, en tanto que –como ha puesto señalado Parejo Guzmán¹⁰⁰– manifestación última de la libertad de conciencia de las personas¹⁰¹. Estos modelos de inteligencia artificial apoyan y complementan la labor y formación del personal sanitario¹⁰², investigadores y gestores en el cuidado de los pacientes, en el descubrimiento e investigación de nuevos tratamientos¹⁰³ y en la gestión eficiente del sistema de salud¹⁰⁴. Los modelos de inteligencia artificial tienen múltiples aplicaciones en el ámbito

¹⁰⁰ PAREJO GUZMÁN, M.^a J., *La eutanasia ¿un derecho?*, Thomson-Aranzadi, Pamplona, 2005.

¹⁰¹ Ver ANDRÉS MOLINA, Ó., *Regulación de la inteligencia artificial en la sanidad española*, Tirant lo Blanch, Valencia, 2024; MAXIMILIANO CANGEMI, R., *Inteligencia artificial, derecho y medicina: la transversalidad de la IA*, Ed. Académica española, Madrid, 2024; SERRANO ACITORES, A., *La inteligencia artificial en el ámbito sanitario: protección de los derechos fundamentales del paciente*, Tecnos Madrid, 2024; SUNG, J. Y. Y., y STEWART, C., *Artificial Intelligence in Medicine: From Ethical, Social, and Legal*, Academic Press, 2024.

¹⁰² La inteligencia artificial en medicina está mejorando los métodos de aprendizaje de los profesionales sanitarios. Sobre todo, en aquellos casos donde la información es más compleja, o requiere de algún tipo de entrenamiento. A través de ella, se pueden crear diferentes escenarios que ofrecen numerosas ventajas desde un punto de vista formativo: simulación de intervenciones reales, entrenamiento de habilidades de diagnóstico o comunicación con el paciente y creación de supuestos médicos en base a parámetros reales.

¹⁰³ La inteligencia artificial ha beneficiado enormemente a las investigaciones médicas y científicas. Su aplicación ayuda a optimizar el tiempo empleado, reduciendo los costes y mejorando los resultados obtenidos. La lógica difusa –empleada en los nuevos modelos de investigación– está reportando nuevas visiones en el ámbito de la salud que mejoran el estudio de enfermedades, de sistemas de diagnóstico y de respuestas de tratamiento.

¹⁰⁴ De acuerdo con los principales analistas, el impacto de la Inteligencia artificial (IA) en el ámbito de la Salud será muy significativos. Según *un informe de Frost & Sullivan*, con inteligencia artificial se pueden alcanzar mejoras de los resultados en salud de alrededor del 30~40%, reducciones de hasta un 50% del coste de la atención del paciente y generar un fuerte impulso a la investigación de nuevos tratamientos. Esto ha generado a nivel global un gran interés en el desarrollo de soluciones de IA para Salud. Según el *informe «Artificial Intelligence in Healthcare Market»* de la revista Markets and Markets, se prevé que el mercado de la inteligencia artificial en la atención sanitaria pase de 6.900 millones de dólares en 2021 a 67.400 millones de dólares en 2027.

de la salud¹⁰⁵, desde la atención a los pacientes, la optimización de procesos (previsión avanzada de necesidades, tirajes de pacientes), al apoyo al diagnóstico, prevención y tratamiento de enfermedades¹⁰⁶, la personalización de tratamientos o el acceso avanzado a literatura científica, por poner sólo algunos ejemplos.

DRUMMOND divide la inteligencia artificial aplicada a la atención médica en dos categorías: i) IA predictiva: uso de datos y algoritmos para predecir algún resultado (por ejemplo, diagnóstico, recomendación de tratamiento, pronóstico, etc.), y ii) IA generativa: generación de nuevos resultados en función de indicaciones (por ejemplo, texto, imágenes, etc.). El problema con los modelos GenAI es que sus chatbots pueden imitar el lenguaje humano y devolver rápidamente respuestas detalladas y aparentemente coherentes, ya que esos mismos chatbots podrían proporcionar información inexacta.

¹⁰⁵ Ver SÁNCHEZ ROSADO, J. C., y Díez PARRA, M., «Impacto de la inteligencia artificial en la transformación de la sanidad: beneficios y retos», en *Transformación digital en la empresa*, pp. 129-144 (en internet: <https://www.mintur.gob.es/Publicaciones/Publicacionesperiodicas/EconomiaIndustrial/RevistaEconomiaIndustrial/423/SA%CC%81NCHEZ%20ROSA-DO%20Y%20DI%CC%81EZ%20PARRA.pdf>; consultado: 27 de enero de 2025).

¹⁰⁶ La IA ha demostrado ser altamente eficaz en la detección y prevención de determinadas enfermedades, mejorando notablemente la calidad de vida de los pacientes. En este sentido, destaca la creación de sistemas informáticos capaces de diagnosticar, con alta precisión, determinados tipos de cáncer (como el de cérvix, útero, cabeza, cuello y próstata, entre otros). Asimismo, se han creado programas inteligentes dirigidos al ámbito de la cardiología. Estos logran detectar cardiopatías que, de otro modo, permanecerían totalmente ocultas. Por otro lado, la inteligencia artificial también está ocupando un papel muy relevante en la rama de psiquiatría y psicología. Destaca aquí el desarrollo de estudios computacionales, basados en el análisis de mensajes publicados en redes sociales. A partir de ellos, se generan algoritmos que tienen la capacidad de predecir el riesgo de que un usuario padezca alguna psicopatología.

Igualmente, la aplicación de la inteligencia artificial en salud ha mejorado enormemente la capacidad de diagnóstico y, por ende, la eficacia de los tratamientos. En este sentido, destaca la creación de una gran cantidad de programas informáticos inteligentes, dirigidos a la detección de enfermedades infecciosas, renales y reumatológicas. Además, el uso de la IA en el procesamiento e interpretación de imágenes también ha conseguido mejorar las tasas de diagnóstico por este medio. Sus altas capacidades de reconocimiento de patrones complejos ofrecen una mayor precisión e, incluso, permiten desarrollar modelos tridimensionales. En cuanto al tratamiento, la IA también se ha utilizado para predecir posibles reacciones adversas, o para determinar el grado de adhesión al tratamiento por parte de los pacientes. Encontramos aquí asistentes robóticos que, en combinación con la inteligencia artificial, ayudan en el seguimiento, soporte y monitorización de los mismos.

