



EMPLEO DE LA INTELIGENCIA ARTIFICIAL Y CONSENTIMIENTO INFORMADO EN LA ASISTENCIA SANITARIA

AI'S USE IN HEALTH CARE AND INFORMED CONSENT

CAROLINA PEREIRA SÁEZ <https://orcid.org/0000-0001-8439-2352>

Universidad de A Coruña

Facultad de Derecho, Campus de Elviña, s/n, 15071, A Coruña. 981 167000 ext. 1640

c.pereira.saez@udc.es

RESUMEN:

Palabras clave:

IA, consentimiento
informado, asistencia
sanitaria

Recibido: 12/07/2025

Aceptado: 24/08/2025

La trascendencia del empleo de la inteligencia artificial (IA) en muchos ámbitos de la existencia humana apremia a reflexionar sobre su relevancia ética. En este trabajo se atiende al significado que su empleo en la asistencia sanitaria tiene para el consentimiento informado del paciente. Para ello se propone, primero, una comprensión de la IA, así como del fundamento del consentimiento informado. A continuación se atiende brevemente a una cuestión fundamental: la condición libre de la persona humana, sujeto tanto de la ética como del Derecho. La conclusión que se propone es que un empleo de la IA respetuoso con la condición libre de la persona exige entenderla siempre como herramienta, en este caso al servicio del cuidado de la salud. Para que esto no sea una afirmación hueca ha de traducirse en la desmitificación de la IA: es preciso considerarla y emplearla no como auténtico sujeto, sustituto de la persona, sino producto o desarrollo humano, teniendo en cuenta que, en último término, es el ser humano quien la diseña y decide usarla y cómo. Una medida que puede ayudar a evitar la personificación de la IA es asegurar la intervención humana ya en las primeras fases del desarrollo de los sistemas de IA. Tratar de resolver las incertidumbres éticas que plantea la IA exigiendo la intervención humana cuando el sistema está ya desarrollado y tomando decisiones en la vida real puede ser poco más que un sello que pretende avalar lo que no puede.

ABSTRACT:

Keywords:

AI, Informed
Consent, Health Care

The implications of the use of artificial intelligence (AI) in many areas of human existence compels us to reflect on its ethical relevance. This paper addresses the signification of its use in healthcare for patient informed consent. To this end, it first proposes an understanding of AI, as well as the basis for informed consent. Next, a fundamental question is briefly addressed: the free condition of the human person, subject both of ethics and Law. The conclusion proposed is that a use of AI respectful to the free nature of persons requires understanding it always as a tool –in this case, in the service of healthcare. To ensure this is not a hollow assertion, it must be translated into a demystification of AI: It is necessary to consider and use it not as an authentic subject, a substitute for the person, but as a product or human development, taking into account that, ultimately, it is the human being who designs it and decides to use it and how. One measure that can help prevent AI personification is to ensure human intervention already in the early stages of AI system development. Trying to solve the ethical uncertainties posed by AI by requiring human intervention when the AI system is already developed and making decisions in the real life may be little more than a rubber stamp that purports to endorse what it cannot.

1. Introducción: la trascendencia ética de la IA

La finalidad de este trabajo es reflexionar brevemente sobre cómo afecta el empleo de sistemas de inteligencia artificial (IA) en la asistencia sanitaria al consentimiento informado del paciente. Para ello se va a considerar, primero, qué es lo que caracteriza a la IA a estos efectos. En segundo lugar, se va a atender al sentido del consentimiento informado en el ámbito de la asistencia sanitaria. Se propone que desde esa perspectiva puede ofrecerse un criterio para pautar el empleo de sistemas de IA que explique el sentido de las recomendaciones concretas más frecuentes –principalmente las exigencias de explicabilidad y supervisión humana¹. Así, se va a proponer que un uso de la IA adecuado al fundamento del consentimiento informado es el que tiene en cuenta las exigencias de la condición libre del ser humano y el carácter interpersonal de la relación entre profesional asistencial y paciente.

Esta reflexión se encuadra, en realidad, en otra, más amplia, que excede el objeto de este trabajo pero que ha de tenerse en cuenta: la que considera los riesgos y los desafíos éticos que plantea el desarrollo y empleo de sistemas de IA, no sólo en el ámbito de la asistencia sanitaria, sino en general, y propone pautas éticas que dirijan ese empleo con el objeto de lograr un uso de la IA centrado en el ser humano y orientado al bien común. Es decir, el punto de partida es que los desarrollos de IA no son una mera técnica éticamente neutral y que, por lo tanto, no sólo su uso, sino su mismo desarrollo requiere unas pautas éticas.

Así, son ya muchos los organismos políticos, legisladores, comités y expertos en ética, asociaciones, empresas y entidades de muy diversa naturaleza que entienden que el empleo de sistemas de IA conlleva riesgos serios y por lo tanto plantea desafíos éticos, en respuesta a los cuales han propuesto pautas éticas y jurídicas para

el empleo de la IA². Entre los que suelen identificarse como riesgos con trascendencia ética, encontramos, por lo menos, los siguientes: a) discriminación, sesgos y problemas de justicia; b) concentración de poder; c) privacidad y datos; d) dominio de intereses económicos de las grandes compañías tecnológicas; e) vigilancia y control por parte de poderes públicos y pérdida de libertades; f) opacidad, problema de las “cajas negras” y falta de explicabilidad; g) incertidumbre acerca de la responsabilidad en la toma de decisiones y por daños; h) seguridad, acierto y solidez técnica de los sistemas; i) aplicación de criterios eugenésicos en el ámbito de la vida humana; j) reducción de lo humano a lo técnico: de la inteligencia humana a la artificial, del bien humano a lo cuantificable, de los problemas humanos a los problemas técnicos, de la decisión libre a la estimación algorítmica; k) pérdida de control humano significativo por disminución de la atención y pérdida de cualificación (exceso de delegación); l) problemas medioambientales; m) afectación a la crisis de la verdad y el debate público; n) afectación a la democracia; ñ) pérdida de puestos de trabajo; o) ciberseguridad; p) problemas de seguridad en aplicaciones a ámbitos críticos³.

