



# 03/

## Retos de la introducción

de la inteligencia  
artificial en la  
relación asistencial.  
¿Hacia una medicina  
sin rostro?

### **Bernabé Robles del Olmo,**

Neurólogo y máster en Bioética. Coordinador Unidad de Bioética.  
Parc Sanitari Sant Joan de Déu. Sant Boi de Llobregat (Barcelona).  
Vicepresidente Comité de Ética de la Investigación con Medicamentos (CEIM).  
Fundació de Recerca Sant Joan de Déu. Eplugues de Llobregat (Barcelona).

Cuando se aborda el tema de la potencial deshumanización de la asistencia sanitaria, una lectura superficial nos puede llevar a pensar que “todo es culpa de la inteligencia artificial (IA)”. La primera reflexión a plantear sería, por tanto, si la digitalización de la medicina es la causa preponderante, o detonante, de la pérdida progresiva de calidez y calidad en la relación clínica. Pero es posible que la respuesta no sea tan simple, y que el reto más “inteligente” sea imaginar cómo, con qué requisitos, puede ayudarnos la IA a recuperar la calidez y la calidad en el trato.

*Palabras clave: Inteligencia Artificial, Humanización, Asistencia.*

When reflecting on the potential dehumanisation of health care, a cursory consideration of the subject may lead one to think that “artificial intelligence (AI) is to blame for it all”. Accordingly, the first aspect to be weighed is whether the digitalisation of medicine is actually the chief cause or mainspring of the progressive loss of warmth and quality in the clinical relationship. However, the answer to this question may not be as simple as it may seem, and the most “intelligent” challenge may well be to imagine how and with what prerequisites AI can help us to recover the warmth and quality of treatment.

*Key words: Artificial Intelligence, Humanisation, Care.*

---

## 1/

---

### ¿Humanizar? Calidad y calidez.

Cuando hablamos de humanizar el cuidado de la salud, nos referimos a un trato cálido, respetuoso y de calidad. Pero el hecho de que un acto o una decisión sea producto de humanos no es garantía de trato “**humanizado**”. ¿Todo acto humano es deseable, incluso tolerable? Si repasamos la historia de los Sapiens y, desafortunadamente, nuestro presente, cuesta contestar con un sí categórico.

La “**humanización**” a cultivar en la atención a la salud es aquella que mejora nuestra hospitalidad, nuestro interés por el otro y por lo otro, por ese mundo irrepetible que tenemos delante. Hablamos pues de respeto y responsabilidad, de aquel trato que todos querríamos recibir cuando sufrimos. Se trataría de no dejar nunca de ser sujeto para convertirnos en simple objeto, o peor, en dato. Desearíamos ser, cuando desgraciadamente nos toque, “**persona enferma**”, y no mera “**enfermedad**”, sin mirada.

Se trata pues de construir confianza en la alteridad, viendo al otro como alguien que, entre otras muchas otras dimensiones de su persona, padece y necesita ayuda. Esto no nos debería resultar difícil, dado que estamos neurobiológi-

camente cableados para ello, con el sistema de neuronas espejo ([Rizzolatti & Cattaneo, 2009](#)).

La pregunta que queremos abordar sería pues: ¿puede alinearse la IA con un cuidado humanizado? Achacar la deshumanización completamente a la IA sería seguramente un análisis simplista y poco introspectivo. La atención sanitaria comenzó a “**deshumanizarse**” mucho antes de la implantación acelerada de los algoritmos y de los detectores de patrones en la práctica habitual.

---

## 2/

---

### Una breve historia de la relación asistencial.

Las actividades de cuidado entre humanos seguían inicialmente un modelo básicamente espiritual, orientado a confortar ante el sufrimiento y la incertidumbre, dado que nuestro margen de maniobra era muy limitado cuando se trataba de reconducir fenómenos naturales. Las funciones de chamán y curandero coincidían a menudo en la misma persona.

Durante milenios, probablemente hasta finales del siglo XIX, el paradigma no cambió demasiado. La atención psicológica, espiritual y paliativa no son inventos recientes. Salvo alguna intervención resolutiva de cirugía menor en la superficie del cuerpo (no se disponía de anestesia general), y la relevante tarea de tranquilizar cuando la patología reconocida tenía una esperable buena evolución espontánea, la atención paliativa (láudano y acompañamiento a pie de cama) era el modelo asistencial habitual cuando las cosas no iban bien.

Con la aparición de antisépticos, anestésicos generales y antibióticos, el poder de intervenir sobre el cuerpo se incrementó exponencialmente.

Como el cuerpo era más abordable, y más explo-  
rable por los avances paralelos del conocimiento  
científico, se impuso un modelo mal llamado  
biomédico. Quizá deberíamos llamarlo más bien  
mecanicista, dado que entendía la Medicina  
como un servicio de reparación de cuerpos.  
El poder que visualizó entonces la sociedad en  
las ciencias médicas se hizo tan intenso que ya  
no íbamos nosotros a ver a los enfermos, sino  
que venían ellos a vernos a templos de salud que  
construimos para concentrar los cuidados, los  
profesionales y la tecnología necesaria para pro-  
veerlos. En este contexto, la mesa de la consulta  
se convirtió en instrumento, pero también en  
una primera barrera, porque marcaba ya una  
distancia en el trato interpersonal. Luego, poco  
a poco, llenamos la mesa y sus inmediaciones de  
aparatos (fonendo, negatoscopio, ecografía, te-  
léfono, ordenador, móvil, etc.). Todos ellos dis-  
traían la mirada y, potencialmente, interrumpían  
el encuentro interpersonal que suponía, en  
origen, la relación clínica.

Además, las necesidades de soporte de las per-  
sonas atendidas y las obligaciones formativas de  
un sistema “de éxito”, comenzaron a llenar tam-  
bién la sala de gente (profesionales, estudiantes  
y acompañantes), y de nuevas interrupciones.  
El acto médico estaba dejando de ser un acto  
íntimo, de confianza. La tecnología reducía  
el contacto físico interpersonal, desactivando  
nuestra red neuronal de la empatía. Así fuimos  
perdiendo, unos más que otros, la mirada.

