

>> Haciendo lazos



Consideraciones éticas respecto de la Inteligencia Artificial

Sofía P. Salas

Introducción

Hace algo más de 100 años, el escritor checo Karel Capek, en su obra *Rossum's Universal Robots*, narra la historia de Rossum, un científico que crea máquinas semejantes al hombre las que, a pesar de que pueden pensar por sí mismas, son explotadas en las industrias para que trabajen como obreros. El origen del término *Robot* viene de la palabra eslava *robot* o *roboti* (en plural), que se refiere precisamente al trabajo forzado (1). Tiempo después, fue el escritor de ciencia ficción Isaac Asimov quien se preocupó de manera especial por las leyes que debían gobernar el mundo de los robots. Tal como lo señala en su obra *Círculo Vicioso* (1942), estas tres leyes establecieron lo siguiente:

“Primera Ley: Un robot no hará daño a un ser humano ni, por inacción, permitirá que un ser humano sufra daño.

Segunda Ley: Un robot debe cumplir las órdenes dadas por los seres humanos, a excepción de aquellas que entren en conflicto con la primera ley.

Tercera Ley: Un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la primera o con la segunda ley.”(2)

Estas leyes rigen los códigos del programa que regula el cumplimiento de los programas que se guardan en la memoria principal de los robots y no pueden eludirse, puesto que son una característica de seguridad propia de la interacción de los robots con los seres humanos. Posteriormente, cuando en sus libros los robots habían adquirido otras responsabilidades a nivel planetario, Asimov agregó una ley cero, que antecede a las tres anteriores: *“Un robot no puede dañar a la humanidad o, por inacción, permitir que la humanidad sufra daños”*. En la mirada de este autor, estas leyes eran necesarias para proteger a los humanos de estas criaturas que,

eventualmente, podían rebelarse ante el control humano. Al estar estas leyes inscritas “en su cerebro”, los robots no podían desobedecerlas (aunque en alguna de sus ficciones no siempre esto quedó tan claro).

Todo esto que nos parece más bien lejano, propio de los cuentos de ciencia ficción que tal vez leímos cuando niños, tiene importancia actual al hablar de los aspectos éticos de la inteligencia artificial (IA). En el presente artículo, revisaremos brevemente el concepto de IA, luego veremos la experiencia adquirida al programar los autos autónomos o inteligentes, para finalizar con su aplicación en medicina y la discusión sobre las consideraciones éticas que debieran regular su uso.

El nacimiento de la Inteligencia Artificial (IA)

Para muchos, John McCarthy es considerado como el padre de la IA, puesto que en una conferencia realizada en el Dartmouth Colleague en 1956, llamó IA a las nuevas técnicas empleadas para poder dotar a las máquinas de inteligencia (3). McCarthy desarrolló programas matemáticos que permitieron que algunas computadoras jugaran ajedrez y, eventualmente, le ganaran a jugadores humanos. Pero antes que se acuñara el término de IA, fue Alan Turing quien ideó un examen, llamado posteriormente prueba o test de Turing, para determinar la capacidad de una máquina de tener un “*comportamiento inteligente similar al de un ser humano*”. Propuso Turing que se evaluara una conversación entre una persona y una máquina diseñada para producir respuestas similares a las de un humano; un observador independiente debía juzgar quién era quién en base a las respuestas de texto. Inicialmente había planteado la pregunta de si las máquinas “*podían pensar*”; pero ante la dificultad de definir lo que es “pensar”, decidió formular otra pregunta: “*¿Existirán computadoras digitales imaginables que tengan un buen desempeño en el juego de imitación?*”. En la base de su pensamiento está la creencia de que las máquinas pueden demostrar un comportamiento inteligente similar al de los seres humanos. Desde entonces se han perfeccionado las máquinas y la IA está inserta en nuestras actividades cotidianas de un modo que muchas veces no estamos conscientes (3). A modo de ejemplo, usamos asistentes personales como Siri o Alexa, cuando buscamos la mejor manera de llegar

de un lugar a otro (un ejemplo es Waze), en video juegos, en la navegación aeroespacial y también en la medicina. Incluso la programación de los así llamados “vehículos autónomos” para tomar decisiones complejas requiere de IA.

