



Informes Anticipando  
**INTELIGENCIA  
ARTIFICIAL EN SALUD:  
RETOs ÉTICOS Y LEGALES**



### Informe Anticipando coordinado por:

#### **Carlos Romeo**

Catedrático de Derecho Penal y Coordinador del Grupo de Investigación Cátedra de Derecho y Genoma Humano de la Universidad del País Vasco/EHU.



### Expertos colaboradores:

#### **Encarnación Guillén**

Presidenta Asociación Española Genética Humana (AEGH). Miembro Comité Bioética de España. Jefa Sección Genética Médica-HCUVA. Universidad de Murcia. IMIB-Arrixa. CIBERER-ISCIII.

#### **José Manuel Jerez**

Catedrático de Lenguajes y Sistemas Informáticos. Departamento de Lenguajes y Ciencias de la Computación. ETSI Informática de la Universidad de Málaga

#### **Iñigo de Miguel**

Investigador Distinguido del Grupo Cátedra de Derecho y Genoma Humano de la Universidad del País Vasco. Research Professor del Ikerbasque Basque Foundation for Science.

#### **Pilar Nicolás**

Investigadora Doctora del Grupo Cátedra de Derecho y Genoma Humano y profesora de la Facultad de Derecho de la Universidad del País Vasco.



### Comité Asesor Observatorio de Tendencias de Medicina Personalizada de Precisión:

#### **Joaquín Arenas**

Director del Instituto de Investigación del Hospital Universitario 12 de Octubre (i+12).

#### **Ángel Carracedo**

Director de la Fundación Pública Gallega de Medicina Genómica (Servicio Gallego de Salud) y Coordinador del Grupo de Medicina Genómica de la Universidad de Santiago de Compostela (CIBERER).

#### **Pablo Lapunzina**

Jefe de grupo de investigación del Instituto de Genética Médica y Molecular (INGEMM) del idipaz y Director científico del CIBERER.

#### **Fernando Martín-Sánchez**

Profesor de Investigación en Informática Biomédica en Instituto de Salud Carlos III.

Nº de depósito legal: M-24974-2020

ISBN edición impresa: 978-84-09-23429-5

©2020 del contenido: Fundación Instituto Roche. Se permite la reproducción parcial, sin fines lucrativos, indicando la fuente y la titularidad de la Fundación Instituto Roche sobre los derechos de la obra.

[www.institutoroche.es](http://www.institutoroche.es)

Con la colaboración de Ascendo Consulting Sanidad&Farma.

# Contenidos

<b>PRESENTACIÓN.....</b>	<b>5</b>
<b>RESUMEN EJECUTIVO .....</b>	<b>7</b>
<b>INTRODUCCIÓN.....</b>	<b>9</b>
<b>LOS DERECHOS .....</b>	<b>12</b>
Derechos respecto a la gestión de los datos empleados en los procesos de aprendizaje de los algoritmos de IA.....	12
Derechos respecto a los resultados ofrecidos por las herramientas de IA.....	13
<b>RETOS ÉTICOS Y LEGALES .....</b>	<b>15</b>
Garantizar el anonimato de los individuos y pacientes .....	15
Protección de la información del individuo por la posibilidad de establecer relaciones entre datos masivos a través de las tecnologías del big data .....	16
Evitar posibles sesgos a la hora de tomar decisiones basadas exclusivamente en perfiles y en procesos de estratificación.....	17
Garantizar que la toma de decisiones no sea automatizada o basada exclusivamente en propuestas del sistema de IA o en perfiles obtenidos de los pacientes.....	17
Asegurar la transparencia sobre los elementos esenciales y sobre el proceso por el cual un sistema de IA llega a una conclusión decisoria .....	19
Establecer a quién atribuir y quién debe asumir responsabilidades a la hora de emplear herramientas de IA en salud .....	19
<b>CONCLUSIONES Y RECOMENDACIONES .....</b>	<b>21</b>
<b>BIBLIOGRAFÍA .....</b>	<b>23</b>



Informes Anticipando

## INTELIGENCIA ARTIFICIAL EN SALUD: retos éticos y legales





# PRESENTACIÓN

---

Los Informes Anticipando, elaborados en el marco del Observatorio de Tendencias de Medicina del Futuro promovido por la Fundación Instituto Roche, surgen con el objetivo de contribuir a la generación y puesta en común del conocimiento, así como a la difusión de los avances que se producen en la evolución de la Medicina Personalizada de Precisión (MPP) para colaborar a traer al presente la medicina del futuro.

El Observatorio cuenta con un Comité Asesor de expertos formado por el **Dr. Ángel Carracedo**, el **Dr. Joaquín Arenas**, el **Dr. Pablo Lapunzina** y el **Dr. Fernando Martín**. Entre sus funciones se incluye la selección de las temáticas que abordan estos informes, la identificación de expertos y la validación de los contenidos.

Este informe que versa sobre los retos éticos y legales de la aplicación de la Inteligencia Artificial en salud está coordinado por el **Dr. Carlos M. Romeo Casabona** y en su elaboración han participado como expertos la **Dra. Pilar Nicolás**, el **Dr. Iñigo de Miguel**, el **Dr. José Manuel Jerez** y la **Dra. Encarnación Guillén**.

El **Dr. Carlos Romeo** es catedrático de Derecho Penal y coordinador del Grupo de Investigación Cátedra de Derecho y Genoma Humano, UPV/EHU. Doctor en Derecho y en Medicina y Diplomado Superior en Criminología. Fue magistrado de la Audiencia Territorial de Zaragoza, decano de la Facultad de Derecho, fundador y director del Centro de Estudios Criminológicos de la Universidad de La Laguna. Miembro del Grupo Europeo de Ética de Ciencias y Nuevas Tecnologías de la Comisión Europea, del Comité de Bioética del Consejo de Europa, de la Real Academia de Jurisprudencia y Legislación, de la Real Academia de Medicina de Zaragoza y presidente del Comité de Ética de Investigación con Medicamentos de Euskadi. Fue vocal del Comité de Bioética de España, de la Comisión Nacional de Reproducción Humana Asistida y del European

Research Area Board de la Comisión Europea. Director de las colecciones de Estudios de Derecho Penal y Criminología, Derecho y Ciencias de la Vida, de la Revista de Derecho y Genoma Humano/Law and te Human Genome Review y de la Enciclopedia de Bioderecho y Bioética.

La **Dra. Encarnación Guillén** es Licenciada en Medicina y Cirugía y Doctora en Medicina, programa de Salud Pública y Medicina Preventiva, por la Universidad de Murcia. Especialista en Pediatría y áreas específicas (MIR, 1990-1994, Hospital Clínico Universitario Virgen de la Arrixaca). Fellowship in Medical Genetics (1995-1997, New York Medical Center-New York University). Máster Universitario en Dirección y Gestión Sanitaria (2019, UNIR). Experta de la OMS en el Grupo de Trabajo de Genética del TAG pediátrico para la 11<sup>a</sup> Clasificación Internacional de Enfermedades (2011 a 2016). Miembro del Comité de Expertos de la Federación Española de Enfermedades Raras (FEDER). Ex Presidenta de la Sociedad Española de Genética Clínica (SEGCD). Exconsejera de Sanidad de la Región de Murcia. Actualmente Presidenta de la Asociación Española de Genética Humana (AEGH), miembro del Comité de Bioética de España y vocal de la Comisión Nacional de Reproducción Humana Asistida. Ha participado en el European Shared Blueprint of Digital Transformation of Health and Care for the Ageing Society (2016) y en el informe de la Región de Murcia para promover la digitalización de los sistemas de salud a nivel europeo, aprobado unánimemente por el Comité Europeo de las Regiones en 2019.

El **Dr. José Manuel Jerez** es Catedrático de Lenguajes y Sistemas Informáticos en la Escuela Técnica Superior de Ingeniería Informática de la Universidad de Málaga. Responsable del grupo de investigación ICB (Inteligencia Computacional en Biomedicina) de la Junta de Andalucía, ha desarrollado su labor investigadora principalmente en diseño y aplicación de algoritmos de inteligencia artificial

## INTELIGENCIA ARTIFICIAL EN SALUD: retos éticos y legales

en el área de la oncología. Ha dirigido tanto proyectos de investigación, orientados a la aplicación de técnicas de minería de datos a problemas de clasificación en el ámbito de la biomedicina y la ingeniería. Investigador Principal de diferentes proyectos dentro del área de IA en salud, co-responsable del diseño e implantación de Galén, sistema de información oncológico desplegado en la UGC Oncología Intercentros de los hospitales Regional y Virgen de la Victoria de Málaga, para la gestión y coordinación de todos los procesos implicados en el funcionamiento de una unidad de oncología médica, y la transferencia de los resultados de investigación derivados de la ejecución de los proyectos centrados en la aplicación de IA al área de oncología, como el modelado predictivo de supervivencia, la minería de texto y procesamiento de lenguaje natural aplicado a campos no estructurados de la HCE, flujos de trabajo para análisis de datos bioinformáticos, inteligencia de negocio para la gestión de datos en Salud, etc.