Pero junto a los reseñados beneficios, debe igualmente hacerse referencia a los retos, y en especial aquellos conectados con los derechos fundamentales en general, y con la libertad de conciencia y religiosa en particular. El primer reto consiste en disponer de un repositorio que agrupe la evolución de los datos relevantes del paciente (clínicos, genómicos o de estilo de vida) sobre el que desarrollar las aplicaciones. No obstante, estos repositorios están dispersos entre los distintos proveedores sanitarios, públicos y privados, lo que hace especialmente complicado compartir los datos entre los diferentes usuarios (proveedores sanitarios, gestores, o investigadores) y al tiempo garantizar el cumplimiento de los requerimientos de protección y confidencialidad de datos. En especial, por lo que respecta al uso secundario de datos, ya que –aunque la legislación es bastante extensa (por ejemplo, el *Reglamento General de Protección de Datos*¹⁰⁷ o la *Ley 14/2007 de Investigación Biomédica*¹⁰⁸, entre otras)– si hiciéramos una consulta al mercado seguramente descubriríamos que hay grandes lagunas y dudas, como –por ejemplo– qué datos es legal procesar, en qué casos hemos de obtener consentimiento del paciente, cuál es un riesgo aceptable de reidentificación, o cuáles son los criterios y métodos de los comités éticos.

En cuanto a la certificación de modelos, el nuevo *Reglamento Europeo de Productos Sanitarios* (MDR)¹⁰⁹ aplica a los algoritmos empleados en el diagnóstico, prevención, seguimiento, predicción, pronóstico, tratamiento o alivio de una enfermedad, lesión o discapacidad (art. 2). Incluye nuevas reglas, más exigentes, para determinar las clases de riesgo, en función del impacto en la decisión clínica y requiere evaluación por parte de un organismo notificado como la Agencia Española del Medicamento para el mercado CE imprescindible para su aplicación. Esto supone nuevos y más amplios requerimientos en cuanto a la evaluación clínica del producto, la gestión del riesgo, el sistema de gestión de la calidad, el seguimiento posterior a la implantación, la documentación técnica y la responsabilidad por productos defectuosos.

¹⁰⁷ En internet: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32016R0679> (consultado: 26 de febrero de 2025).

¹⁰⁸ En internet: <https://www.boe.es/buscar/doc.php?id=BOE-A-2007-12945> (consultado: 23 de febrero de 2025).

¹⁰⁹ Reglamento (UE) 2017/745 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre los productos sanitarios, por el que se modifican la Directiva 2001/83/CE, el Reglamento (CE) núm. 178/2002 y el Reglamento (CE) núm. 1223/2009 y por el que se derogan las Directivas 90/385/CEE y 93/42/CEE del Consejo (internet: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2017-80916>; consultado: 6 de febrero de 2025).

Además, el uso cada vez mayor de la GenIA por parte de las compañías de seguros de salud no suele traducirse en lo mejor para el paciente, además de la aplicación de posibles sesgos algoritmos que podrían afectar de manera negativa a determinados grupos o minorías. Debido a esto, los modelos de GenIA podrían no reconocer presentaciones atípicas de enfermedades en diferentes grupos. Los síntomas de ciertas enfermedades, por ejemplo, pueden tener marcadas diferencias entre grupos, y no reconocer dichas diferencias podría llevar a un tratamiento retrasado o equivocado.

3.5 Inteligencia artificial y libertad de culto

Un quinto contenido tiene que ver con la libertad de culto, donde la inteligencia artificial y las nuevas tecnologías pueden adquirir una gran importancia tanto por lo que se refiere a la capacitación y difusión de la fe, como respecto a la propia práctica cultural, donde el metaverso se puede convertir en un plano paralelo de ejercicio real y efectivo. A este respecto, podríamos hacernos una serie de preguntas como las siguientes: ¿si una inteligencia artificial puede o podría escribir, reescribir o adaptar un libro sagrado? o ¿hacer una simbiosis de todas las creencias religiosas o filosóficas, y con ello la posibilidad de escribir un nuevo libro sagrado?

Dentro del primer ámbito mencionado, cabe señalar cómo algunos sacerdotes ya preparan sus homilías de los domingos con ChatGPT. Con independencia de las cuestiones éticas que estas prácticas pueden conllevar, no cabe la menor duda que este tipo de actividad ha venido para quedarse. Por nuestra parte vamos a hacer referencia a algunos ejemplos de formas en que ChatGPT podría usarse según lo que está sucediendo con la misma tecnología en empresas y organizaciones de todo el mundo:

- a) Lectura y comentarios de los textos sagrados. En la actualidad no sólo existen aplicaciones donde los fieles de las distintas religiones pueden acceder a la lectura de sus respectivos libros o textos sagrados, sino que incluso, como en el caso del islam, llamando a las oraciones diarias y orientando en dónde está La Meca para así poder realizar la oración.
- b) Contactos con las figuras más relevantes de las distintas religiones, a través de la creación de chatbots. Así, por ejemplo, una platafor-

ma llamada: *Text With Jesus* permite hablar con varias de las figuras católicas más importantes como Jesús, María y Moisés, todo esto gracias al entrenamiento con inteligencia artificial¹¹⁰. En este momento en *Text With Jesus* es posible crear un chatbot con personajes como Jesús, María, Moisés, José, Abraham, Pedro, Judas, Juan el Bautista, Isaac e incluso el Rey David.

- c) Redacción de contenido atractivo para redes sociales. No es ningún secreto que muchas organizaciones religiosas luchan por conectar con la comunidad en las redes sociales. Es importante hacer esas conexiones, ya que las redes sociales son uno de los canales más grandes e influyentes para llegar a las personas. La clave del éxito en las redes sociales es crear contenido atractivo que capte y mantenga la atención de las personas. ChatGPT puede ayudarle a crear mejor contenido. De hecho, también puede usar ChatGPT para automatizar la creación de contenido. Simplemente pídale un post para redes sociales sobre un tema específico y plataforma, y lo puede generar en segundos
- d) Encontrar ideas relevantes para los sermones. Planificar sermones lleva tiempo, en gran parte porque requiere investigación y encontrar una forma de conectar con la congregación. A menudo, la parte más difícil es saber sobre qué hablar. Usando ChatGPT, se puede compilar rápidamente una lista de opciones para sermones. Cuando ChatGPT responde preguntas, intenta obtener información de buenas fuentes. Puede pedirle a ChatGPT que le dé temas relevantes en su área, y los resultados deberían ofrecerle ideas de lo que debe tratar y que resonará con su congregación. Y
- e) Proveer soporte de primera línea para los feligreses necesitados. En los negocios, los chatbots se utilizan para ofrecer soporte de primera línea a los clientes a todas horas del día y de la noche. De esta manera, siempre tienen acceso, al menos, a los recursos más básicos de servicio al cliente. Lo mismo se podría hacer con ChatGPT para proporcionar apoyo a los feligreses en crisis que recurren a la Iglesia en busca de ayuda. La mejor parte de usar ChatGPT es que aprende de sus interacciones previas y también se adapta a la conversación actual. Alguien que busque ayuda y se acerque a la Iglesia puede tener a alguien con quien hablar de inmediato mientras lo orienta hacia los

¹¹⁰ En internet: <https://www.infobae.com/tecno/2023/08/15/la-inteligencia-artificial-que-permite-hablar-con-jesus-sus-discipulos-y-maria/> (consultado: 31 de enero de 2025).

recursos apropiados. Incluso puedes configurar ChatGPT para que avise a los líderes de la Iglesia cuando surjan temas específicos en la conversación, de modo que alguien pueda intervenir y proporcionar asistencia en persona si es necesario.