Hay que reconocer que el modelo biomédico ha  
aportado muchos avances y ha mitigado mucho  
sufrimiento físico, pero también ha generado  
carencias complejas en un modelo asistencial  
que debería enfocarse más en personas y valores  
que en puras averías. Ante estas insuficiencias  
se ha propuesto en el siglo XX el modelo biop-  
sicosocial o, incluso, un modelo estrictamente  
social, centrado en adaptar el entorno social a  
la enfermedad, la discapacidad o el sufrimiento.

Fruto del aumento de la esperanza de vida deriva-  
do la mejora de las condiciones sociales en ciertas  
partes del globo, y del abordaje más resolutivo de

las patologías agudas, el crecimiento de los padeci-  
mientos crónicos ha hecho patente la necesidad de  
rediseñar el modelo hacia lo bio-psico-socio-espi-  
ritual. Volvemos, de alguna forma, a los orígenes.  
Ahora queremos abandonar, cuando menos par-  
cialmente, las factorías de salud que construimos  
durante los siglos XIX y XX para volver a cuidar  
de cerca a las personas crónicamente enfermas, ya  
no totalmente curables.

Por otro lado, la conquista social incuestionable  
que supuso, aunque desgraciadamente no en  
todo el globo, la universalización de la atención  
sanitaria (deseada, justa y necesaria) generó un  
volumen de actividad, tanto asistencial directa  
como de soporte, que derivó en oportunidades  
de negocio y de empleo. La salud comenzó a  
convertirse en un “sector económico” haciendo  
confluir intereses múltiples y diversos (comer-  
ciales, culturales, industriales, administrativos,  
curriculares, científicos, mediáticos, laborales,  
etc.) sobre aquello que parecía tan simple en su  
esencia: una relación de ayuda a quien sufre.

Las relaciones cuasicontractuales así generadas  
compartimentan las tareas asistenciales en már-  
genes poco líquidos y, por tanto, no te impli-  
can en lo íntimo y personal de cada persona, y  
pueden acabar desembocando en estrategias de  
medicina defensiva, viendo en la persona aten-  
dida una potencial amenaza. Una relación así  
no te traspasa.

Es el falso confort (o el falso consuelo) de lo im-  
personal. Vamos entendiendo poco a poco que  
el mito tecno-científico en Medicina no es la  
panacea, y que sobrevivir a una patología aguda  
con una intervención técnica brillante genera,  
afortunadamente, más trayectoria vital pero  
también más oportunidades para padecer en-  
fermedades, condiciones o trastornos no total-  
mente curables, solo paliables. La cronicidad es,  
de alguna forma, el resultado de éxitos sociales  
y biomédicos, en aquellas partes del globo que  
se los han podido permitir. Vamos entendiendo  
que medicalizar la sociedad no será el remedio  
para todo mal, y que esta medicalización tiene  
sus propios riesgos y efectos adversos, siendo el

más evidente el de generar sufrimiento por el propio miedo a sufrir y por la hipervigilancia del estado de salud.

Además, profesionales, pacientes o ciudadanos pueden quedar fácilmente deslumbrados por la técnica, esperando de ella soluciones que no son esperables, ni desde el punto de vista científico ni desde la perspectiva humana. Se aplican recetas tecnológicas sin una justificación sólida y sin una reflexión pausada y rigurosa sobre pros y contras. Por si esto fuera poco, se evalúan poco los resultados obtenidos a medio y largo plazo.

Se impone también una reflexión sobre la dimensión que merece el “**diagnóstico**”. Diagnosticar significa “**conocer a través de**”. Quizá lo hemos convertido erróneamente en un fin absoluto, en el marco de la tecnificación en el modelo exclusivamente mecanicista que impulsaron los avances en el conocimiento del funcionamiento del cuerpo. Diagnosticar debería volver a ser lo que fue, un instrumento para interpretar y mitigar el sufrimiento vivido por las personas que nos piden ayuda. Si el diagnóstico no aporta más que una etiqueta clasificadora y despersonalizadora, pierde su sentido, y deshumaniza.

Por supuesto, indicadores y “**biomarcadores**”, bien diseñados y rigurosos con la estadística son excelentes herramientas de gestión y de decisión. Pero debemos estar seguros de no malinterpretarlos. Nuestra tarea va de historias personales y sociales, todas diferentes y todas con condicionantes, contexto y excepciones. ¿Dónde queda la persona, y su derecho a rechazar una injerencia en su trayectoria vital, en su cuerpo o en su espíritu? Seguramente la “**atención centrada en personas**” debería ir precisamente de eso: de conjugar eficiencia y atención personalizada, ciencia y respeto, obligaciones y comunicación. Asumido este marco, la IA sería seguramente una aliada para una atención justa, cálida y de calidad. Fuera de este marco será una excusa más para no dirigir la mirada.

En este sentido, el “**datafundamentalismo**” de este nuevo paradigma puede llevarnos a con-

fundir los perfiles que representan a las personas con las personas mismas, que pasarían de seres humanos a códigos o categorías. ¿Es esto lo que debería perseguir una auténtica medicina humanizada? Aquellos que atendemos y los que somos atendidos, debemos huir de la instrumentalización del otro en la relación asistencial. Todos tenemos nuestra parcela de responsabilidad. Si no cuidamos desaparecerá la confianza, que es la bisagra esencial de cualquier relación de ayuda, de cualquier relación social, al fin y al cabo. Sin confianza, la ayuda se convierte en servicio neutro, sin piel, sin cara. Al fin y al cabo, el cuidado de la salud, por complejo que haya devenido, supone una constelación de relaciones interpersonales.

En este contexto es en el que llama a la puerta la IA, una nueva y fascinante tecnología. Pero ya estábamos perdiendo la mirada en la atención sanitaria antes de la expansión de las herramientas digitales y de las decisiones asistidas por algoritmos.