El **Dr. Iñigo de Miguel** es Doctor en Derecho y Filosofía. Investigador distinguido en el Grupo de Investigación de la Cátedra de Derecho y Genoma Humano del Departamento de Derecho Público e Ikerbasque Research Professor. Ha sido profesor de cursos de posgrado en la UNED, la Universidad de Comillas, la Universidad San Pablo-CEU, la Universidad de Sevilla, la Universidad de Valladolid, la Universidad de Calabria (Italia) y la Jagiellonian University of Krakow, Polonia. Ha participado en más de diez proyectos de investigación financiados por la Comisión Europea y ocho proyectos de investigación nacionales, casi todos orientados a la bioética y el derecho de las nuevas tecnociencias. Cuenta con el Premio de investigación sobre bioética de la Fundación Víctor Grifols i Lucas, obtenido por su trabajo “La clonación, diez años después” y el premio internacional de bioética de la Junta General del Principado de Asturias y la Sociedad Internacional de Bioética 2008. Ha participado como investigador

en el proyecto Bigdatius sobre el uso de datos clínicos ante nuevos retos tecnológicos y científicos (big data), y las oportunidades e implicaciones jurídicas. Actualmente coordina el proyecto PANELFIT (*Participatory Approaches to a New Ethical and Legal Framework for ICT*) desarrollado con fondos del *H2020 research and innovation programme*, financiado por la Comisión Europea.

La **Dra. Pilar Nicolás** es Licenciada y Doctora en Derecho e investigadora de la Facultad de Derecho de la UPV/EHU desde 2006. Ha sido Vicedecana Coordinadora de la Sección de Bizkaia de la Facultad de Derecho (2016-2018) y coordina el Grado en Derecho en Bizkaia. Es miembro del Grupo de Investigación Cátedra de Derecho y Genoma Humano (Grupo A del Sistema Universitario Vasco). Es vocal del Comité de Ética de la Investigación con Seres Humanos de la UPV/EHU y del Comité de Ética de la Investigación del Instituto de Salud Carlos III (Ministerio de Ciencia, Innovación y Universidades). Es miembro del Comité Científico Externo de la Plataforma Red Nacional de Biobancos-del Instituto de Salud Carlos III. Es miembro del Ethics and Policy Committee del International Cancer Genome Consortium. Coordina el International Legal and Ethical Advisory Board de la Rare Anaemias Disorders European Epidemiological Platform. Investigadora principal del proyecto *Bigdatius*, finalizado en 2019, con el objetivo de analizar las implicaciones legales de la gestión de bases de datos con fines biomédicos. Participa en el Proyecto internacional *Eucanimage*, de la Comisión Europea en el que se busca generar una plataforma de imágenes oncológicas para el desarrollo de herramientas de IA con fines diagnósticos. Forma parte del grupo de trabajo sobre aspectos éticos y legales de la iniciativa *1+ Million Genomes* (Comisión Europea), en la iniciativa *Global Alliance for genomics and health* y forma parte del *Ethics, Policy and Governance* del ICGC ARGO.





## RESUMEN EJECUTIVO

---

Hoy en día los **sistemas de información**, y en concreto la **Inteligencia Artificial** (IA), están jugando un papel muy importante en la **transformación digital** de nuestro entorno. La IA busca **emular la inteligencia humana** a través de **algoritmos y herramientas computacionales** para **generar conocimiento de manera autónoma** a partir del análisis de gran cantidad de datos.

El número de **aplicaciones de IA** aumenta exponencialmente, siendo **cada vez más complejas y sofisticadas**. Además, también son más las áreas en las que encuentran aplicación, desde el sector financiero hasta el sector sanitario. En concreto, en este último campo, se prevé que las **herramientas de IA contribuirán a la realización de tareas sencillas y repetitivas** que consumen el tiempo del personal sanitario, liberándoles para que realicen otro tipo de tareas asistenciales de más valor. Pero también a la **resolución de problemas complejos de salud**, como la **detección de la predisposición a desarrollar enfermedades**, la detección de portadores de enfermedades hereditarias o la **predicción de la respuesta individual a tratamientos**.

Estos algoritmos de IA se sirven para su desarrollo y aprendizaje del uso masivo de datos que pertenecen a los individuos. Es necesario tener en cuenta que los individuos tienen derechos sobre sus datos, que hay que respetar, y el hecho de que las herramientas basadas en IA empleen dichos datos puede entrar en conflicto

con los derechos de los individuos, planteando retos desde el punto de vista ético y legal. Además, existen otras cuestiones que, dado el ritmo vertiginoso de aplicación de estas herramientas, hay que tener en consideración y plantear para darles respuesta. Por ejemplo, el **riesgo de identificar los datos del paciente**, la **capacidad** de las herramientas de IA para establecer relaciones entre datos a través de del big data, el **riesgo de que se tomen decisiones sesgadas** con base en estos datos, la tendencia a la automatización y la opacidad sobre los elementos esenciales y sobre el proceso por el que un sistema de IA ha llegado a una conclusión que se suman al debate sobre **quién debe ser responsable de los resultados de la aplicación de IA en el campo de la salud**.

Es por esto que, de cara a poder trasladar estas herramientas a la práctica clínica y contribuir a la digitalización de la Medicina Personalizada de Precisión en el futuro, los **profesionales sanitarios, científicos, juristas, poderes públicos y la sociedad en general** deberán enfrentarse y buscar de manera conjunta soluciones a estos retos.

Soluciones que, sin que estas herramientas supongan la sustitución del profesional sanitario, permitan que los sistemas de IA se empleen en la práctica clínica por profesionales pasando por la elaboración de códigos de conducta internacionales y garanticen los derechos de los individuos.





# INTRODUCCIÓN

---

A pesar de que no existe una definición universalmente aceptada de la Inteligencia Artificial (IA), puede definirse como la ciencia que busca desarrollar sistemas que presenten un comportamiento inteligente, que consiste fundamentalmente en tratar de comprender su entorno y emular la capacidad de manejo de información, razonamiento, percepción, aprendizaje y creación que presentamos los humanos.<sup>1</sup> En este sentido, la IA en el campo de la salud consiste en la aplicación de algoritmos y técnicas que tratan de imitar dichas facultades de la inteligencia humana de manera que sea posible crear herramientas capaces de aprender a partir del análisis de gran cantidad de datos acumulados, de los que extraen conclusiones de manera autónoma, contribuyendo a la resolución de problemas complejos de salud.<sup>2</sup>

En los últimos años, el auge que está experimentando la IA tiene que ver con la **disponibilidad de grandes cantidades de datos en un soporte informatizado** (como por ejemplo las historias clínicas digitales, las imágenes informatizadas, las apps de salud, los *wearables*, etc.) que pueden ser analizados de manera automatizada por **algoritmos cada vez más complejos y sofisticados** (como las redes de aprendizaje profundo o el *deep learning*<sup>a</sup>, y más concretamente, las redes neuronales artificiales<sup>b</sup>), gracias a la **creciente potencia de los recursos y las tecnologías computacionales** (para más información, consultar los Informes Anticipando sobre Biología de Sistemas<sup>c</sup> y Los datos en la era de la Medicina Personalizada de Precisión<sup>d</sup>).<sup>3</sup> Esta posibilidad de analizar gran cantidad de datos (o *big data*) de manera automatizada supone un cambio de paradigma, puesto que permite obtener información que aparentemente no está implícita en los datos o que no se puede obtener o deducir directamente de ellos. Este hecho hace que las herramientas de IA encuentren aplicación en el campo de la salud, entre otros, en áreas como la detección de la predisposición a

desarrollar enfermedades, la detección de portadores de enfermedades hereditarias o la predicción de la respuesta individual a tratamientos.<sup>4</sup>

## EL DESARROLLO DE LA IA EN EL SECTOR DE LA SALUD Y POTENCIALIDADES PARA LA MEDICINA DEL FUTURO

---

Recientemente ha aumentado de forma notable el número de herramientas basadas en IA que se están desarrollando para su aplicación en el campo de la salud y la investigación biomédica, y hay un creciente interés en este ámbito. Dos indicadores del auge de estas herramientas son, por un lado, las convocatorias específicas para financiar proyectos de “Salud, IA, *big data* y Medicina Personalizada” en el marco del Programa H2020 de la Unión Europea, indicando que la combinación de esos cuatro pilares es estratégica; y por otro lado, a nivel nacional, la publicación en 2019 de la Estrategia Española de I+D+i en IA, donde se señala que la atención sanitaria es un sector que se va a beneficiar enormemente de la IA.