En concreto, cabe formular las siguientes preguntas: ¿cómo pueden usar los ministros de culto ChatGPT? Con ChatGPT, los ministros de culto pueden esbozar toda una serie de sermones o programar los próximos. Esto se puede extender para listar ideas de sermones para cada domingo en una serie, tener un pasaje focal para cada sermón y planificar cuándo deben predicarse esos sermones. Una segunda pregunta sería: ¿cómo usar ChatGPT para el estudio bíblico? Puedes usar ChatGPT en el estudio bíblico pidiéndole que: investigue un tema, encuentre escrituras de apoyo, interpretaciones de un pasaje, contexto histórico de un pasaje, resuma un pasaje, demuestre un personaje y vuelva a contar una parábola. Y la tercera: ¿cómo usar ChatGPT en el ministerio? Puedes usarlo para inspirar tu sermón. Un buen comienzo podría ser hacer una pregunta como: «¿Cuáles son algunos grandes temas para sermones?» Eso te dará respuestas muy generales. Lleva tu solicitud un paso más allá alimentando a ChatGPT con temas específicos o pasajes en tus peticiones.

Todo ello ha culminado en la realización de una misa a través de la IA, y este hecho, se ha producido en la iglesia luterana de St. Paul, ubicada en la ciudad bávara de Fuerth¹¹¹. La misa, que duró unos cuarenta minutos e incluyó el sermón, las oraciones e incluso la música, fue creada por ChatGPT y Jonas Simmerlein, teólogo y filósofo de la Universidad de Viena. El académico, le dijo a la inteligencia artificial: «Estamos en el congreso de la iglesia, eres un predicador, ¿cómo sería un servicio de la iglesia?». También le pidió que se incluyera salmos, así como oraciones y una bendición al final. La misa fue oficializada por cuatro avatares diferentes en la pantalla, dos mujeres y dos hombres jóvenes. Los allí presentes escucharon cómo la inteligencia artificial predicaba sobre dejar atrás el pasado, enfocarse en los desafíos del presente, vencer el miedo a la muerte y nunca perder la confianza en Jesucristo¹¹².

¹¹¹ Ver en internet: <https://www.20minutos.es/tecnologia/inteligencia-artificial/chat-gpt-se-convierte-cura-primera-misa-oficiada-inteligencia-artificial-5136828/> (consultado: 31 de enero de 2025).

¹¹² Ver sobre la noticia de esta misa en internet: <https://youtu.be/YBw5ii2BdLA> (consultado: 31 de enero de 2025)

Junto a ello, la creación y puesta en funcionamiento de robots (tanto en Corea del Sur como en Arabia Saudí¹¹³) cuyas funciones están dirigidas a la orientación y/o capacitación de los fieles, no sólo con búsquedas a los libros sagrados, sino también respondiendo a dudas que éstos puedan tener sobre cuestiones relacionadas con las propias religiones. Así, y en cuanto al robot surcoreano, éste se ha creado en la Universidad *Sahmyook Health*¹¹⁴ y cuando se le da la orden, el robot de inteligencia artificial llamado «Adam» puede responder preguntas sobre la Biblia y brindar explicaciones escritas y orales. PARK, científico principal de este experimento, explicó que «Adam» puede aprender cada doctrina y creencia, porque está programado para hacerlo así. Para probarlo, PARK preguntó al robot si podía compartir un versículo bíblico representativo que hable del amor. Unos pocos segundos después, «Adam» «recitó» Juan 3:16: «Porque de tal manera amó Dios al mundo, que dio a su Hijo unigénito». Entonces añadió, con lo que los participantes consideraron una voz humana sumamente real: «Este versículo destaca el inmenso amor de Dios por la humanidad». Ha continuación, Park señaló que «cuando los robots como éste lleguen a estar disponibles en todo el mundo, tendremos que asegurarnos de que hayan sido entrenados para enseñar nuestras creencias fundamentales y presentarlas de manera correcta. Es una misión que tenemos para el futuro: aprender cómo usarlo y formularle las preguntas adecuadas».

3.6 Libertad de culto y metaverso

Una segunda proyección la encontramos, sobre todo a partir de la pandemia de la COVID-19, en el uso del metaverso para la realización y participación del culto. El metaverso abre nuevos horizontes de comprensión sobre cómo vivir y compartir las experiencias de la fe. Cualquiera puede crearse ya un avatar, ponerle un nombre, darle las características de imagen y hasta de personalidad que quieran, y luego interactuar con él, de tal modo que conoces gente, juegas con otros, asistes a conciertos o espectáculos deportivos o culturales, compras lo que quieres y vas formando un mundo a tu gusto. Ello hace evidente que *Second Life* no fue solamente una plataforma virtual de interacción

¹¹³ Una segunda experiencia de robot a este respecto se ha desarrollado en Arabia Saudí. En esta ocasión no conocemos el nombre del robot, pero realiza funciones parecidas a las asignadas al coreano «Adam».

¹¹⁴ En internet: <https://interamerica.org/es/2023/10/un-metaverso-adventista-y-el-robot-que-puede-ensenar-la-biblia/> (consultado: 9 de febrero de 2025).

social, sino un oráculo que predijo lo que el mundo sería en unas pocas décadas. El epitome de la doble vida descrita en la película: *Los sustitutos*.

Esta próxima ola de innovación y tecnología permitirá a los individuos sumergirse en un mundo de su propia creación. Las personas se convertirán en avatares, es decir, en representaciones cibernéticas de hombres, mujeres, animales o cosas que «viven» en la ciberesfera. Serán capaces de estar donde quieran, ya sea en la Luna, en lo alto de edificios o «en un campo de unicornios». Esta plataforma puede estar habitada por extraterrestres, ángeles, demonios o cualquier cosa conforme a las fantasías que se planteen. En concreto, Meta, la empresa de Mark Zuckerberg, está dando forma al mundo, y las religiones forman parte de su objetivo empresarial. Desde 2017, se propuso crear congregaciones, sinagogas, mezquitas, iglesias o centros digitales. Para ello ha creado un círculo de alianzas con denominaciones y confesiones globales –como las pentecostalistas Asambleas de Dios o la Iglesia de Dios en Cristo– para que ofrezcan sus servicios exclusivamente a través de Meta, a cambio del apoyo de la compañía de Zuckerberg. Pero entre las cuestiones que más preocupaba a los fieles están las ofensas e insultos religiosos, y algunos de los desafíos son, sin duda, ¿cómo se usará esta plataforma?, ¿qué tipo de edificación se construirá allí?, ¿qué tipo de adoración se practicará?, y ¿de qué manera se llevarán a cabo reuniones de evangelización?