Son, en realidad, riesgos de muy distinta naturaleza: en algunos casos, se refieren a daños colaterales no intencionados, como los daños para el medio ambiente; en otros, a daños derivados de un uso indebido, como

2 Para una posible aproximación a los desafíos éticos que tiene en cuenta, además, las más relevantes del vastísimo número de propuestas atendidas por escuelas éticas, puede verse Gaudet, M., Herzfeld, N., Scherz, P. y Wales, J. (eds.), *Encountering Artificial Intelligence. Ethical & Anthropological Investigations*, Pickwick, Eugene, 2024, 24-40 y bibliografía ahí citada. Como ejemplo de legislación puede verse, en el ámbito europeo, el Reglamento UE 2024/1689, de 13 de junio de 2024, que se propone “mejorar el funcionamiento del mercado interior y promover la adopción de una inteligencia artificial (IA) centrada en el ser humano y fiable, garantizando al mismo tiempo un elevado nivel de protección de la salud, la seguridad y los derechos fundamentales consagrados en la Carta, incluidos la democracia, el Estado de Derecho y la protección del medio ambiente, frente a los efectos perjudiciales de los sistemas de IA (en lo sucesivo, “sistemas de IA”) en la Unión así como prestar apoyo a la innovación” (cfr. art. 1).

3 Algunos ejemplos de propuestas que identifican riesgos son el Reglamento UE 2024/1689, de 13 de junio de 2024; la *Recomendación sobre la ética de la inteligencia artificial* de la UNESCO (23.11.2021); las *Directrices éticas para una IA fiable* del grupo independiente de expertos de alto nivel sobre inteligencia artificial creado por la Comisión Europea en junio de 2018 (Bruselas, abril de 2019) o el *Libro blanco sobre la inteligencia artificial. Un enfoque europeo orientado a la excelencia y la confianza*, Comisión Europea, Bruselas (19.02.2020).

1 Cfr., por ejemplo, Comitato Nazionale per la Bioética y Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita, *Parere congiunto Intelligenza artificiale e medicina: aspetti etici*, Roma, 2020, 17 y Mittelstadt, B., *The Impact of Artificial Intelligence on the Doctor-Patient Relationship. Report Commissioned by the Steering Committee for Humans Rights in the fields of Biomedicine and Health (CDBIO)*, Bruselas, Consejo de Europa, 2021, 46-49 y 57-60.

ataques y problemas de ciberseguridad; algunos son problemas técnicos, como la opacidad; otros, sociales y políticos, como la afectación a los puestos de trabajo o a la democracia. Para atenderlos se hacen propuestas, también distintas, de pautas éticas: muchas, de inspiración principialista; otras, elaboradas a partir de las exigencias de los derechos humanos, o como desarrollos de distintas escuelas éticas, como la de la ética de la virtud. Pues bien, como se decía, este contexto excede, con mucho, el objeto de este trabajo; sin embargo, conviene tenerlo en cuenta para recordar que esas pautas éticas se proponen a técnicos, programadores, desarrolladores, políticos, usuarios, etc.; es decir, a las personas: sólo el ser humano es sujeto moral, un ser libre y responsable que ejerce su libertad en sus decisiones y acepta las consecuencias de las mismas. Los sistemas de IA son siempre el resultado de procesos y decisiones humanas⁴.

En el contexto particular del cuidado de la salud encontramos también muchas propuestas que atienden a los desafíos éticos del empleo de sistemas de IA y proponen pautas éticas que lo orienten hacia la atención al ser humano y la contribución al bien común⁵. Como sabemos, los beneficios para la salud que pueden llegar a obtenerse empleando sistemas de IA en la asistencia sanitaria y el cuidado de la salud en general parecen significativos⁶, tanto en lo que se refiere a la gestión

pública como en lo que se refiere a la atención personal. Para algunos éste es, de hecho, uno de los ámbitos en los que las posibilidades de que el empleo de sistemas de IA incrementen la eficiencia de los resultados parecen mayores, en tareas que van desde la asistencia al diagnóstico de ciertas enfermedades hasta la prevención, el desarrollo de fármacos o el seguimiento de enfermedades crónicas, por ejemplo.

Al mismo tiempo se es consciente, también, de los riesgos que entraña el empleo de sistemas de IA en este ámbito. Podemos encontrar distintos catálogos de riesgos pero, a grandes rasgos, resultan básicamente coincidentes; Mittelstadt, por ejemplo, los estudia en el ámbito de la relación médico-paciente agrupándolos entorno a seis grandes temas: 1) desigualdad en el acceso a la atención sanitaria de alta calidad; 2) transparencia, tanto para los profesionales como para los pacientes; 3) riesgo de sesgos sociales en los sistemas de IA; 4) disolución de la comprensión del bienestar del paciente; 5) sesgo de automatización o riesgo de confiar automáticamente o en exceso en los resultados generados por sistemas de IA, descalificación y desplazamiento de la responsabilidad; 6) impacto en el derecho a la intimidad. En cualquier caso, sean estos u otros los riesgos previstos, parece claro que el empleo de sistemas de IA en el ámbito de la salud no sitúa en un contexto especialmente delicado, por la afectación que puede suponer para la vida de las personas. Así, por ejemplo, el Reglamento UE 2024/1689, de 13 de junio de 2024, clasifica el uso IA en este ámbito como de alto riesgo, por lo que exige el cumplimiento de unos requisitos específicos, como la incorporación de mecanismos de reducción de riesgos, el empleo de datos de alta calidad, la obligación de ofrecer información clara al usuario, la garantía de supervisión humana o la evaluación del impacto sobre los derechos humanos previa a la implantación⁷.

4 Cfr. Lehr, D. y Ohm, P. «Playing with the Data: What Legal Scholars Should Learn About Machine Learning». *UC Davis Law Review*. 2017; 51: 653-717, 715; cfr., también, Dicasterio para la doctrina de la fe y Dicasterio para la cultura y la educación, *Antiqua et Nova. Nota sobre la relación entre la inteligencia artificial y la inteligencia humana* (28.05.2025), § 39.

5 Por hacer referencia sólo a algunos, Solomonides, A. et alii. «AMIA Position Paper. Defining AMIA's artificial intelligence principles». *Journal of the American Medical Informatics Association*. 2022; 29(4): 585-591; OMS, *Ética y gobernanza de la inteligencia artificial en el ámbito de la salud: orientaciones de la OMS*, Ginebra, 2021; Mittelstadt, B., *op. cit.*, que atiende no a todas las posibilidades de empleo de sistemas de IA en el cuidado de la salud sino sólo a lo que afecta a la relación médico-paciente; Comitato Nazionale per la Bioética y Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita, *op. cit.*; Línea de trabajo de Transformación digital de DigitalEs, *Inteligencia artificial ética en sanidad. Recomendaciones para la adopción de IA respetuosa, transparente, segura y justa*, DigitalEs, 2022.