En general, las herramientas basadas en IA que se están desarrollando para su aplicación en medicina están dirigidas a ayudar en el diagnóstico, en la supervisión de la evolución de pacientes y en la elección de tratamientos,<sup>1</sup> aunque también existen otras aplicaciones para el campo de la investigación biomédica, la gestión del sistema sanitario e incluso dirigidas a los pacientes en ámbitos como el del autocuidado. Para este conjunto de objetivos son decisivas la concepción y la metodología de la que parte la Medicina Personalizada de Precisión, la cual requiere por su parte una reflexión propia sobre sus dimensiones ética y jurídica.<sup>5</sup>

<sup>a</sup>Modalidad de aprendizaje automático basado en estructuras algorítmicas y computacionales complejas.

<sup>b</sup>Modalidad de aprendizaje profundo basado en estructuras computacionales que emulan las redes neuronales del cerebro humano.

<sup>c</sup><https://www.institutoroche.es/observatorio/biologiadesistemas>

<sup>d</sup><https://www.institutoroche.es/observatorio/losdatos>



# INTELIGENCIA ARTIFICIAL EN SALUD: retos éticos y legales

En cuanto a la **actividad asistencial**, la aplicación de la IA en el diagnóstico, la prevención y el tratamiento de patologías humanas supondrá un cambio radical en el escenario de la práctica de la medicina<sup>6</sup> ya que los profesionales sanitarios se encuentran frente a cantidades ingentes de datos de diversas fuentes. Estos datos son difíciles de traducir en información útil mediante análisis “tradicionales”, pero la capacidad de análisis de las herramientas de IA se presenta como la solución más prometedora. Concretamente, existe desde hace tiempo una gran actividad a nivel internacional en el desarrollo de técnicas y herramientas de IA para su aplicación a la solución de problemas de naturaleza biomédica, centrada particularmente en las siguientes áreas:

- **Análisis de imagen médica.** Se trata de la aplicación en salud donde más avances se están produciendo, gracias al hecho de poder compartir muchas imágenes médicas (radiológicas, histológicas de tumores, etc.) y la capacidad de procesarlas, tratarlas, compararlas y analizarlas con herramientas de IA basadas en redes neuronales artificiales<sup>5</sup>. El análisis continuado de imágenes médicas supone un entrenamiento para estas herramientas de manera que, a partir de lo aprendido de otros pacientes, contribuyen a establecer el diagnóstico. Por esto, es previsible que especialidades médicas como la dermatología, la radiología o la anatomía patológica se vean enormemente beneficiadas por estas herramientas a corto plazo.<sup>2</sup>
- **Sistemas de Apoyo a la Decisión Clínica.** El escenario más probable de la medicina del futuro viene construido en torno a un modelo en el que los profesionales de la salud dispondrán de información a tiempo real de determinadas variables provenientes de diferentes fuentes, que se integrarán en modelos automatizados de apoyo a la decisión clínica (para más información, consultar el Informe Anticipando sobre Los datos en la era de la Medicina Personalizada de Precisión<sup>1</sup>) que les asesorarán sobre qué decisión tomar para cada paciente,<sup>6</sup> diagnosticar enfermedades de baja prevalencia o detectar situaciones clínicas graves para actuar de forma más precoz.
- **Procesamiento de lenguaje natural.** Consiste en entrenar a las herramientas de IA en la

extracción de información a partir de la cantidad de datos disponibles en historias clínicas digitales (HCD) en formato de texto abierto, para que sean capaces de “entender” la semántica del texto y así detectar y predecir eventos como, por ejemplo, la progresión de la enfermedad en pacientes de cáncer, o la predicción de la asistencia a urgencias de pacientes con base en el historial de seguimiento.

- **Administración de tratamientos.** Será posible emplear este tipo de herramientas para la predicción de respuesta al tratamiento, como por ejemplo en el campo de la oncología, de la diabetes o en la determinación de resistencias bacterianas a antibióticos, donde se están dando ya los primeros pasos. Esto permitirá la determinación del tratamiento óptimo para cada paciente, contribuyendo a una mayor personalización de la Medicina.
- **Elaboración de perfiles de pacientes.** Una de las aplicaciones para las que se puede emplear la capacidad de la IA de establecer relaciones entre grandes conjuntos de datos para clasificar a las personas en grupos o subgrupos con perfiles similares de enfermedad.<sup>4</sup>

En cuanto a la **gestión sanitaria y de las estrategias de salud pública**, la IA puede emplearse para establecer correlaciones que permitan cambiar o mejorar procesos, protocolos de actuación, etc. con los siguientes fines.

- **La gestión del centro o del sistema sanitario.** Al igual que con los Sistemas de apoyo a la Decisión Clínica, se pueden diseñar o adaptar estos algoritmos para, por ejemplo, la estimación de recursos necesarios en UCIs.
- **Creación de modelos predictivos.** La aplicación de la IA en medicina está permitiendo crear modelos predictivos apenas imaginables hace unos años.<sup>6</sup> Estos modelos predictivos permitirán detectar patrones en múltiples patologías a partir de todo tipo de datos del individuo, no sólo los que a primera vista ofrecen relevancia clínica. Así mismo, la capacidad de las herramientas de IA para detectar asociaciones no perceptibles a simple vista, permitirá su aplicación no solo en el campo de la prevención de enfermedades, sino que se extenderá a la detección de riesgos a padecer,



<sup>5</sup>Modelos algorítmicos de IA que emulan el funcionamiento de las neuronas, conectadas en red para compartir información.  
<sup>6</sup><https://www.institutoroche.es/observatorio/losdatos>



por ejemplo, enfermedades relacionadas con el envejecimiento. Todo ello permitirá diseñar estrategias de prevención de riesgos personalizadas (para más información, consultar el Informe Anticipando sobre Medicina Preventiva Personalizada<sup>9</sup>).

En cuanto a la **investigación biomédica**, la aplicación de la IA consistirá en aplicar herramientas de aprendizaje automático, de análisis probabilísticos, de *deep learning* y de análisis por capas para analizar gran cantidad de datos desde distintas perspectivas. Este tipo de aproximación ofrece numerosas hipótesis, a diferencia del modelo científico de generación de hipótesis basadas en evidencias (para más información ver el Informe Anticipando de Los datos en la era de la Medicina Personalizada de Precisión), para resolver interrogantes presentes, actualmente, a nivel biosanitario.

A nivel de **ciudadanos/pacientes**, la IA puede jugar un papel muy importante en el autocuidado, en la monitorización de enfermedades crónicas, en la detección de situaciones agudas con necesidad de atención, etc.

La IA va a transformar el sector de la salud en muchos aspectos, y aunque existe cierta percepción, por parte de los profesionales sanitarios, de que la incorporación en el futuro de la IA en la práctica médica podría llegar a sustituirles, el personal sanitario va a seguir teniendo un rol central en todo el proceso. La IA será un complemento en la toma de decisiones y en el apoyo al clínico para eliminar el trabajo repetitivo que se puede automatizar, de manera que la transformación más significativa a corto-medio plazo irá en la línea de liberar a los clínicos de carga de trabajo monótona y automatizable, dejándoles tiempo para dedicar a la asistencia, la monitorización y el control de los resultados en salud para lograr una atención más eficiente y segura para el paciente, así como una mayor humanización de la asistencia sanitaria.

La exactitud en la predicción de la IA es proporcional (grado de probabilidad) al número y calidad y exactitud de los datos introducidos y la experiencia almacenada sobre un tema en particular, pero siempre podría fallar si emerge un caso que nunca ha ocurrido antes. Las predicciones de la IA dependen de los datos y los algoritmos con el que el sistema está “entrenado”: las predicciones pueden ser “erróneas” debido a la inexactitud de los datos proporcionados o al uso de suposiciones infundadas.<sup>7</sup>

No obstante, estas nuevas herramientas se enfrentan a limitaciones derivadas del hecho de que los algoritmos de IA se sirven del uso masivo de datos para su desarrollo y aprendizaje, datos que por su naturaleza pertenecen al individuo y sobre los que tiene y ejerce sus derechos. Es por esto que el uso de herramientas basadas en IA que empleen dichos datos puede entrar en conflicto con los derechos de los individuos. En este sentido, profesionales sanitarios, científicos, juristas, poderes públicos y la sociedad en general deberán enfrentarse a una serie de retos para lograr que estas herramientas lleguen a la práctica clínica.