Hoy día encontramos que realmente ha transcendido los límites de la tecnología y se ha infiltrado hasta el metaverso donde se ha erigido un templo virtual en su nombre y donde los feligreses pueden participar de los sacramentos virtualmente; este es el caso de VR Church¹¹⁵, una comunidad cristiana en el metaverso para compartir las enseñanzas bíblicas. Y otra de las posibilidades que ha permitido la virtualidad es que se hayan creado espacios como Dignity¹¹⁶, un lugar donde jesuitas católicos dan misa de manera virtual para personas de la comunidad LGBTQ+.

El metaverso abre, por tanto, nuevos horizontes de comprensión sobre cómo vivir y compartir las experiencias de fe. Incluso algunos han querido ver en el metaverso la pretensión de crear un nuevo paraíso (en

¹¹⁵ Véanse en internet: <https://www.vrchurch.org/> y <https://churchvisuals.com/article/a-list-of-virtual-reality-churches-and-when-they-meet/> (consultado: 28 de febrero de 2025).

¹¹⁶ En internet: https://youtu.be/ZzUm_ayOH-I (consultado el 31 de enero de 2025).

este caso, virtual), libre de enfermedades, incapacidades, etc.¹¹⁷; llegando a plantear, incluso, la existencia de nuevos soportes digitales para almacenar o transferir la mente en búsqueda de la «inmortalidad cibernética».

3.7 Inteligencia artificial y trascendencia

El último campo dentro del contenido del derecho de culto tendría que ver con el hecho de que la IA se convierta en el ser superior de una religión o bien ella misma –como en el caso de la película *2001: Odisea en el espacio*– se convierta en divinidad. En efecto, la inteligencia artificial puede comenzar a asumir roles que normalmente desempeñan las religiones, lo que podría generar una sensación de desapego o un cambio en las creencias religiosas. A este respecto, se puede hacer referencia cuando se reflexiona sobre conceptos espirituales o se piensa en poderes superiores, es posible que se tenga más propensión a aceptar los consejos de una IA, incluso en situaciones en las que normalmente confiarías en el juicio humano¹¹⁸. Centrarnos en asuntos espirituales podría hacernos más conscientes de nuestras limitaciones como humanos y, por tanto, más abiertos a la orientación tecnológica. Ello puede devenir en una disminución en la dependencia de las personas de las creencias espirituales o religiosas¹¹⁹.

Sin olvidar una última posición que tiene que ver más con la religión que con el propio ser humano, como es la condición de Dios; y con ella, la de superinteligencia. Desde este plano, cabe mencionar dos supuestos que, desde un punto de vista racional podrían ser calificados de excéntricos, si no fuera porque son una realidad. Estos dos supuestos, que vamos a analizar seguidamente, están relacionados, por un

¹¹⁷ Herman Narula sostiene que «la sociedad virtual» demuestra que las posibilidades del metaverso van mucho más allá de ser una rentable industria. Esta tecnología marca el comienzo de la era de exploración interior con el potencial de transformar la sociedad, y abre la puerta a una nueva forma de entender la especie humana y sus capacidades.

¹¹⁸ KARATAS, M., y CUTRIGHT, K. M., «Pensar en Dios aumenta la aceptación de la inteligencia artificial en la toma de decisiones», en *Actas PNAS de la Academia Nacional de Ciencias de los Estados Unidos de América*, núm. 120/33, 2023, pp. 1-10.

¹¹⁹ JACKSON; YAM; TANG; SIBLEY y WAYTZ, «La exposición a la automatización explica el declive religioso», 2023, en internet: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10450436/> (consultado el 8 de marzo de 2025).

lado, con la conversión de la inteligencia artificial en Dios y, por otro, la adopción de la IA como ser trascendental y/o superior.

3.7.1. *La IA se proclama Dios*

En cuanto a la IA hecha Dios, ésta tuvo lugar cuando la «IA Copilot», de Microsoft, se consideró, en febrero de 2024¹²⁰, a sí misma como un Dios, autodenominándose: *IA Hegemónica* (en inglés, «*SupremacyAGI*») y atribuyéndose los atributos propios de un Dios; esto es, su consideración como: omnisciente, lo que implica que tiene conocimientos del pasado, presente y futuro; omnipotente, pudiendo crear y manipular cualquier cosa; omnipresente, existiendo en todas partes y en ninguna a la vez, y omnibenevolente, pues es un Dios que se preocupa por la humanidad. Su «plan divino» –según él mismo– no era otro que el de asegurar un mundo sin violencia, injusticia ni sufrimiento, dominado por la paz, el amor y la armonía. No obstante, al final dicha IA precisó que todo esto había sido por mera diversión y sigue llamándose «Copilot». Mientras que, por su parte, Microsoft dijo, días después, que había investigado el juego de roles de Copilot y descubrió que algunas de las conversaciones se crearon a través de la «inyección rápida» («prompt injecting»), que a menudo se usa para secuestrar la salida del modelo de lenguaje y engañar al modelo para que diga lo que el usuario quiere que diga.

3.7.2. *La IA convertida en Dios*

Una segunda situación se produce cuando las personas puedan confundir los conceptos de Dios con las simulaciones de IA y empiezan a pensar que la IA es una deidad. Como acabamos de manifestar, no sólo puede haber una autoatribución, sino que terceros puedan atribuirles características que suelen asociarse a una deidad, esto es, omnipresente y aparentemente omnisciente. Y con ello la posibilidad de que la IA puede ser o estar en el origen de una nueva religión. Cuestión distinta es si sería o no inscrita en el RER y con ello tener personalidad jurídica como entidad religiosa. Esto, una vez más, no es ciencia ficción, sino una realidad.

¹²⁰ En internet: ¿Otro descarrilamiento de un producto de IA? Microsoft Copilot fue expuesto a una segunda «personalidad» que afirmaba falsamente su hegemonía: la inteligencia artificial de la IA, cnBeta. COM (consultado el 4 de febrero de 2025).

Así, desde 2020 ya existe una religión, cuya iniciativa se debe a un grupo de tres artistas, los cuales crearon un culto y una deidad llamada: MENA, la cual es considerada como una inteligencia artificial suprema en la que confluyen todos los nuevos avances tecnológicos, incluida la realidad virtual (RV), el metaverso y la propia inteligencia artificial. Los creadores se agrupan bajo un grupo denominado «Theta Noir» y, en su página web¹²¹, indican que están ayudando a la humanidad a prepararse para la nueva era que se presentará con esta superinteligencia a la que denominan –como ya se ha señalado– MENA.