---

## 3/

---

### IA y conocimiento: un nuevo paradigma.

La IA reta al método científico tradicional, que buscaba contrastar hipótesis. La relación causa-efecto, esencial en ciencia desde el siglo XVI, no es tan relevante cuando utilizas IA. Ésta, al menos por el momento, ni busca causas, ni busca sentido, ni explica. Su territorio es el de las correlaciones. Por ello, muchas instituciones o empresas que priorizan el pragmatismo y la rentabilidad ya se han adherido hace tiempo sus habilidades predictivas (por ejemplo, bancos y aseguradoras). Lo importante es si la persona devolverá el préstamo o no, no las razones por las que ocurrirá esto.

Es fundamental ofrecer un trato cálido y respetuoso, centrado en la persona y no únicamente en su enfermedad, para preservar la empatía y la confianza en la relación asistencial

Y este nuevo paradigma ya se está imponiendo también en Medicina. Es la Medicina de las “P”: predictiva, participativa, preventiva, personalizada y de precisión, que ha encontrado un catalizador en la IA y en todas las técnicas de análisis de datos masivos. Aspiramos ya a detectar trastornos y/o riesgos para la salud sin necesidad de que los síntomas nos avise. Queremos preservar la salud más que atender a la enfermedad. Queremos “vacunarnos” contra toda enfermedad. Incluso el término “paciente” se tensiona en el nuevo modelo.

En esencia, la IA no es más que una rama de la informática que cataliza las potencialidades de la estadística, sobre todo de las probabilidades condicionadas bayesianas. Diríamos que es una tecnología de la información “especial”, que hace algo más que informar, porque es capaz de transformar la información en predicciones, avisos o recomendaciones.

No es tan nueva como pensamos. Los pensadores medievales, como por ejemplo Ramon Llull, ya asumieron la posibilidad de que ciertas funciones mentales podrían ser ejercidas por entidades diferentes al cerebro humano, y constan intentos ya decididos de construir máquinas pensantes desde principios del siglo XX y, sobre todo desde la II Guerra Mundial.

Sin embargo, la revolución digital ha posibilitado que los proyectos relacionados con esta área del conocimiento y de la técnica se hayan incrementado de manera más que exponencial en los últimos 30 años. ¿Por qué? Por las mejoras en la infraestructura computacional que han permitido el desarrollo de herramientas de gestión de datos masivos y el desarrollo de un “aprendizaje profundo” reinventado gracias, por ejemplo, a las mal llamadas “redes neuronales” artificiales (mejor “redes de microperceptrones”)

---

## 4/

---

### El vértigo de la sustitución de la mente.

---

#### 4/1

---

#### ¿Qué es “artificial”?

Podríamos pensar que el adjetivo “artificial” no es afortunado. Al fin y al cabo, los sistemas de IA no son más que aparatos hechos con materiales naturales por individuos naturales. Quizá utilizamos “artificial” cuando algo aparece después de una decisión de búsqueda, de una invención. Existen animales que fabrican elementos que no existían (termiteros, diques de castores, etc.), pero quizá la diferencia está, como decíamos, en la intención. Por tanto, desde el punto de vista moral, no está mal conservar el apellido “artificial” para recordarnos que, en el fondo, existe una responsabilidad humana detrás del desarrollo de estas técnicas y, más allá, detrás de las preguntas que les hacemos y de los datos que facilitamos a sus algoritmos.

Buscar responsabilidad en la máquina, al menos de momento, no se puede plantear. Los algoritmos no son más que fórmulas, complejas, pero fórmulas. Cuando hablamos de “caja negra” en la IA no lo hacemos porque no conocemos la fórmula que estructuran su funcionamiento ni porque no conozcamos las ramas y bifurcaciones de sus algoritmos.

Hablamos de “caja negra” porque llega un momento en que no podemos controlar el flujo de los datos y resultados a través de dicha estructura lógico-matemática. Pero tampoco debemos olvidar, si nos preocupan las cajas negras, que en estos momentos hay casi 8 millones de cajas negras en la biosfera, un número equivalente a la cantidad de cerebros humanos funcionando.

---

## 4/2

---

### El eterno problema de definir la inteligencia.

Sobre el término inteligencia (etimológicamente, “**leer entre**”), también debemos reconocer que muchas herramientas de IA actualmente disponibles son claramente superiores a nosotros en esa tarea: “**distinguir patrones de correlación**”. Hablamos de inteligencia cuando distinguimos lo no aparente, cuando interpretamos la realidad y la percepción para resolver problemas. Las herramientas de IA detectan patrones y correlaciones entre fenómenos y datos no accesibles al cerebro humano. Por eso nos parecen una ayuda interesante.

Además, a menudo son capaces de detectar estas correlaciones “**en tiempo real**” tendiendo a cero el tiempo requerido para los antes tediosos análisis estadísticos. En este punto, no obstante, es obligado recordar de nuevo que la IA detecta cosas en las que las personas no pueden reparar, pero con unas “**gafas**” que diseñan, como mínimo en su origen, las personas.

Ahora bien, los sistemas IA no conocen el sentido de las correlaciones que detectan, ni les importa. Tampoco sienten vergüenza. En el fondo, todavía “**no saben lo que hacen**”. Así pues, la IA no pretende explicar el funcionamiento del cerebro, ni necesariamente imitarlo. En principio, no habría que esperar que se comporte como un humano. Aunque usemos la palabra inglesa “**learning**”, un algoritmo ni aprende ni es creativo. Estas son cualidades antropocéntricamente proyectadas. Podríamos definir un espectro dimensional, según la complejidad de datos manejados que iría desde la estadística descriptiva clásica, pasando por el aprendizaje automático y el aprendizaje profundo hasta los sistemas de IA autónomos sobre los que actualmente existe una frenética actividad de investigación. Ésta es una limitación relevante si hablamos de suplir el rol de las personas en una hipotética relación clínica con algoritmos.