6 Cfr. los documentos citados en la nota 5. Hay interpretaciones entusiastas de las posibilidades de la IA en el ámbito del cuidado de la salud y otras más escépticas. Para una breve descripción de posibles aplicaciones de sistemas de IA en la asistencia sanitaria puede verse, por ejemplo, Russell, S. y Norvig, P., *Artificial Intelligence. A Modern Approach* (4ª ed.), Londres, Pearson, 2022, 48; cfr. también Comitato Nazionale per la Bioética y Comitato Nazionale

per la Biosicurezza, le Biotecnologie e le Scienze della Vita, *op. cit.* 7-9; Mittelstadt, B., *op.cit.* 29-36. Para una visión optimista, que considera que la aplicación de la IA en la asistencia sanitaria permitirá una asistencia, en realidad, más humana, cfr. Topol, E., *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*, Nueva York, Basic Books, 2019. Para una revisión crítica, que lo considera un planteamiento ingenuo, cfr. Sparrow, R. y Hatherley, J., «High hopes for "Deep Medicine"? AI, Economics and the Future of Care». *Hastings Center Report*. 2020; 50(1): 14-17.

7 Cfr. Reglamento UE 2024/1689, de 13 de junio de 2024, capítulo III, secciones primera y segunda.

Como vemos, nos encontramos, otra vez, ante riesgos de distinta naturaleza, con distintas respuestas. La atención a los sesgos sociales y el riesgo de discriminación es muy frecuente; los riesgos que tienen que ver con la privacidad y el uso de datos se atienden en detalle en normas y recomendaciones específicas⁸, porque el desarrollo de sistemas de IA necesita grandes cantidades de datos de calidad, y el empleo de historias clínicas digitales y de sistemas de IA en la asistencia sanitaria pueden facilitar mucho el acceso a esos datos. Además, para pronunciarse sobre los criterios éticos para manejar esos datos han de tenerse en cuenta muchas cuestiones: la naturaleza de los datos y la vinculación con el sujeto, el beneficio económico resultante de su uso, la contribución al bien común que su empleo para la investigación pueda suponer, la eficacia de los sistemas de pseudo-anonimización y las posibilidades reales de re-identificación, por ejemplo. En otros casos, se trata de riesgos especialmente significativos para el Derecho, como los que se refieren al desplazamiento de la responsabilidad; aunque aquí, para algunos, la adopción previa de normas claras puede ser suficiente. Por otro lado, parecen riesgos especialmente significativos para el cuidado de la salud los que tienen que ver con el exceso de confianza en sistemas automatizados y la pérdida de cualificación en el profesional: frente a temores más drásticos, pero más improbables, de dominio de los sistemas de IA sobre las personas, parece más realista el riesgo de exceso de confianza en la IA, dada su complejidad o mayor objetividad; una suerte de complejo de inferioridad respecto de la máquina –o de abandono–, que lleva a descuidar la preparación humana respecto de lo que la máquina hace mejor.

Vale la pena, por tanto, estudiar con detalle los distintos riesgos que implica el empleo de sistemas de IA en la asistencia sanitaria, por distintas razones. Sin embargo, por lo que se refiere al objeto de este trabajo, el consentimiento informado del paciente, el riesgo principal parece ser el que tiene que ver con la falta de transparencia u opacidad de los sistemas de IA, en la medida

⁸ A esta cuestión atiende el Reglamento UE 2025/327, de 11 de febrero de 2025, relativo al Espacio Europeo de Datos de Salud.

en que afectará a la información necesaria para prestar un consentimiento válido. Éste es, por otro lado, uno de los riesgos más característico de los sistemas de IA: estamos habituados a que los desarrollos tecnológicos tengan efectos negativos no previstos o que se usen para fines perniciosos; sabemos que suponen, directa o indirectamente, una forma de poder; somos conscientes del alcance que puede tener la generalización de sesgos o criterios discriminatorios. Sin embargo, nos resulta más alarmante la idea de aceptar las propuestas de un sistema artificial cuando no terminamos de entender cómo las alcanza –lo que lleva a algunos a afirmar que se trata más de arte que ciencia, y a mitificar los sistemas de IA⁹.

Pues bien, como decíamos, en este trabajo nos vamos a limitar a reflexionar brevemente sobre la trascendencia que puede tener el empleo de sistemas de IA en la asistencia sanitaria para el consentimiento informado del paciente. Somos conscientes de la amplitud de los términos de esta relación: por un lado, la inteligencia artificial, de no fácil delimitación, hoy omnipresente, causante tanto de entusiasmo¹⁰ como de inquietud¹¹, generadora, en cualquier caso, de incógnitas, con serias áreas problemáticas¹². Por otro, el consentimiento informado, que, por lo menos en España¹³, a la vista de los últimos desarrollos legales y jurisprudenciales de su fundamento constitucional (la inte-

⁹ Cfr. Lehr, D. y Ohm, P. *op. cit.* 716 y *passim*: todo el trabajo tiene por objeto desmitificar la IA y “descorrer la cortina” que oculta el intenso y complicado trabajo humano de científicos de datos, analistas y programadores que está detrás de lo que puede parecer “modelos mágicos de funcionamiento surgidos de las misteriosas profundidades de una computadora” (717).

¹⁰ “Primero resolvió la IA, después usó la IA para resolver cualquier otra cosa” (Demis Hassabis, CEO de Google DeepMind, según refieren Russell, S. y Norvig, P., *op. cit.* 49. En otro sentido, mucho más trascendente, Kurzweil, R., *La singularidad está más cerca. Cuando nos fusionamos con la IA*, (trad. de Alexandre Casanovas) Barcelona, Deusto, 2025. De todas maneras, también la confianza en la IA en el primero de los sentidos implica que, en último término, los problemas humanos se entienden como problemas técnicos, para los que hay técnicamente una mejor opción posible alcanzable por un sistema de IA.

¹¹ Cfr. Russell, S. y Norvig, P., *op. cit.* 51.