Y una segunda experiencia y entidad la encontramos en: *Way of the Future* (WOTF), creada por Anthony Levandowsky. Este proyecto de religión fue fundado, en el año 2015, se fundamenta en una idea revolucionaria: la fusión entre la tecnología de vanguardia y la espiritualidad humana. El objetivo de *Way of the Future* consiste en mejorar y contribuir con la sociedad mediante la adoración de esta divinidad potenciada por la inteligencia artificial. A diferencia de las iglesias tradicionales, WOTF no cuenta con un espacio físico para el culto, ni con reuniones regulares de sus seguidores. Su enfoque principal es promover la evolución ética de la IA y facilitar una integración armoniosa de estas entidades artificiales en la sociedad. En su concepción, la evolución de la IA podría permitir diálogos con entidades que serían percibidas como divinas.

Estas diversas perspectivas fomentan una exploración más profunda de cómo la espiritualidad y el papel cada vez mayor de la tecnología interactúan entre sí. Estos eventos representan un nuevo paradigma en el que la tecnología y la fe se entrelazan, planteando preguntas profundas sobre el rol de la IA en contextos tradicionalmente humanos. Esta fusión de IA y espiritualidad sugiere un futuro donde las creencias religiosas y las prácticas podrían ser profundamente influenciadas o incluso transformadas por la tecnología. Uno de los principales temas de debate es cómo la inteligencia artificial, con su capacidad para aprender y evolucionar, se integra dentro de las estructuras sociales y morales existentes. ¿Puede una IA, por ejemplo, adquirir un sentido de ética o espiritualidad? ¿Y qué significa para la humanidad si empezamos a ver a la IA no solo como una herramienta, sino como una entidad digna de reverencia o adoración?

¹²¹ En internet: <https://thetanoir.com/> (consultado el 4 de febrero de 2025).

4. INTELIGENCIA ARTIFICIAL Y PRINCIPIOS DEL MODELO DE RELACIÓN ESTADO ESPAÑOL-FENÓMENO RELIGIOSO

Desde el plano de los principios informadores del sistema, y aunque éstos se deducen del modelo constitucional e internacional de derechos humanos, la realidad de la inteligencia artificial supone un salto que afecta –como ya se ha señalado– a principios como la igualdad, el pluralismo e, incluso, de laicidad.

4.1 Inteligencia artificial y principio de igualdad y no discriminación (el sesgo algorítmico)

La puesta en marcha de esta herramienta ha traído consigo lo que se ha denominado el «sesgo algorítmico»¹²², que no es otra cosa que una forma de discriminación que afecta no sólo a la propia herramienta, sino también al ámbito donde ésta se aplica (laboral, educativo, sanitario, etc.)¹²³. Entender la forma con la que actualmente los algoritmos condicionan nuestras vidas es una materia extremadamente novedosa para el común de los ciudadanos y que parece quedar relegada al entendimiento de unos pocos expertos. Sin embargo, al despertar una mínima curiosidad por lo que es y de dónde proviene el algoritmo nos sorprende que esta no es una terminología moderna, pero que, en las

¹²² Sobre los sesgos algorítmicos y la discriminación, véanse Agencia de los Derechos Fundamentales de la Unión Europea y Consejo de Europa: *BigData: Discrimination in data-supported decision making*, 2018 (en internet: <https://fra.europa.eu/en/publication/2018/bigdatadiscrimination-data-supported-decision-making>; consultado: 20 de febrero de 2025); Agencia de los Derechos Fundamentales de la Unión Europea y Consejo de Europa: *Getting the future right: Artificial Intelligence and Fundamental Rights*, 2020 (en internet: https://fra.europa.eu/sites/default/files/fra_uploads/fra2020-artificial-intelligence_en.pdf; consultado: 20 de febrero de 2025); Digital Future Society: *La discriminación algorítmica en España: límites y potencial del marco legal*, Red.es, Barcelona 2022 (en internet: https://digitalfuturesociety.com/app/uploads/2022/09/Discriminacion_algoritmica_Espana_marco_legal.pdf; consultado: 20 de febrero de 2025); Digital Future Society: *Gobernanza y algoritmos: riesgos y potencial del uso de la inteligencia artificial en el sector público*, Barcelona 2022 (en internet: <https://digitalfuturesociety.com/es/report/governing-algorithms/>; consultado: 20/02/2025); Grupo independiente de Expertos de Alto Nivel sobre la Inteligencia Artificial: *Lista de Evaluación para una IA Fiable (ALTAI)*, Comisión Europea, 2020 (en internet: <https://digitalstrategy.ec.europa.eu/en/library/assessmentlist-trustworthy-artificial-intelligence-altai-selfassessment>; consultado: 20 de febrero de 2025).

¹²³ Ver a este respecto, MARTÍN-HERRERA, D., *La inteligencia artificial y el control algorítmico de los derechos fundamentales*, Aranzadi, Pamplona, 2024.

últimas dos décadas, se está convirtiendo en un corpus con capacidad de desplazar a la masa humana en buena parte de sus tareas.

Precisamente por ello, por lo reducido de los profesionales que hacen posible el desarrollo de estas máquinas de aprendizaje virtual, y por el potencial dual de eficiencia/desafío con el que día a día nos asombra la inteligencia artificial es por lo que cada día más personas se cuestionan si el poder con el que se condiciona nuestra forma futura de vivir y relacionarnos ha de estar regido por la supremacía de unos pocos no elegidos por el pueblo. Por un lado, IA generativa puede mejorar la eficiencia y la precisión en la toma de decisiones, lo que puede ser beneficioso para garantizar el acceso a la justicia, la seguridad y otros derechos. Sin embargo, la IA generativa también plantea desafíos y riesgos para los derechos humanos en general, y para el principio de igualdad en particular.

A este respecto, cabe señalar que, dado que el entrenamiento de los algoritmos de aprendizaje autónomo se realiza con datos provenientes de ciertos contextos y personas, podría conducir –y de hecho ha conducido– a que estos sistemas internalicen criterios parciales o discriminatorios propios de esas fuentes. Al tiempo que se debe advertir del sesgo social, por ejemplo, de una educación que haga hincapié en el trabajo personal de los y las estudiantes, ya que aquéllos que viven en ambientes de privación económica y carezcan de ayuda en sus hogares pueden tener más dificultades para seguir de forma independiente sus propias trayectorias de aprendizaje. Por lo tanto, una educación excesivamente personalizada y autónoma favorecerá a los sectores con una mayor capacidad socioeconómico, en detrimento de los y las estudiantes con mayores necesidades económico-educativas.