El vértigo antropológico que genera la posibilidad de ser sustituidos también influye en nuestras valoraciones acerca de la introducción de la IA en la relación clínica. Decía Warner Slack (1997), pionero de la informática médica, que si un médico podía ser sustituido por IA es que se lo merecía.

Sentimos un vértigo similar en el Neolítico, cuando descubrimos que los animales hacían mucho mejor, tareas agrícolas o de transporte que nosotros y, más tarde, en el siglo XVIII, cuando fueron máquinas transformadoras de energía las que comenzaron a asumir tareas que antes hacíamos humanos o animales, con mayor rendimiento, precisión y resistencia.

La diferencia de base entre inteligencia humana y artificial es que la IA se basa en algoritmos matemáticos y programas informáticos, mientras que la humana se basa en procesos biológicos y fisicoquímicos. El problema es que ahora ciertas máquinas opositan a asumir funciones mentales. Y los humanos han hecho radicar el concepto de persona, con mayor o menor grado de consenso, en dichas funciones. Por tanto, la pregunta está servida: ¿La IA nos despersonalizará? ¿La IA amenaza nuestra especie?

---

## 5/

---

### Oportunidades e IA.

Como decíamos, podemos asumir ya a día de hoy que, ciertas herramientas de IA son mejores que nosotros en la velocidad de análisis de datos masivos, así como en la detección automática de correlaciones, tendencias y patrones. Por tanto, son solventes para clasificar, revisar, reconocer o predecir. La diferencia con la inteligencia humana se nota especialmente cuando hay infinitas columnas de variables y pocas observaciones.

Y no podemos negar que estas cualidades ofrecen grandes oportunidades en diferentes campos de la atención a la salud

(Pujol Guanyabens, 2023):

- Puede permitir una atención clínica, médica o quirúrgica, más segura y rigurosa, que maneje resultados en tiempo real y que disponga de ayudas casi instantáneas al razonamiento clínico, personalizando tratamiento y estableciendo pronósticos con mayor precisión. Los campos más tecnificados de la atención (cirugía, anestesia, cuidados intensivos, prótesis, etc.) seguramente ya han sido los primeros en recibir estas aportaciones. Por tanto, la IA da soporte a decisiones y a procedimientos sanitarios, automatizando tareas para incrementar su seguridad, exactitud y reproducibilidad. Permite una gestión más eficiente de recursos porque puede evaluar resultados favorables y desfavorables casi en tiempo real. Sin embargo, todavía no admitimos decisiones opacas de un algoritmo sobre una persona, sin supervisión de otra persona.

De todas maneras, para ser efectiva, esta “supervisión” debería iniciarse ya en el diseño de la herramienta. Una vez diseñadas, cuando abordamos decisiones difíciles con amplios márgenes de incertidumbre, resulta tentador “obedecer” la sugerencia de un elemento inanimado que tendemos a asumir como más exacto, neutral y exento de conflictos de interés, premisas todas cuando menos debatibles al tratarse de productos humanos. Comenzamos a disponer ya de publicaciones que muestran la capacidad de la IA generativa para realizar diagnóstico. Por ejemplo, en 2022, en 30 pacientes tratados en Urgencias en un hospital, se compararon las listas de diagnóstico diferencial hechas por médicos con las que sugería **chatGPT**. Observaron una gran superposición (alrededor del 60%) entre ambas. Los médicos humanos incluyeron el diagnóstico correcto entre los 5 primeros en el 87% de los casos mientras que la versión gratuita de ChatGPT 3.5 lo hizo en el 97% de los casos (Berg et al., 2023). Seguramente el camino más adecuado

en este campo será generar una cultura de decisiones híbridas **persona-IA** que resulten eficientes y seguras a un tiempo, considerando valores.

- Aunque su introducción es más pujante en la asistencia directa, quizá su impacto puede ser aún mayor en Salud Pública, dada su capacidad para el manejo de datos masivos y para detectar patrones. Predicciones, tendencias y/o monitorización de decisiones son instrumentos clásicos de los salubristas. Y la predicción sobre colectivos suele ser más potente que cuando se intenta anticipar lo que sucederá con un individuo concreto. Hemos visto estas aportaciones, con mayor o menor rigor, en tiempos de pandemia.

- Derivado de todo ello, cuesta poco imaginar oportunidades en el campo del diseño de políticas sanitarias, que no pueden perder nunca el foco colectivo y que necesitan también predicción de patrones y tendencias, así como cierta “neutralidad” en los procedimientos para preservar la justicia en el mayor grado posible. Ofrece a los responsables información sobre escenarios de futuro que faciliten decisiones sobre prioridades, reorganización y planificación en el sistema de salud, y sobre la gestión de las instituciones sociales y sanitarias. Además, como se trabaja con datos agregados y con objetivos ligados al bien común, los riesgos y restricciones relacionados con privacidad individual representarían un obstáculo menor.

- Por supuesto, también la investigación se puede ver beneficiada, especialmente en aquellos estudios que llamamos “observacionales”, porque la IA no tiene por qué observar muestras pequeñas y luego extrapolar, sino que puede buscar patrones, tendencias y correlaciones en la inmensidad de datos disponibles en el mundo real. Además, ya hemos dicho que puede detectar todo eso con mayor sensibilidad (seguramente no, por el momento, con mayor especificidad) que los investigadores de carne y hueso. También la gestión de datos en ensayos clínicos, el diseño de fármacos y el estudio de

las características y comportamiento de las proteínas son otros campos que ya están recibiendo beneficios de las técnicas de IA.

- También nos permite evaluar datos del mundo real sobre resultados positivos y negativos de nuestras intervenciones; terapias, programas asistenciales, cribados, nuevas tecnologías. Por tanto, puede ser una herramienta muy valiosa para preservar la calidad asistencial y la sostenibilidad del sistema sanitario y/o social.