¹² Es interesante la distinción que se hace en Gaudet, M., Herzfeld, N., Scherz, P. y Wales, J. (eds.), *op. cit.*, entre problemas antropológicos, que en muchas ocasiones se refieren a una situación futura de la IA, y cuestiones éticas concretas con las que nos encontramos hoy en los contextos particulares en los que se aplican sistemas de IA.

¹³ Aunque, por lo que se refiere a la noción de autodeterminación en el Derecho, no solo en España, como pone de manifiesto el desarrollo en las últimas décadas de nuevos derechos de autodeterminación: cfr. Cartabía, M., ««La edad de los nuevos derechos». *Revista de Derecho Político*. 2011; 81: 61-100, *passim*.

gridad física y moral del art. 15 de la Constitución española) parece entenderse como herramienta de un amplísimo, casi ilimitado, derecho de autodeterminación¹⁴. Por este motivo, a continuación vamos a procurar acotar los términos de esta relación, tratando de ofrecer una comprensión apropiada por lo menos a los efectos de este trabajo.

2. ¿Qué inteligencia artificial para la asistencia sanitaria?

Es conveniente, entonces, especificar a qué nos referimos aquí con la expresión “inteligencia artificial” cuando la aplicamos a la asistencia sanitaria. No es este el lugar para tratar con detalle el significado de esta expresión, pero es necesario saber qué se entiende por IA para poder reflexionar sobre la trascendencia que su intervención en la relación médico-paciente pueda tener para el consentimiento informado de este último.

La definición de IA no es fácil, ni se pretende aquí abordarla. En aquella reunión científica del verano de 1956 que tuvo lugar en Dartmouth College, en Hanover, New Hampshire, John McCarthy, Marvin L. Minsky, Nathaniel Rochester y Claude E. Shannon se proponían estudiar la inteligencia artificial a partir del presupuesto de que todos los aspectos del aprendizaje y otros rasgos de la inteligencia podrían, en principio, describirse de manera tan precisa que se pudiese hacer una máquina capaz de simularlos¹⁵. Así, es frecuente entender la IA a partir de la idea de que es posible que una máquina pueda mostrar un comportamiento que se consideraría inteligente de ser una persona quien lo realizase; esto es, de que es posible desarrollar sistemas o máquinas capaces de imitar comportamiento humano inteligente. En realidad, esta comprensión de la IA es muy amplia, además de dependiente de qué se entienda por inteligencia, por lo que puede ser difícil de delimitar con precisión en contextos particulares; por otro lado, no distingue con precisión IA de aprendizaje automático¹⁶,

ni atiende a otras diferencias dentro del amplio campo de la IA. Tiene, sin embargo, la ventaja de reflejar con más flexibilidad que comprensiones más técnicas la percepción social de la IA, además de permitir incorporar sistemas o modelos nuevos que puedan ir desarrollándose. Esto puede ser especialmente útil en el contexto jurídico-político; así, por ejemplo, el Reglamento Europeo sobre inteligencia artificial sigue ofreciendo una definición de IA relativamente amplia¹⁷.

En cualquier caso, esta comprensión es interesante, además, por su perspectiva funcional: la inteligencia de los sistemas de IA se entiende por su capacidad para ofrecer una respuesta adecuada, asociada a la razón humana, con independencia de cómo se generen esas respuestas¹⁸. Por tanto, aunque se emplee la expresión “inteligencia” para aplicarla a un artefacto, no se pretende necesariamente sostener que la máquina “piense” en sentido humano¹⁹. Sí simula ciertas expresiones de la inteligencia humana²⁰, aunque su contexto original,

17 Cfr., por ejemplo, el art. 3 del Reglamento UE 2024/1689, de 13 de junio de 2024: “ Sistema de IA: un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos y virtuales”.

18 Si tenemos en cuenta los criterios de clasificación que proponen Russel y Norvig respecto de las posibles formas de entender la IA (entenderla por referencia al ser humano o por referencia a un estándar de racionalidad, por un lado, y entenderla por referencia al proceso interno o entenderla por referencia a su caracterización externa, por otro) resultarían cuatro modelos posibles de IA, según se entienda que es inteligente en el sentido de que: a) piensa como un ser humano; b) actúa como un ser humano; c) piensa racionalmente, o d) actúa racionalmente. A su juicio, el enfoque más extendido de la IA hoy es el que la entiende desde la perspectiva d), la perspectiva que estos autores llaman “del agente racional”: cfr. Russell, S. y Norvig, P., *op. cit.* 19-22. Se trata, entonces, no de pronunciarse sobre si la máquina “piensa” o no (entendamos ese pensar por comparación con el pensamiento del ser humano o lo entendamos por comparación con un proceso de racionalidad o corrección abstracta) sino de si actúa u obtiene un resultado “inteligente”, en el sentido de correcto o acertado.

19 Sobre esta cuestión puede verse Dicasterio para la doctrina de la fe y Dicasterio para la cultura y la educación, *op. cit.*, *passim*.

20 Hace algunas décadas se propuso distinguir entre IA débil, para referirse a la IA que podía funcionar como si pensase, y la IA fuerte, para referirse a máquinas que no solo simulan pensar sino que efectiva y conscientemente piensa. Hoy, sin embargo, lo más frecuente es entender la IA fuerte no por referencia a ese “efectivamente, pensar” sino como la IA de nivel humano o general, capaz de resolver una muy amplia variedad de tareas tan bien como una persona. La cuestión de si efectivamente piensa o no es, para muchos investigadores en IA, tan poco relevante como la cuestión de si un submarino nada: lo importante es la proximidad funcional (cfr. Russell, S. y Norvig, P., *op. cit.* 1032-1037). A mi juicio, lo relevante

14 Cfr. STC 19/2023, de 22 de marzo, especialmente FJ. 6 y voto particular de la magistrada doña María Luisa Balaguer Callejón.

15 McCarthy, J. Minsky, M., Rochester, N. y Shannon, C. «A Proposal for the Dartmouth Summer research project on Artificial intelligence, August 31, 1955». *AI Magazine*. 2006; 27(4): 12-14.