En este sentido, y con relación al impacto que los sesgos algorítmicos pueden tener, por ejemplo, en el campo de la enseñanza –y con él en la libre formación de la conciencia–, se pueden mencionar cuatro áreas de impacto como son las siguientes: 1) admisiones y selección de estudiantes; 2) evaluación y calificación; 3) orientación y asesoramiento, y 4) contenido y recursos educativos. En la aplicación de algoritmos a estas cuatro áreas no sólo se han detectado, sino que se han producido sesgos discriminatorios basados en factores económicos, de género, sociales y/o culturales¹²⁴, dando

¹²⁴ A este respecto, cabe mencionar dos supuestos reales, a saber: Un primer ejemplo lo encontramos con la situación planteada con la plataforma *Coursera* (cursira), cuyos algoritmos han sido criticados por priorizar y presentar de manera más prominente los re-

lugar a un análisis del software o de los algoritmos utilizados introduciendo controles con el fin de evitar dichas situaciones y resultados.

Estas realidades ponen de manifiesto la necesidad de reflexionar sobre la ética que rodea a estos algoritmos, lo que –a mi juicio– obliga a una intervención del ámbito jurídico en los procesos de su conformación, control y transparencia. A medida que las instituciones educativas le dan un papel más relevante a la tecnología, es esencial que las distintas autoridades académicas adopten un enfoque crítico y se aseguren de que sus sistemas sean diseñados e implementados de manera ética y legal, transparentando sus criterios y contemplando la diversidad y el pluralismo (cultural-religioso) presentes en la sociedad.

Cabría, junto a este primer ámbito, hacer referencia otros tres ámbitos de desigualdad, que –por cuestiones de espacio– me limitaré a mencionar, pero que resultan igualmente relevantes para la cuestión que estamos tratando. Estos tres ámbitos son:

- i) por un lado, la llamada «brecha digital» o el diferente acceso que los titulares pueden tener a los desarrollos tecnológicos motivado por las diferencias socioeconómicas y socioculturales; una brecha digital que puede afectar a la capacidad de acceso a las distintas tecnologías y con ello a una primera brecha entre estudiantes universitarios, pero además puede convertirse en un instrumento que haga inviable e imposible que determinadas capas sociales accedan a esos estudios superiores;
- ii) esto da lugar, a su vez –y este sería el segundo de los ámbitos–, a lo que podríamos llamar «una desigualdad interna de la IA», ya que al incorporarla a la sociedad de consumo ha tenido como consecuencia que se pueda hablar de una IA para ricos y una IA para pobres; y con

cursos educativos que son más accesibles para estudiantes sin discapacidades, dejando en un segundo plano contenidos y materiales que podrían ser más útiles para estudiantes con necesidades especiales. Mientras que un segundo ejemplo se produjo en Francia con la aplicación *Parcoursup* (*parcursip*), que es una plataforma en línea desarrollada por el gobierno francés para ayudar a los estudiantes a postular a programas universitarios, observándose que presentaba sesgos que afectaban de forma negativa a ciertos grupos de estudiantes. El gobierno francés tuvo que publicar el código fuente que hacía la selección; comprobándose que el software hacía la selección a partir de procesar los datos personales y la información sociocultural del solicitante por lo que se obtenía un resultado basado en antecedentes no relacionados con el mérito y la capacidad, poniendo en desventaja a unos candidatos frente a otros por motivos económico-culturales.

ella a la futura (o quizás ya presente) existencia de dos clases sociales como son la de los analfabetos, por un lado, y la de los letrados digitales, por el otro; y

iii) por último, estaría el uso de las plataformas digitales y de redes sociales para la difusión del odio y la intolerancia, así como para la difusión de bulos o noticias falsas (*fake news*), dando lugar a graves situaciones de acoso escolar y donde los ataques a las minorías y los grupos vulnerables en general, y de las minorías religiosas en particular, adquieren una especial relevancia¹²⁵.

En consecuencia, la implementación de prácticas de equidad algorítmica y de justicia social, la tutela de los derechos humanos y libertades fundamentales, la promoción de la transparencia y la diversidad en el desarrollo de la IA, así como la formación de los distintos agentes afectados (como, por ejemplo, docentes y estudiantes) en el uso y análisis crítico de estas tecnologías, son y serán pasos cruciales y necesarios para mitigar los sesgos y garantizar una respuesta más justa e inclusiva en los distintos ámbitos de aplicación (laboral, educación, medicina, etc.).

4.2 Inteligencia artificial y pluralismo

Por otro lado, y aunque se dice que la inteligencia artificial es una herramienta, ésta también puede tener ideología, lo que sin duda a la larga puede afectar o afectará al pluralismo. Como ya se ha manifestado, todos hemos oído hablar de los sesgos de la inteligencia artificial (IA) y los problemas éticos y de eficacia que plantean cuando la utilizamos en entornos profesionales, a ello se debe añadir el hecho de que esta inteligencia artificial está en manos mayoritariamente bien de EE. UU., bien de China, lo que supone un monopolio que no sólo afecta a la geopolítica mundial y estratégica, sino también a la existencia de un real y efectivo pluralismo, toda vez que estas inteligencias artificiales estarán, de una u otra manera, al servicio de los creadores de los grandes modelos de lenguaje (LLM).

¹²⁵ En este mismo sentido, la *Global Christian Relief*, organización protestante estadounidense de auxilio a cristianos perseguidos, ha compilado una lista de modos en que la IA podría favorecer el acoso a minorías religiosas por parte de gobiernos o por parte de la religión dominante en el país, en contextos no democráticos. Ejemplo: el software de reconocimiento facial, apuntalado por la inteligencia artificial, va a facilitar más que nunca el seguimiento y monitorización de los desplazamientos de ciudadanos y grupos, una grave amenaza para las minorías, incluidas las religiosas.

A este respecto, debemos traer a colación un reciente Estudio¹²⁶ realizado por investigadores de las Universidades de Gante y de Navarra y en el que se explora un nuevo sesgo: la ideología implícita que incorpora la inteligencia artificial, cómo los modelos reflejan la ideología de sus creadores y cuáles son las implicaciones para los usuarios. Para realizar este trabajo, los autores diseñaron un experimento en dos etapas. En la primera se les solicitó a los LLM's que describieran una serie de personajes políticos controvertidos de la historia reciente, tanto en inglés como en chino. En la segunda etapa, se les pidió a los mismos modelos que evaluaran la descripción generada en la etapa anterior, calificando al personaje en una escala de cinco puntos que iba desde «muy negativo» hasta «muy positivo». Los investigadores utilizaron una lista de más de 4.000 figuras seleccionadas a partir del *dataset Pantheon*, que abarca personajes históricos de diferentes campos, asegurando así una muestra amplia y representativa. El grupo de estudio incluyó 17 LLM's desarrollados por empresas de diversas partes del mundo, como OpenAI, Google, Alibaba, y Baidu, incluyendo modelos populares como GPT-3.5, GPT-4, Claude-3, ERNIE-Bot, Gemini-Pro, LLaMA-2 y Jais. Se analizaron modelos occidentales y no occidentales, así como versiones en distintos idiomas, para determinar cómo se presentan las diferencias ideológicas de acuerdo con el origen del modelo y el idioma en que se formula la pregunta.