Vehículos autónomos e IA

Actualmente existe la tecnología para dotar a los vehículos de un sistema de conducción que ya no requiere de un ser humano al volante. Si bien su sistema computacional estará cargado con indicaciones precisas respecto de qué hacer frente a una luz roja o en un cruce de calles, donde sabrá reconocer con exactitud qué vehículo tiene la preferencia, es más difícil que pueda ser equipado con un sistema de decisiones morales que le permita decidir respecto de qué vida preservar -la de los ocupantes del vehículo o los peatones que tiene enfrente- ante un choque que parece ser inevitable. Este verdadero dilema moral obligará a los programadores de estos vehículos a dotarlos de un sistema de decisión para que puedan evaluar cómo dividir el riesgo del daño entre los distintos actores que se encuentran en una carretera. En este tipo de escenario, es evidente que las leyes de Asimov son insuficientes para resolver estas situaciones (4). En el año 2018, un grupo de investigadores publicaron los resultados de un experimento realizado en línea, que llamaron la “Máquina moral”, donde pudieron recoger cerca de 40 millones de decisiones, generadas en 233 países o territorios, respecto de si un auto debía seguir su curso o cambiar de dirección, según qué obstáculo tenía en frente y, de manera importante, cuál era el perfil de personajes que podía colisionar (humanos o no humanos, jóvenes o viejos, hombres o mujeres, criminales o doctores, con conducta cuidadosa o imprudente al cruzar la calle, entre otras variables). Lo interesante de este experimento es que las decisiones morales demostraron tener que ver con la cultura de quienes respondían, lo que implica que es difícil que se pueda lograr un consenso respecto de cómo se pueden programar estos vehículos (4, 5). En condiciones reales, es el conductor el que en milésima de segundo toma la decisión, probablemente de manera instintiva; en cambio, los vehículos autónomos deben ser programados de antemano para seguir en la ruta trazada inicialmente o hacer un giro e impactar a otro vehículo, a seres vivientes, a contra una muralla. Dado que no es fácil lograr una mirada consensuada

respecto de cómo resolver estos escenarios, algunos autores consideran que un estándar mínimo es que los eventuales compradores sean informados respecto de cómo ha sido programado el vehículo, para así poder elegir entre las distintas opciones, aquella que está más de acuerdo con su propio sistema de valores.

IA en medicina

Así como se ha desarrollado IA para el uso de diversas tecnologías, también ha crecido su uso en medicina, como una herramienta más puesta al servicio de la atención clínica. Ya existe suficiente evidencia que muestra que sistemas de IA son más eficientes en realizar diversos diagnósticos médicos que los mejores expertos en las distintas áreas clínicas, en la medida que sean programados de manera adecuada. Ejemplo de esto son sus usos en áreas como el diagnóstico de imágenes radiológicas, dermatológicas o en el examen de fondo de ojo para pesquisar retinopatía diabética, y también en el análisis de electrocardiogramas para diagnosticar ciertas condiciones cardíacas, entre otras muchas aplicaciones médicas (6). Por lo tanto, se espera que el uso de IA impacte en la práctica médica y en la entrega de servicios de salud en un futuro próximo (7). Lo verdaderamente interesante del uso de IA en el campo de la medicina, es que el aprendizaje de las máquinas (*machine learning*) puede entregarle a los computadores la capacidad de aprender aun cuando no hayan sido explícitamente programados para hacerlo, pudiendo crear estrategias que les permiten ser sobresalientes -más rápidos y con mayor certeza que los propios profesionales médicos- en una determinada tarea. Pero el uso clínico de la IA, al menos por ahora, no es ni será perfecto, por lo que se requiere de un alto nivel de supervisión y también de regulación (8). Además, es importante señalar que para que su uso clínico se expanda, se requiere de profesionales sanitarios que interpreten e integren de manera adecuada la IA a su práctica, puesto que es inevitable que exista una especie de “caja negra”, que pertenece oculta al clínico, relacionada con la forma en que las máquinas aprenden, más allá de su programación inicial.

Aspectos éticos del uso de IA en medicina

El 2021, los miembros de la UNESCO adoptaron un histórico acuerdo que define los valores comunes y los principios que se requieren para desarrollar el uso de IA en diversas áreas relevantes, como son la educación, la ciencia, la cultura y la comunicación y la información(9). En la Tabla 1 se muestran los principales valores y principios identificados en dicho documento. En lo que respecta al ámbito de la salud y bienestar social, la UNESCO señala que los Estados Miembros deberían esforzarse porque el uso de la IA mejore la salud humana y proteja el derecho a la vida, junto con preocuparse porque los distintos actores involucrados en los sistemas de IA relacionados con la atención en salud “tengan en cuenta la importancia de las relaciones del paciente con su familia y con el personal sanitario”. Asimismo, insta a que se haga “participar activamente a los pacientes y sus representantes en todas las etapas pertinentes del desarrollo del sistema”, aspecto que a nuestro juicio aún es incipiente.