16 Cfr. Russell, S. y Norvig, P., *op. cit.* 19, nota 1. Cfr. también Mittelstadt, B., *op. cit.* 11.

lógico-matemático, le imponga condiciones; sin dejar de ser, entonces, antes que una forma artificial de inteligencia, un artificio de la inteligencia humana²¹. Sus elementos básicos son los datos y los algoritmos, secuencias de instrucciones expresadas matemáticamente, que le permiten, a partir de esa información almacenada y archivada, descubrir relaciones o conexiones ocultas²². Los datos son de una importancia radical: “un algoritmo es, al fin y al cabo, tan bueno como sean sus datos”²³. Así, estos dos elementos, algoritmos y datos, matizan el carácter autónomo de la “toma de decisiones” de los sistemas de IA.

Como se decía, no vamos a tratar de ofrecer una definición más precisa, pero sí es necesario atender a los aspectos de los sistemas de IA más relevantes para el consentimiento informado del paciente cuando se usan en la asistencia sanitaria. Por eso es necesario tener en cuenta que, aunque en sentido propio las diferencias entre inteligencia humana e inteligencia artificial sean significativas²⁴, desde una perspectiva funcional, y según el ámbito al que nos estemos refiriendo, pueden no serlo: los complejos sistemas actuales de IA pueden llevar a cabo tareas muy sofisticadas en ámbitos determinados. Los que usan aprendizaje automático (*machine learning*) tienen la capacidad de “aprender” y actuar en función de lo aprendido, de manera que, con el tiempo y a partir de la exposición a datos, modifican su comportamiento: por ejemplo, pueden revisar una recomendación basándose en patrones de datos que están evolucionando²⁵. Se trata de sistemas que pueden percibir su entorno y por eso “aprenden” a través de la “experiencia” y, a

partir de ahí, actúan maximizando sus posibilidades de alcanzar su fines²⁶. Mediante un proceso automatizado de descubrimiento de correlaciones entre variables en un conjunto de datos, puede hacer predicciones o estimar un resultado; esto es, “tomar una decisión”²⁷. Para ello operan no limitados por una lógica explícita, de manera transparente, sino que, a partir de sus marcos matemáticos y utilizando ingentes cantidades de datos, son capaces de encontrar asociaciones u obtener resultados más allá de lo que se les había programado.

Pues bien, es aquí donde nos encontramos, como se adelantaba, con el rasgo de la IA que parece más significativo a los efectos de este trabajo: la posibilidad de que el resultado que ofrezca el sistema de IA vaya más allá de la previsión de sus programadores y que resulte imposible o enormemente difícil para el ser humano conocer las razones de las estimaciones del sistema de aprendizaje automatizado²⁸. Esto es, no sólo es que vaya más allá de lo previsto por su programador, sino que, además, los pasos a través de los cuales interpreta los datos no siempre son transparentes. Nos encontramos, así, con el problema de las cajas negras o de la opacidad o falta de explicabilidad de los sistemas de IA que emplean aprendizaje automático, típicamente identificado como riesgo de la IA tanto en general como, en particular, para el ámbito del cuidado de la salud²⁹.

Es posible, sin embargo, si atendemos, por ejemplo, a Lehr y Ohm³⁰, que sea un error considerar los sistemas de aprendizaje automatizado como sistemas monolíticos, de una sola pieza; cajas negras completamente formadas e inescrutables —a juicio de estos autores, error típico de los juristas³¹. Estos autores proponen, por el contrario, tener en cuenta que nos encontramos ante un producto de una serie de procesos humanos, por lo que resulta po-

para el ámbito de la IA sería, en el ejemplo del submarino, que es el ser humano quien le hace nadar (entendiendo nadar en sentido funcional).

21 Cfr. Dicasterio para la doctrina de la fe y Dicasterio para la cultura y la educación, *op. cit.* §§ 10-12.

22 Cfr. Comitato Nazionale per la Bioética y Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita, *op. cit.* 5-6.

23 Cfr. Lehr, D. y Ohm, P. *op. cit.* 653-717, 677 y 712.

24 Sin ánimo de entrar aquí en la comparación entre IA e inteligencia humana, sí parece que se puede sostener que los sistemas de IA no comprenden, no tienen intimidad subjetiva ni conciencia de la propia realidad; serían más parecidos a inmensísimas memorias, gestores de datos súper eficientes, que a partir de esa gestión pueden calcular o predecir, en determinados ámbitos, en un nivel superior al humano.

25 Cfr. Solomonides, A. *et alii.* *op. cit.* 586.

26 Cfr. Mittelstadt, B., *op. cit.* 10-11.

27 Cfr. Lehr, D. y Ohm, P. *op. cit.* 671.

28 Cfr. *ibid.*, 705-706.

29 Mittelstadt expone con detalle y claridad este riesgo en la relación médico-paciente, distinguiendo interpretabilidad, transparencia y trazabilidad y distinguiendo también la perspectiva del experto de la del usuario no especializado. La opacidad puede deberse a la inaccesibilidad para el usuario (profesional, paciente) o a la condición del sistema en cuestión de ser fundamentalmente inescrutable o ininteligible: cfr. Mittelstadt, B., *op. cit.* 46-49 y 57-60.

30 Cfr. Lehr, D. y Ohm, P. *op. cit. passim.*

31 Cfr. *ibid.*, 655-656.

sible ganar en explicabilidad³². También para Mittelstadt la explicabilidad resulta no una cuestión de blanco o negro, sino de grados y que debe atenderse en función de la persona (técnico, usuario experto, usuario no experto, etc.); así, hay alternativas para atender a sus exigencias, aunque no siempre se pueda garantizar por completo y en bloque. Sin embargo, esto no excluye que, para lo que nos interesa ahora, en el contexto de la asistencia sanitaria, no siempre sea posible para los usuarios de esos sistemas de IA (esto es, los profesionales asistenciales) conocer cómo —siguiendo qué lógica o en base a qué correlaciones— ha llegado al máquina a los resultados que propone³³. La solución no es simplemente garantizar la intervención humana en la fase final³⁴, ya que en esas condiciones es posible que la persona usuaria no encuentre razones para validar pero tampoco para desechar justificadamente la propuesta del sistema de IA³⁵.

3. Sobre el consentimiento informado en el ámbito de la asistencia sanitaria

Vamos a considerar, entonces, cómo podría afectar el empleo de un sistema de IA así caracterizado al consentimiento informado del paciente. Para ello es conveniente considerar brevemente cuál es la razón de ser del consentimiento informado en la relación médico-paciente, para tenerla a la vista como criterio desde el que desarrollar ese juicio.