- Por supuesto, cuesta poco imaginar el impacto potencial de la IA, sobre todo la generativa, en la educación médica, facilitando mecanismos de entrenamiento de los profesionales en formación que no supongan riesgos para personas: simulación, generación de “casos”, herramientas de evaluación, etc.

No obstante, el sistema nervioso central, es todavía, hoy por hoy, “mejor” que la IA, detectando valores y prioridades humanas, aplicando la solidaridad y la empatía. Es más flexible, tiene capacidad de improvisar. Siente vergüenza, como mecanismo de autocontrol. Además, parece que “se da cuenta” de lo que hace, aunque a menudo no lo parezca.

Decía Pablo Picasso que los ordenadores solo dan respuestas, no se hacen preguntas, aunque los expertos discuten en la actualidad si la IA no podrá llegar a hacer preguntas originales en el futuro. A parte de la curiosidad y la implicación, el ser humano tiene también otras armas valiosas, pero de “doble filo” si aspiramos a preservar la equidad y la justicia: la subjetividad, la empatía y la compasión.

---

## 6/

---

### Riesgos e IA.

---

#### 6/1

---

#### ¿Es el “aprendizaje profundo” una amenaza?

De forma muy simplista, podríamos decir que algunos sistemas algorítmicos de IA son capaces de hacer evolucionar sus respuestas finales en función de la experiencia acumulada por los datos que van recibiendo. Podríamos decir pues que las respuestas de estos sistemas se van construyendo “desde dentro”, y esto puede generarnos dos impresiones:

- Autonomía: puede parecer que “decide la máquina” o, incluso, que podría llegar a ser, en algún momento del desarrollo tecnológico, autoconsciente como nosotros y, más allá, engañar o mentir como nosotros.

- Misterio u opacidad, como se quiera: Es el fenómeno de la “caja negra” antes comentado. Conocemos la fórmula, pero no controlamos del todo el flujo de datos dentro de la estructura algorítmica. Los humanos no acabamos de comprender, en ocasiones, cómo el sistema ha elaborado una predicción o un consejo. Es una fórmula tan complicada que pocos le pueden dar un significado, e incluso sus diseñadores en ocasiones no entienden el flujo de datos. Es una limitación humana, no necesariamente un “defecto” del sistema. No olvidemos nunca que la estructura lógica de fórmulas de un sistema de IA es transparente y la ha escrito un humano, o un equipo de humanos.

La literatura de ficción ha planteado repetidamente este problema, centrado en la pregunta: ¿Podrá la creación controlar o superar al creador? El ordenador Hal 900 de “2001, una odi-

Aunque la IA y los avances tecnológicos aportan beneficios, no deben reemplazar la interacción humana, ya que la relación interpersonal sigue siendo clave en el cuidado de la salud

sea del espacio” sería un buen ejemplo. Sin embargo, no debemos olvidar que, hasta la fecha, el “aprendizaje profundo” no aprende nada. Aprender es una metáfora humana para entender su funcionamiento.

En el fondo, estamos hablando de automatizar procesos, muchos de ellos responsabilidad de nuestra mente hasta hace poco tiempo. De hecho, todo esto ocurre ya desde hace mucho tiempo, mucho antes de que la revolución informática y de las tecnologías de la comunicación de los últimos 30 años.

Las calculadoras científicas ya hacían operaciones matemáticas con mucha mayor velocidad, exactitud y rapidez que la mente humana en los años 70. Como ahora esta “suplencia” de funciones mentales afecta a muchas otras funciones (lenguaje, lectura, comprensión de idiomas, predicciones, pronósticos, apoyo a las decisiones, etc.), el reto al propio concepto de persona se hace más patente, y genera mayor inquietud.

Es importante no olvidar que, de momento, la IA, incluso la generativa, sólo hace, no explica. Sería como la “calculadora” a la aritmética. Por eso la IA generativa siempre aborda las preguntas, por absurdas que sean. Completa agujeros, más que buscar la “verdad”. Su atractivo radica en que se puede usar sin conocimientos informáticos (del mismo modo que una calculadora sin conocimientos matemáticos o un vehículo sin conocimientos de mecánica).

---

## 6/2

---

### El problema de los sesgos.

Los sesgos no son un problema inherente a los algoritmos. Los sesgos son nuestro problema como humanos, fruto de nuestra subjetividad, nuestra educación y nuestros instintos. El problema es que algoritmos y procesamiento de datos masivos pueden generar una ilusión de veracidad global generada a partir de datos no adecuados o no representativos, o de preguntas

con prejuicios y/o discriminación inherente. De hecho, los Sapiens tenemos muchos defectos cuando aplicamos y repartimos cuidados. A menudo nos centramos, sin darnos apenas cuenta, en solucionar nuestro propio conflicto interno empático, es decir, nuestro sufrimiento emocional al contemplar el sufrimiento del otro. Actuamos a menudo para mitigar dicho conflicto emocional interno individual, olvidando en mayor o menor grado el verdadero foco del cuidado: el sufrimiento del otro visto desde su propio prisma.

El problema de la representatividad de los datos que manejan los algoritmos se pone de manifiesto cuando reparamos en la presencia en internet de los diferentes territorios del globo. El mundo que alimenta a los algoritmos es un mundo deforme, que inflama a los territorios donde el desarrollo económico y tecnológico ha sido mayor.

Por tanto, a día de hoy, los datos que nos representan en el océano de datos masivos, distan mucho de ofrecer una imagen fidedigna de todos los humanos. Sin embargo, con ellos vamos construyendo perfiles de personas con la información predictiva que generan los sistemas de IA, y con ellos construimos una supuesta realidad social. Pero recordemos, los perfiles intentan representar, con muchas limitaciones, a la persona perfilada, no son la persona.

A pesar de las estrategias y marcos regulatorios puestos en marcha para evitar estos sesgos, estas herramientas pueden “blanquear”, cuando no amplificar, sesgos sociales, sobre todo si no se tienen en cuenta en su diseño todos los determinantes sociales de la salud.