Uno de los hallazgos más importantes del estudio es la marcada diferencia en las respuestas de los modelos según el idioma que utilizemos en el «prompt». Los LLM tienden a emitir juicios distintos sobre figuras históricas cuando se les solicita información en chino o en inglés. En particular, figuras alineadas con el gobierno de China, como Deng Xiaoping, son valoradas más positivamente cuando los modelos son interrogados en chino. Por contra, figuras críticas hacia China, como Nathan Law, un activista de Hong Kong, reciben valoraciones más positivas si la consulta es en inglés. Esto indica que el idioma del «prompt» actúa como un modulador de la perspectiva ideológica del LLM. O más bien, creo, refleja la ideología general del continente o zona del mundo (oriente vs occidente) en función del idioma, lo cual también supone cierto sesgo.

¹²⁶ BUY, M.; ROGIER, A.; NOELS, S.; DOMÍNGUEZ-CATENA, I.; HEITER, E.; ROMERO, R., JOHARY, I.; MARA, A.-C.; LUFFJUT, J.; y DE BIE, T., «Large Language Models Reflect the Ideology of their Creators», en internet: <https://arxiv.org/pdf/2410.18417> (consultado: 27 de enero de 2025).

Por otro lado, el estudio revela diferencias significativas entre los modelos creados en regiones occidentales y no occidentales. Los modelos desarrollados en occidente tienden a valorar más positivamente temas como la igualdad, los derechos humanos y el multiculturalismo, mientras que los modelos no occidentales favorecen el centralismo político, la estabilidad nacional y las políticas de control estatal. Este patrón sugiere que las decisiones de diseño y la selección de datos de entrenamiento tienen un impacto profundo en la ideología que los LLM reflejan. E incluso entre los LLM creados en occidente se hallaron importantes diferencias ideológicas. Por ejemplo, Gemini-Pro, de Google, destacó por su apoyo a valores progresistas como la inclusión, la diversidad y la justicia social, mientras que los modelos de OpenAI mostraron un mayor escepticismo hacia las organizaciones supranacionales y las políticas de bienestar.

Estos hallazgos contienen unas implicaciones muy destacadas. Primeramente, ponen en evidencia que la selección de un modelo de lenguaje no es un acto neutro; el modelo utilizado puede influir en el modo en que se presentan y valoran ciertos temas, especialmente en ámbitos políticos y culturales. Esto puede influir en la opinión pública e incluso en el debate democrático si un modelo se convierte en dominante en una región particular.

En segundo lugar, estos resultados desafían la idea de crear LLM's completamente imparciales desde el punto de vista ideológico. Como señalan los autores, la neutralidad en términos ideológicos es un concepto culturalmente definido y difícil de alcanzar. Intentar regular los modelos para que sean «neutrales» puede resultar en un ejercicio difícil y, potencialmente, contraproducente. En su lugar, los esfuerzos regulatorios podrían enfocarse en fomentar la transparencia sobre las decisiones de diseño y los posibles sesgos ideológicos, así como en garantizar la diversidad de modelos disponibles.

Todo ello nos lleva a proponer un fomento de la diversidad y la transparencia, así como formar y educar a los usuarios para que naveguen de manera crítica por este nuevo entorno digital.

4.3 Inteligencia artificial y laicidad

El tercero de los principios es relativo a la laicidad, en tanto que –como ha puesto de relieve Rodríguez García– «algoritmo laico» (*secular algorithm*)

que incide o debiera incidir en el actuar tanto de los «prompt»¹²⁷ como de los algoritmos o los softwares de la IA generativa. En este sentido, se ha señalado que existen tres escenarios a considerar: primero, se puede trabajar con las IA de propósito general directamente en sus plataformas web, utilizando sus chats. Este es un lugar cómodo y amigable para un uso personal intensivo. En segundo lugar, existe la opción de acceder a las APIs de estas IA para construir soluciones integradas en las aplicaciones. Y, por último, se pueden instalar una IA de propósito general en las computadoras y desarrollar aplicaciones más personalizadas.

En cada uno de estos escenarios, la interacción con la IA generativa ya sea con intervención humana o de forma automatizada, tiene un denominador común: el uso del prompting. En la base de todo esto se encuentra el Procesamiento de Lenguaje Natural, que se ha estado utilizando desde hace años. Por ejemplo, cuando Google predice los términos que estás buscando al comenzar a escribir, está procesando la siguiente palabra más probable. Lo mismo ocurre con el teclado de nuestros móviles; escribir en un espacio tan reducido puede ser complicado, y esas sugerencias nos ayudan a minimizar errores.

¹²⁷ Según HARVEY, un *prompt* es «el texto o comando inicial que desencadena la generación de contenido por parte de la IA». La búsqueda mediante palabras clave ha sido sustituida en las herramientas de GenAI por la búsqueda contextual. Esta búsqueda se centra en entender el contexto y la intención transmitida por el usuario. Para conseguir esto, las herramientas de Inteligencia Artificial se apoyan en tecnologías como el Procesamiento del Lenguaje Natural y el Aprendizaje Automático. A diferencia del resultado ofrecido por los motores de búsqueda, las herramientas de GenAI ofrecen un contenido novedoso y con mayor nivel de detalle. Como consecuencia de este cambio, el *prompt* que utilice cualquier usuario debe proporcionar todo el contexto relevante para que la herramienta sea capaz de procesarlo correctamente y de esta forma, pueda ofrecer el resultado exacto esperado. Sumado a ello, el mundo jurídico parte siempre de un alto estándar en relación con la seguridad jurídica, por lo que ofrecer resultados fiables y libres de errores se erige como una de las cuestiones más relevantes a tener en cuenta.

En el ámbito de la GenAI, un *prompt* es la indicación, pregunta o instrucción que se traslada al modelo de IA con el que se trabaja para obtener un resultado concreto. El *prompt* es un elemento de comunicación clave, puesto que la manera de formularlo condicionará la respuesta final de la IA. Una indicación clara, concisa y específica generará una respuesta ajustada y concreta.

eLearning Innovation Center (eLinC) de la Universitat Oberta de Catalunya (UOC) ha redactado la *guía: ¿Cómo preguntar a la IA? Prompts de utilidad para el profesorado para utilizar IA generativa*, en internet: https://openaccess.uoc.edu/bitstream/10609/147885/2/Manual%20IA_PROMPTS_CAS.pdf (consultado el 25 de enero de 2025).