He sostenido en otro lugar que existen dos visiones básicas del consentimiento informado en el ámbito de la asistencia sanitaria³⁶. Una lo entiende como expresión

del derecho a la vida y a la integridad física y moral, fundamentado en la dignidad e inviolabilidad de la persona. Para esta visión, el consentimiento informado funciona principalmente como un límite negativo, de manera que la intervención médica necesita, para ser legítima, ser consentida por el paciente (o, por lo menos, no expresamente rechazada, en los casos en que el paciente no puede prestar consentimiento). Su fundamento es, como decíamos, la dignidad del ser humano, que implica la exigencia de ser tratado como persona, no como cosa; como sujeto y no como objeto. En el contexto de la relación asistencial esta exigencia se concreta, en particular, en el deber de respeto a la corporeidad de la persona, que lleva a no intervenir en esa persona si no es con su consentimiento (que puede entenderse presunto en determinadas circunstancias, como adelantábamos). Así, a través del consentimiento informado, el paciente acepta la intervención del profesional asistencial, haciéndola legítima, o la rechaza. A mi juicio, ésta es la función propia del consentimiento informado.

La segunda visión del consentimiento informado es la que lo considera como un mecanismo para garantizar la realización de la elección autónoma. Ésta, a su vez, se entiende como manifestación de la facultad de autodeterminación de la persona, concebida como la capacidad de darse a uno mismo normas que orienten el propio comportamiento libre. En esta visión el consentimiento informado es un instrumento para que el paciente decida qué han de hacer otros (los profesionales) respecto de su propio cuerpo; de esta manera, se autodetermina. Su fundamento es, entonces, la autonomía entendida como la condición de la voluntad de ser ley para sí misma.

La segunda comprensión del consentimiento informado está muy extendida entre académicos y es muy frecuente en propuestas bioéticas. Entre los juristas era menos frecuente: sí se empleaba el lenguaje autonomista, pero el sentido del consentimiento informado solía entenderse al modo de la primera de las visiones expuestas: como un instrumento para aceptar o rechazar un tratamiento o intervención, y no como instrumento de autodeterminación, que permita exigir de un tercero una intervención determinada (entre otras cosas, por

32 Cfr. *ibid.*, 705-706. A juicio de estos autores hay medios técnicos para aumentar la explicabilidad, lo que ocurre es que exige atender a las fases previas a la del algoritmo ya desarrollado y tomando decisiones en el mundo real: esas fases previas conllevan mucha más implicación humana (657-658). Los autores insisten en entender el *machine learning* como el resultado de una serie de procesos humanos, no como un hecho estático, completo (715).

33 Cfr. Mittelstadt, B., *op. cit.* 16. Cfr. también Sparrow, R. y Hatherley, J. *op. cit.* 16.

34 Cfr. Lehr, D. y Ohm, P. *op. cit.* 716.

35 Cfr. Comitato Nazionale per la Bioética y Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita, *op. cit.* 10-11.

36 Cfr. sobre estas dos visiones y para un desarrollo de lo que se propone a continuación, Pereira Sáez, C. «Two Visions of informed Consent». *Medicine and Law*. 2016; 35: 165-172, y «Constitución y consentimiento informado: España». *Julgar*. 2014; número especial: 47-65.

las dificultades prácticas que conlleva entenderlo como instrumento para exigir en el ámbito de la asistencia sanitaria la realización de la voluntad de una persona respecto de sí misma, cualquiera que sea esa voluntad). Sin embargo, en los últimos años encontramos en España sentencias en las que se afirma expresamente que el fundamento del consentimiento informado –el derecho a la vida e integridad del art. 15 de la Constitución³⁷– es un derecho de autodeterminación: por ejemplo, la sentencia del Tribunal Constitucional español 19/2023, de 22 de marzo, donde se afirma que ofrecer la eutanasia y el suicidio asistido es necesario para garantizar la autodeterminación de la persona, esto es, para dar cumplimiento a la voluntad de la persona respecto de sí misma, igual que lo es respetar el rechazo de tratamiento a través del consentimiento informado³⁸. Parecería, entonces, que, por lo menos en la práctica del Derecho en España, la diferencia entre esas dos visiones del consentimiento informado resulta superflua: la razón de que atendamos al consentimiento (o rechazo) de una intervención no sería el respeto a la corporeidad de la persona, sino a su capacidad de autodeterminación. Por otro lado, esta perspectiva podría parecer especialmente interesante para tratar la relevancia del empleo de sistemas de IA en la asistencia sanitaria, teniendo en cuenta que el riesgo más relevante en este contexto parece ser el de la imposición (por inescrutable) de una decisión ajena, la del sistema de IA. Esto lo corroboraría, por otro lado, la frecuencia con que aparece la referencia a la autonomía del paciente en las propuestas de pautas éticas para el empleo de sistemas de IA en la asistencia sanitaria.

Sin embargo, no es la autodeterminación así entendida lo que pretende garantizar el Derecho, en la práctica; además, resulta teóricamente débil. Por lo que se refiere a lo primero, si tratara de garantizar la autodeterminación,

el Derecho afirmaría el deber jurídico de dar cumplimiento a todos nuestros deseos respecto de nosotros mismos (con la única excepción de aquellos que afectasen negativamente a terceros, como exigencia de la coordinación de autonomías); pero no lo hace, como sabemos: por ejemplo, el Derecho español no atiende cualquier solicitud de muerte, sino sólo la de quienes se encuentren en contexto eutanásico. Por lo que se refiere a lo segundo, lo cierto es que la comprensión de la autonomía como auto-legislación, en la que se basa esa noción de autodeterminación, puede resultar menos convincente de lo que parece a primera vista. En la medida en que implica la total indeterminación de la voluntad, esto es, su indiferencia radical y originaria, supone que el obrar libre es, en realidad, incomprensible: la única razón de la determinación de la voluntad es la propia voluntad³⁹. En último término, el problema de entender la libertad como autonomía es ignorar que la libertad, en un cierto sentido y quizá en su estructura más radical, no consiste tanto en un acto de la voluntad como en un reconocimiento. Esto es, deliberamos para decidir libremente y nuestras acciones libres se deben a razones, pero que las razones sean tal no se debe a una decisión libre nuestra, sino que cuando deliberamos nos movemos en un horizonte dado de antemano. Hacemos lo que, desde puntos de vista dados de antemano, nos parece mejor, y nos parece así porque somos como somos⁴⁰. La noción de libertad como autonomía niega este contexto de la deliberación, esta estructura radical de la libertad⁴¹.