<sup>9</sup><https://www.institutoroche.es/observatorio/medicinapreventivapersonalizada>



# LOS DERECHOS

De cara a abordar la aplicación de la IA en el ámbito de la salud, hay que establecer que, jurídicamente, el punto de partida deben ser los derechos fundamentales y libertades públicas<sup>h</sup> de los individuos o del conjunto de la sociedad, y que en un futuro podrían reconocerse nuevos derechos derivados del uso del *big data* y la IA.<sup>3</sup>

Existen dos referencias fundamentales en cuanto a la protección de estos derechos de los individuos que conforman el marco jurídico básico: el Reglamento General de Protección de Datos (RGPD)<sup>i</sup>, y la Ley Orgánica de Protección de Datos y Garantía de los Derechos Digitales (LO 3/2018 del 5 de diciembre). Este marco básico recoge una serie de garantías que se tienen que atender en cuanto a la gestión de los datos para el diseño de este tipo de herramientas de IA junto a los principios básicos en los que se asienta la regulación actual, que son:

- La dignidad de la persona humana, el respeto a la autonomía de su voluntad y a su intimidad.
- El consentimiento del paciente previo a toda actuación en el ámbito de la sanidad, que debe obtenerse tras recibir una información adecuada y por escrito en los supuestos previstos en la Ley.
- El derecho a decidir libremente del paciente o usuario entre las opciones clínicas disponibles, después de recibir la información adecuada.
- El derecho del paciente a negarse, en este caso, a la aplicación de herramientas de IA en su diagnóstico, tratamiento u otras aplicaciones, como emanación del derecho de rechazo a cualquier intervención médica en general.
- El deber que los pacientes tienen de facilitar los datos sobre su estado físico o sobre su salud de manera leal y verdadera, así como el de colaborar en su obtención, especialmente cuando sean

necesarios por razones de interés público o con motivo de la asistencia sanitaria.

- La obligación de todo profesional que interviene en la actividad asistencial a la correcta prestación de sus técnicas, al cumplimiento de los deberes de información y de documentación clínica, y al respeto de las decisiones adoptadas libre y voluntariamente por el paciente.
- La obligación a guardar la debida reserva por parte de la persona que elabore o tenga acceso a la información y la documentación clínica.

Estos principios se materializan a través del reconocimiento de un conjunto de derechos a favor del paciente, que en el caso de la IA en salud podríamos clasificar en los derechos respecto a la gestión de los datos empleados en los procesos de aprendizaje de los algoritmos de IA y en los derechos respecto a la toma de decisiones basada en los resultados obtenidos con herramientas de IA, donde se engloban el derecho a la información sanitaria, el respeto al derecho a la intimidad y a la protección de los datos personales y a la autonomía del paciente, entre otros.

## DERECHOS RESPECTO A LA GESTIÓN DE LOS DATOS EMPLEADOS EN LOS PROCESOS DE APRENDIZAJE DE LOS ALGORITMOS DE IA.

- **Derecho a la intimidad y derecho a la protección de los datos personales en relación con el big data.** La disponibilidad de big data para ser procesado por las herramientas de IA puede suponer una inmensa ventaja en cuanto al



<sup>h</sup>Derecho de una persona o de un ciudadano, que emana de la dignidad humana, del libre desarrollo de la personalidad y de otros valores; se ejerce individualmente o de forma colectiva. Sus contenidos vinculan a todos los poderes públicos; su reconocimiento se establece en normas dotadas de supremacía material y su regulación y restricción vienen reservadas a la ley, que ha de respetar el contenido esencial.

<sup>i</sup>El RGPD o Reglamento (UE) 2016/679 es un reglamento europeo de 2016 que entró en vigor en España en el 2018, que tiene eficacia directa en los países miembros.



avance de la medicina, pero permite establecer relaciones entre los datos que a simple vista no eran previsibles, haciendo posible la obtención de información de carácter personal que no ha sido aportada por el individuo.<sup>8</sup> Por esto es necesario informar a la persona de manera previa sobre la posibilidad de descubrir datos inesperados relativos a su salud, para que pueda tomar una decisión antes de comenzar el análisis (Ley de Investigación Biomédica de 2007).<sup>4</sup>

- **El principio de transparencia.** Las personas tienen derecho a saber qué se hace con sus datos (para qué se utilizan, quién los utiliza, cuánto tiempo, desde dónde va a acceder, dónde se almacenan, etc.), por lo que deberán ser debidamente informados de estos aspectos relacionados con su tratamiento digital, en el caso de que se usen sus datos, y especialmente cuando se empleen herramientas basadas en IA.
- **La autonomía del paciente.** En este contexto, alude a la capacidad de las personas para decidir sobre sus propios datos. En el modelo en el que nos encontramos son necesarias las autorizaciones de los pacientes para el empleo de sus datos, ya sea en investigación (para la creación de modelos o bases de datos, por ejemplo) como en la práctica clínica, ya que puede darse el caso de que el individuo no quiera compartir sus datos ni ser sujeto de las acciones derivadas de sistemas inteligentes. Esto está contemplado en la Ley de autonomía del paciente, que señala que los pacientes tienen el deber de dar los datos que sean necesarios para realizar el diagnóstico,<sup>6</sup> pero no está recogido que estén obligados a revelar otros datos privados para que sean incluidos en el pool de datos a analizar, que podrían ser útiles en el proceso de aprendizaje de los algoritmos de IA.

## DERECHOS RESPECTO A LOS RESULTADOS OFRECIDOS POR LAS HERRAMIENTAS DE IA

Por otro lado, una vez la herramienta de IA está desarrollada, tras aplicarla en una situación concreta, se

va a llegar a una serie de resultados que conllevarán a la toma de decisiones que pueden afectar al individuo, para lo cual se recogen los siguientes derechos:

- **Principio de no discriminación y estigmatización.** Esta toma de decisiones no puede basarse en criterios discriminatorios de cualquier origen o naturaleza, por lo que es necesario conocer cuáles son los parámetros empleados por los algoritmos para alcanzar un resultado u otro. De manera previa a la aprobación para el uso en la práctica clínica de las herramientas de IA, el organismo competente deberá comprobar y garantizar que, de manera general, no existan sesgos que den lugar a situaciones de discriminación y/o estigmatización. Esto es así porque el profesional sanitario deberá tomar decisiones a partir de dichos resultados ofrecidos por la herramienta de IA y por lo tanto, es necesario que sea capaz de comprobar que no sean discriminatorios antes de aplicarlos.
- **Derecho a no ser objeto de decisiones basadas únicamente en procesos automatizados.** Al igual que sucede con otros sectores productivos de la sociedad, el futuro de la medicina tiende a la automatización de los procesos.<sup>4</sup> Sin embargo, está recogido en el artículo 22 del RGPD que las herramientas de IA tienen que ser un apoyo o una ayuda para la toma de decisiones, pero no el único criterio. Es por esto que las decisiones que se tomen sobre los individuos no pueden derivar exclusivamente de procesos automáticos sin que haya la intervención de un humano que tenga en cuenta otro tipo de parámetros que no sean meramente mecánicos: es necesario un control humano significativo (*meaningful human control*), en el que basar, en su caso, la responsabilidad humana moral y reforzar la colaboración integrada del ser humano y la tecnología de la IA.<sup>7,9</sup>
- **Derecho a la información.** En el plano asistencial, el derecho a la información al paciente está ampliamente garantizado en nuestra legislación,<sup>8</sup> en la Ley 41/2002. Sin embargo, habrá que hacer ciertas consideraciones en el marco de la IA en salud, como por ejemplo sobre la capacidad de estas herramientas de establecer relaciones que ayuden a predecir un elevado número patologías

que los individuos podrían llegar a desarrollar. Estas asociaciones no dejan de ser hipótesis planteadas por el sistema, y el profesional tiene que decidir en cada caso de cuáles informar al paciente, aplicando criterios de gravedad, grado de probabilidad, existencia de tratamiento, etc.<sup>8</sup>

- **Derecho a la explicación al paciente.** La Ley 41/2002 regula los derechos del paciente, y en particular el **derecho a recibir una información adecuada**, previa al consentimiento sobre diversos actos clínicos. Sin embargo, esta ley no indica que tenga que ser información completa o exhaustiva, es decir, que el paciente tiene derecho a conocer, en líneas generales, qué opciones de tratamiento hay, si existen efectos adversos, las consecuencias del tratamiento para su salud, etc.; dicho de otro modo, será una información adecuada a sus necesidades y le ayudará a tomar decisiones de acuerdo con su propia y libre voluntad.

En el campo de la IA, implicaría, en primer lugar, el derecho del paciente a conocer que el clínico va a utilizar estas herramientas como otro recurso adicional, las ventajas y limitaciones de este tipo de herramientas o que el paciente conozca por qué se toman las decisiones sobre su salud de manera general, sin entrar en detalle en el proceso de autoaprendizaje de la herramienta de IA.