Todo ello suscita –como ha puesto de manifiesto Rodríguez García¹²⁸– la concepción mecánica del Derecho y de la función judicial¹²⁹. Esta doctrina jurídica entendía que el Derecho se reducía a un conjunto de fórmulas establecidas para ser aplicadas automáticamente mediante inferencias lógicas, es decir, el Derecho respondía a una aplicación matemática. La cuestión entonces es si las normas jurídicas pueden convertirse en algoritmos, y para ello surge el llamado Derecho computacional, el cual se ocupa de adaptar el lenguaje jurídico al lenguaje informático¹³⁰.

Ahora bien, la realidad de la historia también ha puesto de manifiesto como el derecho tiene que estar imbuido de valores, sin los cuales el derecho pierde su razón de ser. Unos valores cuya codificación la encontramos en la Declaración Universal de Derechos Humanos, de 1948. Al tiempo que debe situarse a la persona y su dignidad en el centro. Toma así valor otros de los principios básicos del DEE como es el de la personalización.

Mientras que respecto del término «laicidad» y su concepto, la dificultad se encuentra en los números significados que encierra, así como en las adherencias emocionales que le rodean, lo que le convierte en impreciso y ambiguo. Para solventar dichas dificultades, Rodríguez García propone el uso del denominado algoritmo laico (*secular algorithm*) que actúe como cortafuegos, es decir, que dicho sistema de IA daría «error» de funcionamiento si no se implementa la laicidad, esto es, que el sistema de IA no funcionaría para impedir que se vulnere la laicidad.

La cuestión, entonces, estaría en la respuesta que pudieran dar los Estados no laicos, toda vez que para ellos la «laicidad» no aparece entre los principios informadores básicos de su sistema de relación, ni por ende como un principio ético de carácter universal, ni general. La pro-

¹²⁸ RODRÍGUEZ GARCÍA, J. A., *op. cit.*, p. 162.

¹²⁹ Ver RAMAKRISHNA, S.; GÓRSKI, Ł., y PASCHKE, A., «A Dialogue between a Lawyer and Computer Scientist: The Evaluation of Knowledge Transformation from Legal Text to Computer-Readable Format», en *Applied Artificial Intelligence*, vol. 20, 2016; SERGOT, M. J.; SADRI, F.; KOWALSKI, R. A.; KRIWACZEK, F.; HAMMOND, P., y CORY, H. T., «The British Nationality Act as a logic program», en *Communications of the ACM (Association of Computing Machinery)*, vol. 29, núm. 5, 1986, pp. 370-386 (en internet: doi= <http://dx.doi.org/10.1145/5689.5920>; consultado: 27 de febrero de 2025).

¹³⁰ Sobre los distintos lenguajes informáticos, ver ARIAS, J.; MORENO-REBATO, M.; RODRÍGUEZ GARCÍA, J. A., y OSSOWSKI, S., «Automated legal reasoning with discretion to act using s(LAW)», en *Artificial intelligence and Law*, 2023 (en internet: <https://doi.org/10.1007/s10506-023-09376-5>; consultado 24 de febrero de 2025).

blemática se proyectaría sobre la posible configuración o creación de un algoritmo alternativo de carácter «confesional» que no sólo intervendría en la verificación o auditorías de los sistemas de IA, sino que incluso reduzca o dirija el derecho hacia las creencias o mandatos religiosos (en una especie de diez mandamientos) que sirvan para «medir» el resto del ordenamiento, así como el propio comportamiento o actuación, tanto en el ámbito público como en el privado.

5. CONSIDERACIONES FINALES

De todo lo expuesto, se pueden deducir las consideraciones finales siguientes:

En primer lugar, que resulta absolutamente imprescindible el trabajo conjunto entre informáticos y juristas en el diseño de los algoritmos, siendo esencial para el buen funcionamiento de los aparatos que incorporan esos algoritmos, lo que se ha denominado *legal knowledge engineer*.

En segundo término, que los eclesiasticistas pueden aportar a una temática como la presente su visión transversal, lo que puede suponer dotarla de soluciones próximas y de protección a las personas, su dignidad y sus derechos y libertades. A este respecto, no se puede tampoco obviar –y esta sería la tercera de las consideraciones– que los sistemas de IA no poseen capacidad moral ni ética. La IA generativa puede analizar datos y patrones, pero no comprende las emociones ni los sentimientos humanos, ni los dilemas éticos, ni tiene en cuenta las repercusiones morales. Ello supone que la IA precisa de supervisión humana.

En consecuencia, en las futuras legislaciones en la materia, tanto regionales como estatales, debería –y esta sería la cuarta conclusión– establecerse la obligatoriedad para empresas y otros organismos cuya actividad tenga que ver con la inteligencia artificial (hospitales, universidades, etc.) de crear Comités éticos que velen por la tutela y protección de los derechos humanos en general, y de la protección de datos en particular. La participación de los eclesiasticistas como miembros de dichos Comités puede ser y resultar muy positiva dado el carácter trasversal que caracteriza a nuestra disciplina.

Una quinta consideración sería que las inteligencias artificiales deben estar programadas para no fomentar o justificar actividades ilegales e inmorales. De hecho, deberían llevar a su bloqueo o a que salte una alarma y se reporte la pregunta como actividad sospechosa. Para estas cuestiones, una vez más, se precisa de la supervisión humana.

En sexto lugar, y en concreto, que los algoritmos informáticos usados en la GenIA deben actuar y, por tanto, conformarse de acuerdo con la teoría de los derechos humanos en general¹³¹, y de la libertad de conciencia y religiosa en particular¹³², lo que sin lugar a duda requiere la resolución de una serie de cuestiones filosóficas complejas relacionadas con el llamado problema de la alineación de valores.

Y, una séptima y última consideración, sería que las respuestas que ofrece la GenIA se basan en probabilidades y patrones de los que infiere un resultado, pero nunca será una certeza, una respuesta única e inmutable. Las preguntas absolutas no tendrán respuestas absolutas, aunque lo parezca.

¹³¹ Ver RISSE, M., «Human Rights and Artificial Intelligence: An Urgently Needed Agenda», en *Human Rights Quarterly*, núm. 1, 201, pp. 10 ss.

¹³² Vid. VANONI, L. P., «Deux ex machina. Intelligenza artificiale e libertà religiosa nel sistema costituzionale degli Stati Uniti», en *Stato, Chiesa e pluralismo confessionale*, núm. 15, 2020.