Por lo tanto, a pesar de lo más o menos extendidas que estén una y otra visión del consentimiento informado, la que lo considera expresión de la capacidad de au-

37 El art. 15 de la Constitución española establece que “todos tienen derecho a la vida y a la integridad física y moral, sin que, en ningún caso, puedan ser sometidos a tortura ni a penas o tratos inhumanos o degradantes. Queda abolida la pena de muerte, salvo lo que puedan disponer las leyes penales militares para tiempos de guerra.” La sentencia del Tribunal Constitucional español 37/2011, de 28 de marzo, afirma expresamente que este artículo es el fundamento constitucional del consentimiento informado.

38 Cfr. STC 19/2023, de 22 de marzo, FJ 6 C, d) y e), donde la integridad moral se entiende como genérica capacidad de autodeterminación.

39 Esta manera de entender la libertad implica el riesgo, como escribe Taylor, de perder completamente el sentido, porque la razón de ser de la libertad no puede ser la libertad misma, sino que es, más bien, el tipo de bien que necesita de la libertad para ser elegido. La auto-elección como ideal solo tiene sentido porque hay ciertas cuestiones que son más significativas que otras. Y qué cuestiones sean significativas no es cosa que uno mismo determine; si fuera uno mismo quien lo decidiese, ninguna cuestión sería significativa. Cfr. Taylor, C., *La ética de la autenticidad* (trad. de Pablo Carbajosa), Barcelona, Paidós, 1994, 74-75.

40 Cfr. Spaemann, R., *Persons. The Difference Between ‘Someone’ and ‘Something’*, (trad. de Oliver O’Donovan), Oxford, Oxford University Press, 2006, 209-220.

41 Estructura que, a su vez, no impide la libertad de la deliberación porque la “meta-voluntad” de ser feliz no fija acciones concretas ni actos de voluntad. Cfr. Spaemann, R., *op. cit.*, pp. 209-220.

todeterminación resulta limitada. El triunfo del lenguaje autonomista puede explicarse porque, efectivamente, el consentimiento informado es un instrumento que ha de ejercerse *libremente*; sin embargo, esto no significa que su única razón de ser sea proteger la libertad, como si afirmar que protege algún otro bien supusiera negar esa libertad que se quiere proteger. La comprensión autonomista de la libertad se caracteriza porque, para afirmar la libertad, niega que existan bienes que atraigan a la voluntad, porque entiende que sólo así es libre, aunque haya de quedar reducida a una voluntad aleatoria, que quiere sin causa. Hay, no obstante, otro sentido posible de la libertad, que no la condena a la irracionalidad, que la entiende no como voluntad que es ley respecto de sí misma, cosa imposible⁴², sino como voluntad que elige, porque no está determinada, pero en un contexto previo, a partir de algo dado: las formas básicas del bien humano. Así, la libertad no es tanto el fin como la condición de la persona humana, siendo el bien de la persona un bien que necesita de la libertad para ser elegido⁴³.

4. Una propuesta

A mi juicio, éste es el contexto en el que se entiende adecuadamente el sentido del consentimiento informado en el ámbito sanitario, en donde funciona como instrumento de respeto a la integridad de la persona, particularmente en su corporeidad, precisamente en cuanto sujeto libre, no mero cúmulo de células y órganos. Pero tal respeto a la libertad de la persona no consiste en poner por obra su voluntad, sino más bien en considerar que su bien necesita de la libertad para ser elegido. Por este motivo el consentimiento del paciente necesita la información previa por parte del profesional: sólo se puede prestar consentimiento respecto de algo que se conoce, porque la voluntad sólo puede querer lo que la inteligencia le presente como bueno⁴⁴. Pues bien, en el ámbito de la asistencia sanitaria, a diferencia de lo que ocurre en otros ámbitos, no se presume que el paciente conoce y

comprende aquello respecto de lo que ha de consentir (o no), por ser de carácter especializado o científico; por eso se entiende que es obligación del profesional transmitir la información relevante⁴⁵. En estas condiciones, en el contexto de una relación interpersonal, el consentimiento del paciente legitima la intervención del profesional.

La atención al fundamento del consentimiento informado nos da un punto de vista desde el que observar la intervención de sistemas de IA en la asistencia sanitaria —intervención que, por ahora, será habitualmente mediante sistemas parciales de apoyo en la toma de decisiones. Evidentemente, es necesario que los sistemas respeten las exigencias éticas aplicables en general, como la solidez técnica, etc. Por lo que se refiere en concreto al consentimiento informado para la asistencia sanitaria⁴⁶ el criterio básico es, a mi juicio, que la intervención de la IA sea tal que no impida que el consentimiento informado así entendido pueda tener sentido; para ello es necesario que la relación no pierda su carácter interpersonal. En efecto, sólo procede otorgar un consentimiento ante otra persona, ya que la razón de ser no es la autodeterminación de quien consiente, sino la legitimación de quien interviene. Esto exige como pauta básica que el sistema de IA funcione siempre como instrumento en manos del profesional, no como sustituto. En este sentido han de entenderse las propuestas más frecuentes en relación con el consentimiento informado del paciente⁴⁷, de supervisión humana y de explicabilidad. Así, la supervisión humana no ha de ser un mero sello al final del proceso, sino que exige revisar todos los pasos previos de los desarrollos de los sistemas de IA, donde la intervención humana resulta mucho más factible y eficaz, como se adelantaba. Por lo que se refiere a la explicabilidad, el profesional asis-

42 Cfr. Anscombe, E. «Modern Moral Philosophy». *Philosophy*. 1958; 33(124): 1-19, 2.

43 Cfr. Taylor, C., *op. cit.*, 74-75.

44 Cfr. Finnis, J., *Aquinas. Moral, Political and Legal Theory*, Oxford, Oxford University Press, 1998, 62 y ss.

45 Para esto es esencial cómo se entienda la relación médico-paciente, ya que ese “transmitir la información” es, en realidad, mucho más: explicar, ofrecer alternativas, ayudar a ponderar riesgos, etc. Mittelstadt estudia la afectación de la IA a esa relación teniendo en cuenta justamente esa dimensión de relación interpersonal y su dimensión ética: cfr. Mittelstadt, B., *op. cit.* 35-43 y 66-67.