No obstante, y en segundo lugar, el proceso técnico de toma de decisiones tiene que ser transparente, lo que comporta el llamado **derecho a la explicación**, es decir, el derecho a recibir explicaciones adecuadas e inteligibles al menos **sobre la lógica utilizada en el proceso decisional del sistema de IA**, sobre los procesos que condujeron a esos resultados, por ejemplo si el sistema de IA propone mediante un algoritmo que una

persona con una enfermedad grave no debe ser hospitalizada: “el responsable del tratamiento facilitará al interesado... información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento [de los datos] para el interesado”, recalca el Reglamento Europeo. Este es, precisamente, el aspecto específico que presenta el derecho a la explicación en relación con el derecho más genérico a la información.

- **Derechos decisionales del profesional: su salvaguarda ante la automatización de la práctica clínica.** Por otro lado, en cuanto a la aplicación de las herramientas de IA en la práctica clínica y en relación con el derecho a la explicación, los sistemas de IA no están en condiciones de proporcionar al clínico, en la actualidad, el proceso detallado por el cual un sistema inteligente ha llegado a un resultado determinado y, por consiguiente, aquél no podrá transmitir al paciente información suficiente sobre este proceso algorítmico. Sin embargo, dado que es el clínico el que tiene y ha de mantener la última palabra a la hora de tomar decisiones, para que pueda validar, o no, un determinado resultado tiene derecho a conocer los elementos esenciales en los que se ha basado la conclusión del sistema de IA.<sup>8</sup> En este sentido, los profesionales clínicos deben adquirir la formación adecuada para comprender cómo funciona la herramienta y cómo manejarla de manera adecuada; sin embargo el proceso interno de las herramientas de IA es complejo y difícil de analizar, por lo que la necesidad de profundizar en la comprensión variará según cada caso, y a profundizar en el conocimiento de estos procesos internos deberán dedicar sus esfuerzos los ingenieros de sistemas.



# RETOS ÉTICOS Y LEGALES

Está claro que las nuevas tecnologías de salud digital, y en concreto la IA, pueden contribuir al desarrollo de la medicina, proporcionando a los clínicos información hasta ahora inimaginable. El aprendizaje automático tiene un potencial indiscutible, pero el acceso al *big data*, su correspondiente análisis inteligente y la incorporación de los resultados de estos algoritmos en el campo de la salud son cuestiones que habrá que tener en cuenta, ya que plantean una serie de retos de carácter ético y legal. Si bien es cierto que de momento la traslación a la práctica clínica de todas las aplicaciones que ofrecen estas herramientas de IA no es completa, es pertinente plantear y analizar los retos e implicaciones ético-legales en profundidad, ya que, tal y como señala la OMS, constituyen desafíos urgentes en salud a los que se enfrentará la sociedad en la próxima década. La tecnología avanza tan rápidamente que la regulación debe adaptarse para afrontar todas estas cuestiones y dotarlas de marcos legales que tengan en cuenta, sobre todo, al paciente/usuario del sistema de salud, articulando a través del Derecho las propuestas y soluciones basadas en la ética.<sup>3</sup>

## GARANTIZAR EL ANONIMATO DE LOS INDIVIDUOS Y PACIENTES

El primer aspecto a tener en cuenta, desde el punto de vista del Derecho, tiene que ver con el *big data* y la privacidad de los datos, puesto que las herramientas de IA utilizan y aprenden a partir de datos de pacientes e individuos sanos. A día de hoy, ni la LO 3/2018 ni el RGPD hacen referencia al manejo de datos relacionados con la salud en el marco asistencial o clínico, sino que solo se contempla el uso de datos en la investigación biomédica.<sup>8</sup> En este caso sí que se impone la protección de los datos de los individuos, de modo que la privacidad debe

garantizarse desde el inicio en la gestión y ciclo de vida del tratamiento de datos.<sup>3</sup>

La IA puede ofrecer múltiples beneficios a la sociedad siempre que se mantenga el respeto por los derechos de las personas, su privacidad y la protección de sus datos personales.<sup>10</sup> Este aspecto es crítico sobre todo en el campo de la salud, donde los sistemas de información disponen de grandes cantidades de datos. Pero estos datos son de naturaleza heterogénea y en la mayoría de los casos sin estructurar, lo que dificulta su procesamiento y anonimización, con el consiguiente riesgo de que sea posible la reidentificación<sup>j</sup> del individuo fuente.

El procedimiento de anonimización puede ser tan sencillo como eliminar la información personal de una base de datos. Sin embargo, la anonimización completa de los datos de pacientes no siempre se puede garantizar y en la realidad el proceso es bastante complejo. Por ejemplo, se pueden emplear técnicas de procesamiento de lenguaje natural sobre texto abierto de la HCD para excluir los nombres propios de los pacientes. El problema radica en que pueden existir ambigüedades en el lenguaje que son difíciles de explicar a las máquinas (por ejemplo, el término “sol” puede ser un factor de riesgo de melanoma o el nombre del paciente).

Es previsible que la legislación vaya avanzando en este sector. En un marco regulativo diferente, la Agencia Española de Protección de Datos define una serie de procedimientos que hay que tener en cuenta para el acceso a la información personal. Por ejemplo, cuando se solicitan datos al sistema de salud el RGPD exige redactar un documento de impacto de privacidad de acuerdo a la “Guía práctica para las evaluaciones de impacto en la protección de datos”<sup>3</sup>; y también es necesario realizar un análisis de riesgos de seguridad según el esquema nacional de seguridad a través de la plataforma proporcionada por

<sup>i</sup>Capacidad de volver a identificar a un individuo a partir de sus datos, a pesar de que han sido previamente anonimizados.



el Centro Criptológico Nacional. Por lo tanto, el **principal reto** para los expertos en IA es garantizar que no se pueda reidentificar a los individuos a partir de los datos disponibles en distintos recursos, puesto que no se podría garantizar el anonimato de las personas en el futuro.

Incluso habiendo solventado el problema de la anonimización, existe un elevado **riesgo de identificar datos anónimos o reidentificar** datos como pertenecientes a una determinada persona con relativa facilidad, debido a la capacidad de la IA de establecer relaciones entre grandes cantidades de datos.<sup>11</sup> En estos casos es necesario llegar a soluciones en las que el coste del proceso de reidentificación de los pacientes sea muy superior a los beneficios que se puedan obtener en caso de conseguirlo. Desde el punto de vista legal, queda mucho por avanzar. En la LO 3/2018 sobre el RGPD únicamente se requiere un compromiso por parte de los investigadores en el campo de la salud para no reidentificar al individuo cuyos datos les ha proporcionado, y no se prevé ninguna consecuencia jurídica o sanción específica por su infracción, por lo que sería necesario verificar si es aplicable una sanción administrativa más general.<sup>4</sup> Por otro lado, los datos anonimizados no están protegidos por el RGPD y pueden transferirse sin restricciones desde el punto de vista jurídico. Sin embargo, si estos datos son reidentificados sí que están protegidos.

## **PROTECCIÓN DE LA INFORMACIÓN DEL INDIVIDUO POR LA POSIBILIDAD DE ESTABLECER RELACIONES ENTRE DATOS MASIVOS A TRAVÉS DE LAS TECNOLOGÍAS DEL BIG DATA**

Se han comentado los beneficios que supone la posibilidad que ofrece la IA de establecer relaciones dentro del análisis de datos masivos, pero ésta entraña ciertos riesgos que hay que tener en cuenta, dado que los datos no se reducen solo a información directa sobre la salud<sup>k</sup> de las personas, sino que también incluyen aquellos datos que pueden dar lugar a conclusiones sobre la salud de las personas tras su procesamiento. En este sentido, el RGPD contempla la protección reforzada de los datos de salud ,

cuyo acceso masivo para su uso en la Medicina Personalizada de Precisión está en auge.

Los algoritmos de IA utilizarán todo tipo de variables que puedan ofrecer alguna correlación con respecto al acontecimiento sobre el que se pretende ofrecer un pronóstico. Por ejemplo, un sistema de IA para la ayuda en el diagnóstico clínico puede requerir información sobre dónde vive un individuo, que a priori no es un dato sobre salud, sino un dato “circunstancialmente relevante” para la salud, cuya naturaleza jurídica no está muy definida. Sin embargo, este dato puede aportar información relevante sobre condicionantes ambientales de la salud, como los niveles de contaminación o de salubridad a los que se está expuesto. Facilitar esta información al clínico para el diagnóstico puede no suponer un problema una vez se explica la relación entre ambas variables, pero esto puede de extrapolarse a datos como el género, etnia, profesión, ideología, tendencia sexual u otros que tienen una protección singular<sup>6</sup> y que no todo el mundo está dispuesto a proporcionar. Además, una herramienta de IA puede analizar múltiples datos de distintas fuentes para extraer conclusiones sobre la salud de un individuo, que a su vez se pueden incluir en el análisis como nueva información de la que sacar nuevas conclusiones, de manera que estas herramientas de IA se retroalimentan en cierto sentido.<sup>12</sup> El problema radica en que, si un individuo no quiere proporcionar información personal, dada la capacidad de la IA de establecer relaciones, las herramientas podrían llegar a obtener dicha información, aunque no lo deseemos, e incluirla en futuros análisis.