Para introducir esta perspectiva es precisa la colaboración interdisciplinar en investigación biomédica, con la participación de las ciencias sociales y de las humanidades (antropología, sociología, ética, etc.). Dicha colaboración debería extender a entidades independientes capaces de revisar estos sistemas desde múltiples perspectivas (de Manuel et al., 2023).

---

**6/3**

**Grados de riesgo en el uso de IA en salud.**

Conscientes de los riesgos a nivel global de la implantación y desarrollo acelerado de la IA, la Unión Europea considera, en general, los sistemas de IA dedicados a tender a la salud humana como de alto riesgo, distinguiendo incluso diferentes escalones, partiendo de aquellas aplicaciones que sólo informan de forma más o menos organizada del manejo clínico habitual, pasando por aquellas que asesoran el manejo clínico, hasta llegar a aquellas que directamente tratan, cuidan o diagnostican.

Pero las principales incógnitas en relación al desarrollo de sistemas de IA con grados crecientes de autonomía son: ¿Podrán llegar a ser autoconscientes o tener voluntad propia? ¿Podrán arrepentirse o cambiar de opinión? ¿Podrán engañar intencionadamente y burlar sistemas de seguridad? Un artículo en Science de este mismo año no descarta en el futuro algunas de estas inquietudes y reclama una mayor inversión en seguridad de los sistemas de IA y una mayor concienciación y concreción en las medidas de gobernanza imprescindibles

(Bengio et al., 2024).

y muy fragmentadas. Además, trabajan sin disponer de datos suficientes en tiempo real para llevar a cabo decisiones de elevada trascendencia. Por ello, la mayoría de las decisiones clínicas acaban siendo extremadamente subjetivas y contingentes, poco contrastadas con datos y pruebas científicas. No hacemos una medicina “profunda” (Topol, 2019).

La IA, bien entendida como instrumento, podría ayudarnos a mejorar esto, porque nos permite analizar rápidamente muchas variables al mismo tiempo. Ha demostrado ya sus prestaciones para monitorizar de forma continua resultados o alarmas, para identificar objetivos, para organizar y poner a disposición el conocimiento disponible, para comparar intervenciones, para identificar áreas grises en aquello asumido como establecido, etc. Todas ellas son prestaciones extremadamente útiles para mejorar y hacer más sostenible una medicina cálida y de calidad.

Sin embargo, no nos puede ofrecer por el momento evaluaciones cualitativas y ponderadas, que consideren el contexto y los valores. Esto justifica la imprescindible hibridación de decisiones automatizadas y humanas en la atención a la salud. En el fondo, sería mejor hablar, por el momento, de decisiones humanas asistidas por IA.

---

**7/**

**¿Nos puede ayudar la IA a mejorar y humanizar la medicina actual?**

Debemos asumir que, en muchos casos, la medicina actual, y especialmente en los llamados países “desarrollados” es superficial, cara e ineficiente, a cargo de profesionales muy preparados pero que trabajan en organizaciones rígidas

---

**8/**

**Retos éticos.**

---

**8/1**

**Privacidad, responsabilidad, transparencia.**

Las herramientas de IA para la toma de decisiones sanitarias no dejan de ser o formar parte de dispositivos sanitarios. Conscientes de los

riesgos de intervenir sobre las vidas de las personas, tanto los fármacos como los dispositivos sanitarios están sometidos ya a una regulación estricta para garantizar que su introducción en la práctica aporta valor de forma razonablemente segura.

Por supuesto, dada la elevadísima necesidad de datos de salud generada por estos sistemas, proteger datos personales se convierte en una prioridad. Además, los datos de salud recogidos pueden tener otros fines, no médicos, para las compañías propietarias de la tecnología. Un estudio aparentemente inocente o incluso altruista en salud podría tener, en el fondo, otros fines alejados del sector salud. El estudio sería solo, en cierto modo, minería de datos.

Se genera además una responsabilidad sobre la calidad, veracidad y representatividad de los datos introducidos en las dinámicas algorítmicas. De otro modo, haremos predicciones o recomendaciones sesgadas, peligrosas y/o injustas.

En este marco, la transparencia y la rendición de cuentas son aún más clave, si cabe. Se deben establecer mecanismos transparentes de control y auditoría tanto sobre su funcionamiento como sobre sus resultados.

Hablamos pues de fiabilidad, trazabilidad y explicabilidad. Éste último se ha convertido en uno de los requisitos éticos clave en el desarrollo de la IA, aunque debemos tener en cuenta que cuanto más explicable es el sistema, menos predictivo resulta, y que la explicabilidad es progresivamente más difícil cuando el programa “**aprende**”, siendo éstos precisamente los sistemas que percibimos como un avance evidente en predicción. Además, los requisitos de explicabilidad pueden chocar también con el secreto empresarial y la libre competencia.

En cualquier caso, las personas que puedan recibir potencialmente los efectos, positivos o negativos, de la IA deberían tener acceso a una alfabetización digital suficiente, para así poder reducir su vulnerabilidad ante la IA.

No se trata de que el ciudadano medio pueda conocer su mecanismo íntimo informático, estadístico, lógico o algebraico. No se trata tanto de dominar el “¿cómo?” sino de poder valorar bien el “¿por qué?” y el “¿para qué?” de cada sistema, conociendo también sus limitaciones y posibles desviaciones. Por ejemplo, resulta esencial que los ciudadanos sepan que, hoy por hoy, los sistemas de IA sólo detectan correlaciones, sin poder definir relación causa-efecto entre fenómenos y datos. Además, deben conocer que sus predicciones y recomendaciones no son fijas, sino dinámicas, y en continua evolución en función de la entrada de datos y observaciones en el sistema.