46 Como se adelantaba, atendemos aquí sólo al consentimiento informado para la asistencia sanitaria, no al consentimiento informado para el uso secundario de los datos.

47 Cfr., entre otros, Mittelstadt, B., *op. cit.* 57-60; Comitato Nazionale per la Bioética y Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita, *op. cit.* 14-16, OMS, *op. cit.* 25-26 y 45-54.

tencial sólo puede hacer suyo, o rechazar con razones suficientes, el resultado que obtenga el sistema de IA, si es capaz de comprender mínimamente sus razones; y sólo así podrá explicarlo y proponerlo al paciente. Para que la explicabilidad no resulte poco más que una pretensión ingenua, panacea aparente para todos los problemas⁴⁸, a la que se pueda responder, escépticamente, que lo que define a los sistemas de aprendizaje automatizado de IA es justamente su falta de explicabilidad, es interesante atender a perspectivas como la que veíamos que proponen Lehr y Ohm: no entender el aprendizaje automatizado de manera monolítica, sino por fases, de tal manera que el sistema pueda resultar más o menos explicable, también en función de para quién. Para los usuarios, el conocimiento del sistema de IA necesario es el conocimiento en cuanto instrumento, como ocurre en tantos otros ámbitos técnicos: más que conocer cómo funciona, es necesario conocer qué hace: qué determina y con qué criterios o razones. Esto puede no ser tarea fácil; habrá que ver hasta dónde es posible conocer, cuál es el ámbito de inescrutabilidad o incertidumbre, cuáles son los riesgos de error o los sesgos y, en general, el resto de limitaciones conocidas que puedan acompañar la propuesta del sistema de IA. En esas condiciones, el profesional podrá decidir el peso que otorga a la propuesta del sistema de IA y explicarlo al paciente.

En conclusión, entonces, el empleo de la IA en la asistencia a personas pasa por entender la IA como un sistema desarrollado por personas, desmitificándolo. Esto no supone ignorar la capacidad que puedan llegar a tener los sistemas de IA ni la seriedad de los riesgos que implica sino, más bien, no perder la confianza en el ser humano, no renunciar a lo que es propio de la persona, libre y responsable, manteniendo, también funcionalmente y no sólo en teoría, la diferencia respecto del artificio.

* * *

Este trabajo es uno de los resultados de la Ayuda para grupos con potencial de crecimiento número ED431B 2023/59 de la Xunta de Galicia.

Conflictos de intereses: no existen

⁴⁸ Cfr. Mittelstadt, B., *op. cit.* 16.

Referencias

- Anscombe, E. «Modern Moral Philosophy». *Philosophy*. 1958; 33(124): 1-19.
- Cartabria, M., «La edad de los nuevos derechos». *Revista de Derecho Político*. 2011; 81: 61-100.
- Comitato Nazionale per la Bioética y Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita, *Parere congiunto Intelligenza artificiale e medicina: aspetti etici*, Roma, 2020.
- Dicasterio para la doctrina de la fe y Dicasterio para la cultura y la educación, *Antiqua et Nova. Nota sobre la relación entre la inteligencia artificial y la inteligencia humana*, Roma, 2025.
- Finnis, J., *Aquinas. Moral, Political and Legal Theory*, Oxford, Oxford University Press, 1998.
- Gaudet, M., Herzfeld, N., Scherz, P. y Wales, J. (eds.), *Encountering Artificial Intelligence. Ethical & Anthropological Investigations*, Pickwick, Eugene, 2024.
- Grupo independiente de expertos de alto nivel sobre inteligencia artificial creado por la Comisión Europea en junio de 2018, *Directrices éticas para una IA fiable*, Bruselas, 2019.
- Kurzweil, R., *La singularidad está más cerca. Cuando nos fusionamos con la IA*, (trad. de Alexandre Casanovas) Barcelona, Deusto, 2025.
- Lehr, D. y Ohm, P. «Playing with the Data: What Legal Scholars Should Learn About Machine Learning». *UC Davis Law Review*. 2017; 51: 653-717.
- Línea de trabajo de Transformación digital de DigitalEs, *Inteligencia artificial ética en sanidad. Recomendaciones para la adopción de IA respetuosa, transparente, segura y justa*, DigitalEs, 2022.
- McCarthy, J. Minsky, M., Rochester, N. y Shannon, C. «A Proposal for the Dartmouth Summer research project on Artificial intelligence, August 31, 1955». *AI Magazine*. 2006; 27(4): 12-14.
- Mittelstadt, B., *The Impact of Artificial Intelligence on the Doctor-Patient Relationship. Report Commissioned by the Steering Committee for Humans Rights in the fields of Biomedicine and Health (CDBIO)*, Bruselas, Consejo de Europa, 2021.

- OMS, *Ética y gobernanza de la inteligencia artificial en el ámbito de la salud: orientaciones de la OMS*, Ginebra, 2021.
- Pereira Sáez, C. «Constitución y consentimiento informado: España». *Julgar*. 2014; número especial: 47-65 — «Two Visions of informed Consent». *Medicine and Law*. 2016; 35: 165-172.
- Russell, S. y Norvig, P., *Artificial Intelligence. A Modern Approach* (4ª ed.), Londres, Pearson, 2022.
- Spaemann, R., *Persons. The Difference Between 'Someone' and 'Something'*, (trad. de Oliver O'Donovan), Oxford, Oxford University Press, 2006.
- Solomonides, A. et alii. «AMIA Position Paper. Defining AMIA's artificial intelligence principles». *Journal of the American Medical Informatics Association*. 2022; 29(4): 585-591.
- Sparrow, R. y Hatherley, J., «High hopes for "Deep Medicine"? AI, Economics and the Future of Care». *Hastings Center Report*. 2020; 50(1): 14-17.
- Taylor, C., *La ética de la autenticidad* (trad. de Pablo Carbajosa), Barcelona, Paidós, 1994.
- Topol, E., *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*, Nueva York, Basic Books, 2019.
- UNESCO, *Recomendación sobre la ética de la inteligencia artificial*, 2021.