Por consiguiente, lo importante en el tratamiento de la información no es tanto el carácter de los datos a primera vista, sino las conclusiones que se pueden derivar y el contexto en el que se pueden tratar. En este sentido, el concepto del “dato de salud” y el procesamiento del mismo es muy amplio, y la naturaleza de los datos en salud que se pueden requerir a los individuos es algo relativo, por lo que habrá que valorar y determinar si un dato concreto se puede incluir en el pool de información a analizar, o no, o en qué condiciones.



<sup>k</sup>El RGPD, art. 4.15 define los datos en salud como aquellos datos personales relacionados con la salud física o mental de una persona física, incluida la prestación de servicios de salud, que revelan información sobre su estado de salud.

<sup>l</sup>De acuerdo con lo dispuesto en el artículo 16.2 de la Constitución Española o con el artículo 9.2.a) del Reglamento (UE) 2016/679 o el artículo 9.1 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.



## **EVITAR POSIBLES SESGOS A LA HORA DE TOMAR DECISIONES BASADAS EXCLUSIVAMENTE EN PERFILES Y EN PROCESOS DE ESTRATIFICACIÓN**

Por lo general, debido a la gran velocidad a la que los algoritmos analizan gran cantidad de datos, generando nueva información y aprendiendo de ella, se llega a un punto en el que el que se supera la capacidad humana de comprobar o hacer seguimiento de cómo se llega a conclusiones, ni por parte del usuario ni del desarrollador. En consecuencia, no se sabe cómo una herramienta de IA ha llegado a determinada conclusión por la complejidad del proceso algorítmico. Además, a esta imposibilidad de seguir el proceso por su complejidad se le suma el hecho de que a veces dicho proceso está sujeto a secreto empresarial. El propio Parlamento Europeo señaló que esto podía dar lugar a situaciones de discriminación y sesgo algorítmico, por lo que propuso un marco ético común de máxima prudencia, en referencia a la necesidad de evaluaciones periódicas sobre la representatividad de los conjuntos de datos y de examinar la exactitud e importancia de las predicciones.<sup>3</sup>

Teniendo en cuenta lo comentado anteriormente sobre la naturaleza de los datos que se pueden pedir a un individuo para que los procese un sistema de IA, cabe la posibilidad de que algunas personas, ejerciendo libremente sus derechos, no quieran facilitar dicha información y por lo tanto esos datos no se incluyan para que un algoritmo aprenda de ellos, disminuyendo la representatividad de la muestra. En un futuro, al emplear un determinado algoritmo para tomar decisiones sobre la salud de la población en el que no se haya incluido en la muestra información de todos los individuos (porque así lo han deseado), podría ocurrir en una toma de decisiones sesgada, incumpliendo el principio de igualdad en cuanto a tratamiento dentro del sistema de salud.

Además, la **toma de decisiones basadas en procesos automatizados o en perfiles obtenidos de pacientes**, supone el riesgo de que al clasificar o estratificar a los pacientes con base en criterios que establecen los desarrolladores, por lo que podrían tomarse **decisiones discriminatorias, estigmatizadoras o arbitrarias si estos**

criterios están sesgados o si los perfiles son el único aspecto que se tiene en cuenta a la hora de tomar una decisión.<sup>11</sup>

Por otro lado, los algoritmos aprenden de los datos históricos acumulados. Esto supone un problema porque, por ejemplo, se puede incurrir nuevamente en otro tipo de sesgos (por ejemplo, por razón de género o de etnia) como resultado de la reproducción de modelos pasados (lo que se conoce como dictadura de los algoritmos). Esto plantea la posibilidad de que la decisión o la conclusión a las que se lleguen tras el análisis de datos, por parte de algoritmos, puede, o no, ser la más óptima, y por ello, a la hora de aplicarlos en clínica es necesario que se puedan validar. En este sentido es necesario regular esta tecnología con normativa similar a la de los ensayos clínicos o a la de implantación de nuevas tecnologías, acreditación de las herramientas, diseño de protocolos que ayuden a evitar sesgos de los datos, protocolos de buenas prácticas, etc. En definitiva, es necesaria su validación y posteriormente su vigilancia, porque la aplicación de una decisión guiada por un algoritmo va a tener muchas implicaciones y consecuencias desde el punto de vista legal y ético.

Una idea interesante para abordar este reto sería extender el juramento hipocrático<sup>m</sup> de los médicos a los ingenieros y desarrolladores de IA aplicada en salud, de manera que se aseguren los principios bioéticos de todos los profesionales que jugarán un papel relevante en la medicina del futuro; o la elaboración de los actuales códigos de conducta por el sector concernido. Esto es importante, porque tanto para el profesional como para el sistema de salud es imprescindible llegar a las conclusiones más óptimas con las mayores garantías.

## **GARANTIZAR QUE LA TOMA DE DECISIONES NO SEA AUTOMATIZADA O BASADA EXCLUSIVAMENTE EN PROPUESTAS DEL SISTEMA DE IA O EN PERFILES OBTENIDOS DE LOS PACIENTES**

La medicina basada en la evidencia supone un paradigma que ha favorecido la progresiva estandarización de las prácticas clínicas en las últimas décadas, como se ve

<sup>m</sup>Se trata de un texto ético que recoge las obligaciones morales de los médicos para con sus pacientes, con el objetivo de orientarlos en su labor profesional.

# INTELIGENCIA ARTIFICIAL EN SALUD: retos éticos y legales

en la adopción de protocolos y *guidelines* para el abordaje de diferentes patologías.<sup>13</sup> La tendencia a la **automatización del proceso de toma de decisiones basadas exclusivamente en propuestas algorítmicas de sistemas de inteligencia artificial**, implica cierto riesgo, ya que puede anular la iniciativa y la independencia en la toma de decisiones del profesional,<sup>11</sup> y favorecer, además, caer en la rutina profesional. Por otro lado, es necesario revalidar la fiabilidad de la conclusión del sistema de IA, no ya sólo por su finura tecnológica, sino también por los sesgos en los que puede incurrir su propuesta decisional, debido a las posibles incorrecciones o inexactitudes de los datos de base incorporados al sistema para su tratamiento correspondiente.

Por consiguiente, la toma de decisiones totalmente automatizadas no es legítima, tiene que mediar una intervención humana y en concreto una decisión de los profesionales. Es cierto que los profesionales se basan en protocolos médicos a la hora de decidir cómo proceder con un paciente con una enfermedad determinada, pero siempre emitiendo un juicio acorde con cada situación concreta. En este contexto, las herramientas de IA deben ser un eslabón de esos protocolos, nunca sustituyendo al médico, dando cabida a la atención asistencial personalizada o individualizada que propugna la Medicina Personalizada de Precisión.<sup>14</sup>

Las herramientas de IA permiten que los datos sean tratados con algoritmos de comparación y de correlación estadística, con el objetivo de crear perfiles que puedan utilizarse, por ejemplo, para una mejor clasificación de los pacientes.<sup>15</sup> Sin embargo, a pesar de que este tipo de estrategias puede resultar tremadamente útil en el campo de la medicina, cuando un **proceso de toma de decisiones está basado exclusivamente en procesos de automatización o en perfiles de pacientes** puede verse afectada la autonomía del paciente o el derecho a una explicación, en los términos señalados más arriba.<sup>11</sup> Para garantizar que se preservan los derechos fundamentales de los individuos ante estas prácticas, es necesario validar continuamente los perfiles generados con estas herramientas y los algoritmos subyacentes.<sup>15</sup>

Las herramientas de IA deben constituir para el médico elementos de ayuda y de simplificación de tareas, evitando que el profesional dedique su tiempo, por ejemplo, al desarrollo de tareas repetitivas, dejándole tiempo a

aquellas actividades en las que el profesional aporta valor. Estas herramientas son muy poderosas en tareas de clasificación y predicción, pero son tan complejas que resulta muy complicado extraer el conocimiento del modelo, y siempre cabe la posibilidad de error. Aquí es donde el médico tiene que usar información adicional para tomar la decisión que considere conveniente, puesto que será el que siempre tendrá la última palabra. Es decir, las herramientas de IA deben ir dirigidas al desarrollo de protocolos de trabajo o procedimientos en los que la colaboración entre la herramienta y el clínico suponga un salto de calidad significativo en la labor asistencial.