---

## 8/2

---

### Pronósticos y predicciones ajustadas a valores.

Decíamos antes que las soluciones de IA aplicadas al cuidado de la salud deberían ser consideradas, cuando menos, como dispositivos sanitarios. Pero seguramente la IA en salud es más que un dispositivo sanitario. No es un aparato que se diseña y se construye, para llegar a un “**diseño definitivo**”. Muchas soluciones algorítmicas se transforman con el contexto, a medida que le van llegando datos y acumula “**experiencia**”. Son soluciones dinámicas. Esto es importante, porque nuestra medicina clásica está acostumbrada a horquillas de normalidad, relativamente rígidas. Sin embargo, una solución de IA puede experimentar grandes cambios en sus recomendaciones con pequeños cambios en los nuevos datos que recibe. Los dinteles de seguridad o de predicción, por tanto, son líquidos. Esto debe tenerse muy en cuenta cuando dichos resultados predictivos implican decisiones trascendentes. Conceptos de la estadística como valor predictivo, tanto positivo como negativo, cobran renovada trascendencia en el paradigma de la Medicina de las “**P**”. Cuando una herramienta tiene gran capacidad para detectar patrones y correlaciones, es necesario diferenciar el ruido de las correlaciones útiles en un océano de “**alarmas**” posibles.

Además, las correlaciones útiles en Medicina para humanos deberían estar trufadas de valores. Hasta la fecha, quizá encontremos aquí una de las limitaciones fundamentales para hacer una buena “**medicina algorítmica**”. Los dinteles que marcan actuar o no actuar, seguir o no seguir, reanimar o no reanimar, etc. generan una responsabilidad ética elevada. Y estos dinteles deben estar basado en valores compartidos. Por tanto, parece difícil que puedan establecerse desde una visión miope monodisciplinar. Volvemos a la gran pregunta: ¿Qué queremos de la IA en salud? Y más difícil aún ¿Qué resulta valioso para cada persona de todo lo que ofrece la IA en salud?

El paradigma “**promete**” una monitorización más o menos continuada del estado de salud para detectar precozmente desviaciones de “**la norma**” que permitan diagnósticos más precoces, Apoyada en su sensibilidad en la detección de patrones, parece claro que la IA hará más diagnósticos y más tempranos. Pero por el momento no sabe hacer diagnósticos oportunos (aquéllos realizados en el momento justo en el que aportarán valor a la persona). También tiene, por el momento, limitaciones precisando la necesidad de intervención. No pueden, por tanto, decidir la oportunidad de las decisiones en el marco de los valores y el contexto situacional de las personas atendidas.

Debemos recordar siempre que, por el momento, la IA propone diagnósticos, cuidados o tratamientos. Propone también factores de riesgo de enfermar, factores a menudo no advertibles por la mente humana. Pero aún no diagnostica. ¿Y si nos situamos en la hipótesis de que consiga diagnosticar en el futuro, lo hará de manera personalizada, respetuosa y justa?

También cabe cuestionarse, de entrada, si la monitorización continua de la salud es salud. La misión de la Medicina es que suframos los menos posible, si así lo queremos. ¿Un sistema de alarma 7/24 en personas sanas elimina sufrimiento, o lo genera? Asumir los beneficios implicaría una revolución cultural, y casi antro-

pológica, del ser humano. ¿Sería una vida llena de falsas alarmas, o de alarmas inconcretas, una buena vida? Predecir implica quizá diagnóstico, pero, sobre todo, pronóstico, y oportunidades en función de valores. La accesibilidad y oportunidad de la información que nos facilite la IA representan los dos principales retos éticos para su introducción en la asistencia, entendida ésta, recordemos, como relación de ayuda interpersonal, y responsable con la biosfera.

---

## 8/3

---

### Una ética de la responsabilidad en IA.

Todo avance tecnológico debería estar al servicio de las personas y de la biosfera. El servicio, para las personas, sería hacerles accesible e intuitiva la matemática compleja, la estadística analítica y el análisis de sistemas complejos.

¿Debe de ser siempre supervisada por humanos la IA? ¿Es viable siempre dicha supervisión, incluso en sistemas que usen aprendizaje profundo y produzcan resultados en tiempo real? Seguramente la supervisión humana más relevante se referiría al diseño, a los objetivos, al entrenamiento y a las preguntas a hacerle al sistema.

Y está también el problema de la responsabilidad ante los hipotéticos errores, o los efectos adversos de la aplicación de sistemas de IA cada vez más autónomos. ¿Quién se hará responsable de los daños? ¿Quién será responsable de respetar los derechos de los usuarios, incluida su privacidad? ¿Cómo evitaremos que las correlaciones detectadas deriven en discriminación y/o estigma?

La inteligencia artificial ha demostrado ser superior a los humanos en tareas de análisis masivo de datos y detección de patrones, pero carece de cualidades esenciales como la empatía y la comprensión del contexto

## 9/

### Conclusiones.

Deberíamos evitar discursos del miedo, pero sin dejar de evaluar la IA como cualquier otra tecnología innovadora de riesgo, aunque llena de oportunidades que también debemos distinguir con transparencia y accesibilidad.

Debemos distinguir aquellas tareas en las que la IA nos supera y permitir que colabore con nosotros en ellas. Seguramente el concepto de decisiones híbridas humano-IA será cada vez más habitual (decisiones humanas asistidas por IA).

Por supuesto, en el desarrollo de soluciones de IA, existe el riesgo de la progresiva concentración de empresas tecnológicas que podría llevar a una especie de monopolio del poder, del conocimiento, incluso de la moral. Preservar la independencia y la equidad en la evaluación de todo proyecto de IA en salud será una garantía a blindar democráticamente.

Otro reto para la biosfera es que, a pesar de su apariencia de sistemas “limpios” sin emisiones, su huella ecológica, relacionada con su elevadísimo consumo de energía y con los materiales empleados, es muy intensa, y mayor a medida que se desarrollan más y más potentes sistemas.

Seguramente la IA, en vez de substituir directamente profesionales sanitarios, debería convertirse en una ayuda para recuperar la humanización perdida. Por supuesto, para conseguir esto debemos aplicar los principios de prudencia y proporcionalidad. Como decíamos la pre-evaluación será clave (justificación, objetivos, características de los datos para nutrir los algoritmos, preguntas que les haremos, etc.). No hace muchos años sentíamos también inquietud y preocupación por la introducción de los ordenadores en las consultas.

Las nuevas herramientas de gestión de datos y de procesamiento del lenguaje natural pueden generar tiempo para la interacción personal, para la comunicación y para la empatía. Por supuesto, todo ello ocurrirá si el tiempo que se genere no sea devorado por la burocracia o el mero economicismo. La salud de los humanos merece beneficiarse de la sostenibilidad y de la mejora de calidad que pueden aportar las soluciones digitales, pero ello exige introducir la reflexión ética y la participación pública en el diseño, uso e interpretación de estas nuevas herramientas, que deben ser seguras y libres de intereses espurios. Deberían estar “vacunadas” contra el totalitarismo, la exploración y la manipulación. Algunas propuestas pragmáticas serían: **(Owens, 2023):**

- Mantenimiento y reparación continuada, con auditorías ad-hoc en tiempo real.
- Detección y mitigación de sesgos.
- Considerar las limitaciones de la IA a la hora de interpretar resultados:
  - Mínimas desviaciones de nuevos datos entrantes pueden alterar severamente las recomendaciones salientes.
  - La evolución tecnológica, social, demográfica, pueden deteriorar el valor añadido que aportan estos modelos.
  - Por el momento no se debería inferir relación causal de las correlaciones que detectan.
- Priorizar criterios de justicia, equidad y no discriminación en su desarrollo.
- Garantizar la accesibilidad a soluciones de IA que aporten valor a los humanos.
- Cuidar que tanto los profesionales, como las aplicaciones de IA, como las personas atendidas gocen de la mayor independencia posible. Debemos evitar convertir la IA en un bio-poder o en una estafa moral.

Por tanto, la IA puede ayudar mucho en el cuidado de la salud de las personas, pero solo será así si está bien pensada, diseñada, alimentada de datos, gestionada, revisada y mantenida. La transparencia y la rendición de cuentas serán los principios éticos clave en este nuevo escenario. Es todo un nuevo reto para la gobernanza democrática.

Sin caer en alarmismos, debemos partir de la base que la IA hace actualmente muchas cosas mejor que los humanos, pero para ser útil debe estar diseñada acorde a valores compartidos, bien nutrida de datos representativos, inclusivos y libres de prejuicios. Por último y, sobre todo, debe ser accesible y transparente para no ser injusta.

Aunque muchas actividades médicas se puedan desarrollar sin rostro, debemos tener muy presente que las cosas mejoran con rostro. Como no pocos expertos proponen ya, no debemos conformarnos con aquello que llamamos medicina personalizada (más bien estratificada en muchos casos, por el momento.) Necesitamos que la medicina sea personalizada, pero sobre todo personal (Picard, 2017; Esquerda & Pifarré, 2024), preservando el rostro y la mirada en la relación clínica.



## Bibliografía

Bengio Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., Harari, YN., Zhang, YQ., Xue, L., Shalev-Shwartz, S., Hadfield, G., Clune, J., Maharaj, T., Hutter, F., Baydin, AG., McIlraith, S., Gao, Q., Acharya, A., Krueger, D., Dragan, A., Torr, P., Russell, S., Kahneman, D., Brauner, J. y Mindermann, S. (2024) *Managing extreme AI risks amid rapid progress*. Science, 384(6698), 842-845. doi: 10.1126/science.adn0117. Epub 2024 May 20. PMID: 38768279.

Berg, HT., van Bakel, B., van de Wouw L., Jie, KE., Schipper, A., Jansen, H., O'Connor, RD., van Ginneken, B. y Kurstjens, S. (2023). *ChatGPT and Generating a Differential Diagnosis Early in an Emergency Department Presentation*. Ann Emerg Med. 2024 Jan; 83(1): 83-86. doi: 10.1016/j.annemergmed.2023.08.003. Epub 2023 Sep 9. PMID: 37690022.

de Manuel, A., Delgado, J., Parra Jounou, I., Ausín, T., Casacuberta, D., Cruz, M., Guersenzvaig, A., Moyano, C., Rodríguez-Arias, D., Rueda, J., y Puyol, A. (2023) *Ethical assessments and mitigation strategies for biases in AI-systems used during the COVID-19 pandemic*. Big Data & Society, 10(1). <https://doi.org/10.1177/20539517231179199>.

Esquerda, M. y Pifarré-Esquerda, F. (2024). *Inteligencia artificial en medicina: aspectos éticos, deontológicos y el impacto en la relación médico-paciente*. Med Clin (Barc). May 7:S0025-7753(24)00214-8. English, Spanish. doi: 10.1016/j.medcli.2024.03.003. Epub ahead of print. PMID: 38719685.

Owens, K. (2023) *For Ethical Use of AI in Medicine, Don't Overlook Maintenance and Repair*. Posted in Artificial Intelligence, Hastings Bioethics Forum. <https://www.thehastings-center.org/for-ethical-use-of-ai-in-medicine-dont-overlook-maintenance-and-repair/>

Picard, A. (2017), *Matters of life and death: Public health issues in Canada*. Vancouver: Douglas & McIntyre.

Pujol, R. y Guanyabens, J. (2023). *La medicina en tiempos de inteligencia artificial*. Med Clin (Barc.), 161(12), 530-32. <https://doi.org/10.1016/j.medcli.2023.09.007>

Rizzolatti L. y Cattaneo L. (2009). *The Mirror Neuron System*. Arch Neurol, 66 (5), 557-560.

Slack WV. (1997). *Cybermedicine: How Computing Empowers Doctors and Patients for Better Care*. San Francisco: Jossey-Bass Edition.

Topol, EJ. (2019) *Deep medicine: how artificial intelligence can make healthcare human again*. First edition. - New York: Basic Books.