Será tarea del clínico asegurar que las decisiones sobre la atención clínica al paciente nunca estén basadas exclusivamente en perfiles o en las propuestas decisionales del sistema de IA o que sean automáticas, especialmente si se refieren a la adopción o no de un determinado tratamiento sin ofrecer a cambio una alternativa, incluso aunque sea menos eficiente. Esto es así porque no se dejaría al paciente la posibilidad de decidir sobre el tratamiento, podría incurrirse en prácticas arbitrarias, sin tener en cuenta otros elementos relevantes personales de toma de decisiones, o incluso, como se ha visto antes, podría llevar a prácticas discriminatorias, y que una persona sea excluida de un tratamiento o le sea prescrito otro con base en consideraciones no estrictamente orientadas a su salud.

Los crecientes procesos de automatización no pueden ser completamente independientes del ser humano, pues requieren vigilancia constante. En consecuencia, no deben excluir la especificidad de la relación entre médico y paciente, que atiende a las particularidades que pueda presentar este último. Esta relación es propia de una medicina humanizada, bien que digitalizada, que se viene reclamando insistentemente desde hace años. En este sentido, la máquina no puede reemplazar al ser humano en una relación que se construye a partir de la reunión de áreas complementarias de autonomía, competencia y responsabilidad.<sup>7</sup>

Estos nuevos escenarios requerirán de formación específica para los médicos, porque tendrán que conocer las posibles limitaciones de estas técnicas o, por ejemplo, el sesgo en la interpretación de los resultados; e, indudablemente, estar al tanto de los avances que se produzcan.<sup>7</sup>



## **ASEGURAR LA TRANSPARENCIA SOBRE LOS ELEMENTOS ESENCIALES Y SOBRE EL PROCESO POR EL CUAL UN SISTEMA DE IA LLEGA A UNA CONCLUSIÓN DECISORIA**

Una limitación que se encontrarán los profesionales sanitarios cuando empleen este tipo de sistemas es que, de manera inherente a ellos, son herramientas poco transparentes en el proceso decisional que siguen. Por un lado, porque muchas veces están amparados bajo el secreto profesional, y por otro, por la enorme complejidad del diseño del algoritmo. Normalmente, cuanto más complejo es el modelo, más complicado resulta para el usuario extraer información útil o interpretable del mismo. Los humanos no tenemos la capacidad de cómputo que permite al algoritmo procesar datos y aprender de ellos (generando nueva información y aprendiendo también de ella), por lo que con estas herramientas obtenemos resultados muy potentes, pero no sabemos cómo se ha llegado a ellos, funcionando como una “caja negra”.<sup>11</sup> Esta **opacidad en los elementos y en el proceso por el cual un sistema de IA ha llegado a una conclusión**, conlleva el riesgo de que el profesional de la salud no pueda validar y confirmar o descartar la propuesta del sistema porque desconoce la lógica interna por la cual se ha llegado a dicho resultado, como se indicó más arriba.<sup>11</sup>

El clínico (y el paciente) sentirá más seguridad respecto a la decisión clínica si puede entender la lógica detrás de estos sistemas de IA, más allá de que el resultado sea lo importante y condicione claramente la toma de decisiones. En la medida de lo posible, el conocer cómo funcionan estas herramientas por parte del profesional sanitario puede contribuir a solventar este problema, poniendo en relieve la importancia de incorporar a la formación del profesional conocimientos para la utilización de herramientas, entendiendo sus limitaciones y así poder transmitírselas al paciente. Sin embargo, hoy por hoy muy pocos profesionales sanitarios están capacitados para explicar los por menores del proceso algorítmico. En este sentido, dado que es responsabilidad del clínico informar al paciente de las decisiones tomadas con estos sistemas, no solo es necesario formar a los profesionales sanitarios, sino que, ligado al principio de transparencia, de cara al paciente será necesario incorporar consentimientos informados adecuados sobre el uso de estas herramientas.

Por otro lado, se está avanzado notablemente en el diseño de técnicas para extraer información y conocimiento de estos modelos tan complejos. Esta información es fundamental para conocer más en detalle la relación entre las variables de un modelo y el evento que se quiere predecir, o puede proporcionar nuevos datos en el futuro. De hecho, existen líneas de investigación muy relevantes en la actualidad dirigidas a solventar este aspecto, como la aplicación de técnicas de explicabilidad de modelos, es decir, que sea el propio sistema quien se “autoexplique”, y pueda informar al clínico (y al paciente) de los criterios que se han seguido para tomar una determinada decisión. Estas aproximaciones son muy relevantes, dadas las grandes implicaciones de los resultados en el área de la salud, y más concretamente, en la planificación de tratamientos o toma de decisiones relacionadas con la calidad de vida del paciente.

## **ESTABLECER A QUIÉN ATRIBUIR Y QUIÉN DEBE ASUMIR RESPONSABILIDADES A LA HORA DE EMPLEAR HERRAMIENTAS DE IA EN SALUD**

Como se ha comentado anteriormente, existe una serie de aspectos a tener en cuenta sobre la atribución de responsabilidades a la hora de emplear las herramientas basadas en IA en el campo de la salud. Por ello, se plantea la necesidad de establecer un marco ético común y sólido. De la opacidad de los sistemas de IA deriva la discusión sobre la responsabilidad en la toma de decisiones, si ha dado lugar al deterioro de la salud del paciente.<sup>11</sup> En el caso de que se llegue a un resultado erróneo hay que discernir si se trata de un fallo en el sistema o si el profesional mismo ha cometido el error, y por lo tanto recae sobre él la responsabilidad de los daños causados al paciente. Se trata de una cuestión compleja, porque la atribución de responsabilidades tiene que ver con quién toma una decisión sobre la salud de un paciente y con el correcto funcionamiento de los sistemas o herramientas “inteligentes” que se apliquen.

Según el sistema de responsabilidad vigente, el **médico tiene que asumir en primer lugar la responsabilidad sobre daños producidos en la salud del paciente**

**en caso de mal uso de estas herramientas, puesto que es él quien debe garantizar que la recomendación o la conclusión que se alcanza con IA sea razonable o que al menos no sea perjudicial para el paciente.** Es por tanto responsabilidad del clínico utilizar aquellos sistemas o herramientas que estén acreditados y aceptados para su aplicación clínica, pero no incurrirá en responsabilidad si el sistema funciona de manera incorrecta, en cuyo caso la responsabilidad se trasladará a quien haya diseñado el sistema, siempre y cuando no haya advertido del posible mal funcionamiento, o del encargado de su mantenimiento y actualización. En el caso de que el desarrollador haya advertido al usuario (el profesional de la salud) de los posibles fallos y haya sido totalmente transparente, hay que garantizar que alguien correrá con la responsabilidad, para lo que se ha planteado la implantación de una responsabilidad objetiva (siempre que no sea de naturaleza penal), que desde el punto de vista jurídico significa que el centro sanitario o el médico responden de los daños que se causen, independientemente o en lugar de la culpa personal que se pudiera imputar. Esto podría plantearse cuando los riesgos de daños sean desproporcionados, igual que ocurre con otros riesgos que se asumen en la atención sanitaria. Tampoco debe olvidarse la responsabilidad directa, asimismo de base objetiva, que pueden tener que asumir las administraciones sanitarias en el sistema público de salud como consecuencia de los actos de sus empleados.

En este sentido, cuando se produzca un fallo o mal funcionamiento evitable de la herramienta basada en IA, no cabe hablar de responsabilidad jurídica de las máquinas, pero sí de los desarrolladores y de los usuarios (los profesionales). Es responsabilidad de éstos tomar las medidas oportunas para conseguir que se realice una validación dinámica<sup>6</sup> del proceso o disponer de mecanismos para comprobar que no se producen disfunciones. Para evitar esto, es necesario que la validación se realice de dos maneras diferentes: una validación dinámica interna

del instrumento (validando que en su desarrollo no se han introducido sesgos que alteren su funcionamiento de acuerdo con los intereses o las inclinaciones de los programadores) y otra relativa a la precisión de su análisis y a la exactitud de sus predicciones, lo que muchas veces dependerá de la propia calidad de los datos utilizados en su confección.<sup>6</sup>

La obligación de que haya acreditaciones de estos procesos es semejante a la que debe obtener un producto sanitario para que pueda ser usado en la clínica. Este proceso de acreditación como tecnología sanitaria es importante y supondrá un reto, puesto que son tecnologías novedosas y muy complejas, lo cual puede entrañar cierta dificultad a la hora de determinar su naturaleza jurídica y el régimen de acreditación que, en consecuencia, corresponda. En relación con esto, hay que mencionar el Reglamento (UE) 2017/745 de 5 de abril de 2017 sobre los productos sanitarios, por el que se modifican la Directiva 2001/83/CE, el Reglamento (CE) nº 178/2002 y el Reglamento (CE) nº 1223/2009 y por el que se derogan las Directivas 90/385/CEE y 93/42/CEE del Consejo. Este reglamento entrará próximamente en vigor y se aplicará a las herramientas de IA concebidas como software que se identifiquen con el concepto de producto sanitario.

En cuanto a la utilización de herramientas de IA en investigación, el sistema de gestión de riesgos y de responsabilidad es diferente, puesto que en este ámbito, en principio, no se van a tomar decisiones clínicas, salvo que si se trata de un ensayo clínico se encuentre en la fase de experimentación sobre seres humanos. Esto implica también que las conclusiones o perfiles que se generen con algoritmos no estarán necesariamente validados por un profesional, por lo que en este campo estaría permitida en algunos casos la opción de tomar decisiones de manera automatizada. El procedimiento consiste en que la Agencia Española de Protección de Datos debe evaluar el nivel de riesgo al que se somete a los sujetos si se emplean estas herramientas.



# CONCLUSIONES Y RECOMENDACIONES

---

La IA se está convirtiendo en una valiosa herramienta de trabajo en numerosas actividades humanas y en particular en el sector de la salud, tanto en la asistencia clínica como en la investigación biomédica, por lo que será necesario plantearse las oportunidades y los riesgos que la tecnología implica. En este sentido, es importante tomar en serio los desafíos éticos asociados con la IA y no verlos como meros obstáculos a la innovación, sino como algo que puede estimular y guiar el desarrollo de aplicaciones que fomenten objetivos y valores comunes.<sup>10</sup>

Estas herramientas de IA deben constituir para el médico elementos de ayuda y de simplificación de tareas, evitando que el profesional dedique su tiempo, por ejemplo, al desarrollo de tareas repetitivas, dejándole tiempo a aquellas actividades en las que el profesional aporta valor. Con el objetivo de fomentar la traslación e implantación en el campo de la salud de estas herramientas, se proponen a continuación las siguientes recomendaciones:

- **Los sistemas de IA**, incluso los más avanzados, podrán ser instrumentos de ayuda en la toma de decisiones del profesional sanitario, sin llegar nunca a reemplazarlo, en una relación que se construye a partir de la confluencia de áreas complementarias de autonomía, competencia y responsabilidad, que deben continuar presidiendo la relación médico – paciente.
- Será necesario formar a los profesionales en este ámbito de conocimiento, además de adaptar el contenido de especialidades médicas.
- Las decisiones que tomen los profesionales en el ámbito asistencial podrán apoyarse en las conclusiones o propuestas de los sistemas de IA, pero deberán evitar adoptarlas de manera rutinaria, automática o a partir exclusivamente de perfiles de los pacientes o de su estratificación, sin atender a cada situación individual, es decir, en el marco de las propuestas de la Medicina Personalizada de Precisión.
- Es necesario ocuparse de la “alimentación” o “entrenamiento” correctos de la IA. En este campo el usuario conserva una función central. En este sentido no podemos – hasta la fecha – hablar todavía de “autonomía de decisión” de la máquina. Cuando esto ocurra, habrá reflexionar sobre nuevos desafíos en todos los sectores en los que la IA pueda llegar a desempeñar un papel esencial, incluyendo la actividad sanitaria asistencial.
- Seguir avanzando en la seguridad y en la protección de los datos. Realizar estudios de impacto del uso de los datos personales en la privacidad de los sujetos fuente para tratar que sea mínimo, e implementar medidas que aseguren que se cumpla lo estipulado en la normativa sobre protección de datos. Por ejemplo, a través del diseño de procedimientos y algoritmos que permitan deidentificar<sup>11</sup> o anonimizar los datos con ciertas garantías que permitan salvaguardar el derecho a la privacidad de los pacientes.
- Es necesario diseñar mecanismos de validación y de formación a los miembros de los comités de ética de la investigación y de ética asistencial en materia de IA para que puedan evaluar estas herramientas de una forma rigurosa.

<sup>10</sup>Término que hace referencia al proceso de eliminar de los datos aquella información que permita identificar al sujeto fuente.



## INTELIGENCIA ARTIFICIAL EN SALUD: retos éticos y legales

- Deberá realizarse una reflexión sobre cómo se deben proteger o comercializar estas herramientas desde el punto de vista del creador, por ejemplo, a través de patentes o mecanismos de secreto empresarial ya que, al tratarse de algoritmos integrados en un sistema o dispositivo con una aplicación práctica, son susceptibles de ser protegidos.
- Es necesaria la elaboración de documentos oficiales que de manera clara y práctica proyecten el marco legal a la realidad más concreta a través de guías de actuación, documentos de resolución de dudas frecuentes, protocolos, modelos estandarizados de homologación, validación y asunción de responsabilidades concretas, etc.
- Existe la necesidad de elaborar **códigos de conducta internacionales** que amparen las bases de datos, de manera que se facilite el compartir y acceder a la información que “alimente” este tipo de herramientas, optimizando los recursos y que a la vez sea garantista respecto a la protección de los derechos de los sujetos.
- Introducir una normativa con guías de desarrollo de las herramientas de IA, de manera que quede explicitado cómo deben construirse para evitar sesgos o fallos del sistema.
- Garantizar el principio de transparencia y de información para minimizar las posibles reticencias a las herramientas de IA que surgen de la percepción de no conocer cómo toman las decisiones, en este caso, sobre la salud de las personas. Para ello será imprescindible potenciar la difusión de información y de transparencia a la hora de aplicar estas herramientas.



## BIBLIOGRAFÍA

1. Del Río Solá M, López Santos J, C VP. La inteligencia artificial en el ámbito médico. *Rev Española Investig Quirúrgicas.* 2018;XXI(3):113-116.
2. Obermeyer Z, Emanuel E. Predicting the Future – Big Data, Machine Learning, and Clinical Medicine. *Physiol Behav.* 2016;375(13):1216-1219. doi:10.1016/j.physbeh.2017.03.040
3. Cotino L. Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales. *Big Data and Artificial Intelligence. An Approach from a Legal Point of View about Fundamental Rights. Dilemata.* 2017;24(9):131-150.
4. Romeo-Casabona CM. Intelligenza artificiale, salute e Diritto. Presented at the Comitato Nazionale per la Boetica: Roma; 2019.
5. Romeo-Casabona CM, Nicolás P, de Miguel I. Retos Éticos y Necesidades Normativas En La Actividad Asistencial En Medicina Personalizada de Precisión. 2018.
6. Beriain I de M. Medicina personalizada, algoritmos predictivos y utilización de sistemas de decisión automatizados en asistencia sanitaria. Problemas éticos. *Rev Int Éticas Apl.* 2019;30:93-109.
7. Presidenza del Consiglio dei Ministri. Comitato Nazionale per La Bioética e Comitato Nazionale per La Biosicurezza, Le Biotecnologie e Le Scienze Della Vita, Intelligenza Artificiale e Medicina: Aspetti Etici.; 2020:1-17.
8. Romeo-Casabona CM, Lazcoz G. Inteligencia Artificial aplicada a la salud: ¿qué marco jurídico? In: Sánchez Caro J, Abellán F, eds. *E-Salud y Cambio Del Modelo Sanitario.* Madrid: Fundación Merk; 2020:60-100.
9. European Group on Ethics in Science and New Technologies. *Artificial Intelligence, Robotics and 'Autonomous' Systems.*, Brussels; 2020.
10. The Swedish National Council on Medical Ethics. *Artificial Intelligence in Healthcare.*; 2020.
11. Abellán-García F, Alcalde G, Alfonso I, et al. Uso de datos clínicos ante nuevos escenarios tecnológicos y científicos. Oportunidades e implicaciones jurídicas. *Rev Derecho y Genoma Humano Genética, Biotecnol y Med Av.* 2019;Número Ext(41):1-646. <https://dialnet.unirioja.es/servlet/articulo?codigo=969804>.
12. Floridi L. What the Near Future of Artificial Intelligence Could Be. *Philos Technol.* 2019;32(1):1-15. doi:10.1007/s13347-019-00345-y
13. Perin A. Estandarización y automatización en medicina: El deber de cuidado del profesional entre la legítima confianza y la debida prudencia. *Rev Chil Derecho y Tecnl.* 2019;8(1):3-28. doi:10.5354/0719-2584.2019.52560
14. Romeo C, Nicolás P, De Miguel I. Retos Éticos y Necesidades Normativas En La Actividad Asistencial En La Medicina Personalizada de Precisión. Madrid; 2018.
15. Garriga Domínguez A. La elaboración de perfiles y su impacto en los derechos fundamentales. Una primera aproximación a su regulación en el reglamento general de protección de datos de la Unión Europea. *Derechos y Lib Rev del Inst Bartolomé las Casas.* 2018;(38):107-139. doi:10.14679/1058

## NOTAS

## NOTAS



## NOTAS





[www.institutoroche.es](http://www.institutoroche.es)