Otras preocupaciones se refieren a la forma, aún insospechada, mediante la cual la IA impactará la relación clínica, especialmente en lo que relacionado con el proceso de consentimiento informado(3). Una duda aún no resuelta es si el profesional médico debe o no revelar a su paciente que se está usando IA y en qué ámbitos del quehacer clínico. Otro aspecto relevante tiene que ver con temas de seguridad y transparencia, puesto que es fundamental que la forma en que se entrena a la máquina (cómo “se la alimenta” para que diseñe los algoritmos necesarios), sea lo más real posible y no en base a casos hipotéticos. Para lograr esto, es imprescindible acceder a base de datos de pacientes reales, lo que puede poner en juego la debida privacidad y confidencialidad de la información clínica.

Adicionalmente, se requiere que las máquinas aprendan de una manera no sesgada. A modo de ejemplo, puede haber diferencias raciales en la forma de presentación de ciertos cánceres cutáneos, que podrían pasar inadvertidas en un sistema de diagnóstico mediante IA si los datos con que son alimentados las máquinas no contienen suficiente diversidad étnica o racial para dar cuenta de estas diferencias.

Conclusión

El uso de IA en el ambiente clínico puede permitir sistemas de diagnósticos y terapéuticos más confiables, con menor uso de recursos humanos y en un menor tiempo. Pero como ha sucedido otras veces, los avances tecnológicos plantean nuevos desafíos éticos para los cuales no siempre estamos preparados. En este sentido, se ha sugerido que se puedan crear comités de ética específicos para evaluar el impacto de estas nuevas tecnologías en el quehacer clínico (10). Entre las preguntas que pueden intentar resolver, es cuál es la real utilidad de esta nueva herramienta tecnológica, que justifique su uso en la clínica. Pero también deberá examinarse si los algoritmos entrenados para evaluar la salud del paciente están o no sesgados; si el proceso de consentimiento informado se llevó a cabo de manera adecuada y, yendo más lejos, cuánto debiese saber el paciente del uso de IA para resolver su situación de salud. Asimismo, antes que el uso de IA se expanda, es necesario proveer de adecuados mecanismos de supervisión ética y de aseguramiento de la privacidad de los datos. En consecuencia, el uso de IA en salud sólo será segura y confiable si se incorpora la mirada ética en las distintas etapas del proceso.

Sofía Salas es médico cirujano de la Pontificia Universidad Católica de Chile. Actualmente se desempeña como Profesor Titular del Centro de Bioética de la Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, donde realiza docencia de pre y posgrado. Integra además el Departamento de Ética del Colegio Médico de Chile. sofiasalas@udd.cl

Las opiniones vertidas en esta sección son de exclusiva responsabilidad de quienes las emiten y no representan, necesariamente, el pensamiento de Bioeticar Asociación Civil.

Mayo 2022

Tabla 1. Valores y principios sobre la ética de la inteligencia artificial

Valores	Principios
Respeto, protección y promoción de los derechos humanos, las libertades fundamentales y la dignidad humana.	Proporcionalidad e inocuidad
Prosperidad del medio ambiente y los ecosistemas.	Seguridad y protección
Garantizar la diversidad y la inclusión.	Equidad y no discriminación
Vivir en sociedades pacíficas, justas e interconectadas	Sostenibilidad
	Derecho a la intimidad y protección de datos.
	Supervisión y decisión humanas.
	Transparencia y explicabilidad.
	Responsabilidad y rendición de cuentas.
	Sensibilización y educación.
	Gobernanza y colaboración adaptativas y de múltiples partes interesadas.

Adaptado de referencia (9)

Referencias

1. Viso E (2021). El origen de la palabra robot. Xatakaciencia.com. Recuperado 4 abril 2021, de <https://www.xatakaciencia.com/robotica/el-origen-de-la-palabra-robot>.
2. Tres leyes de la robótica. https://es.wikipedia.org/wiki/Tres_leyes_de_la_rob%C3%B3tica. Recuperado el 4 abril 2022.
3. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. *Artificial Intelligence in Healthcare*. 2020:295-336.
4. Awad E, Dsouza S, Kim R, Schulz J, Henrich J, Shariff A, et al. The Moral Machine experiment. *Nature*. 2018;563(7729):59-64.
5. Maxmen A. Self-driving car dilemmas reveal that moral choices are not universal. *Nature*. 2018;562(7728):469-70.
6. Larentzakis A, Lygeros N. Artificial intelligence (AI) in medicine as a strategic valuable tool. *Pan Afr Med J*. 2021;38:184.
7. Mesko B, Gorog M. A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit Med*. 2020;3:126.
8. Maddox TM, Rumsfeld JS, Payne PRO. Questions for Artificial Intelligence in Health Care. *JAMA*. 2019;321(1):31-2.
9. UNESCO, 2021. Recomendación sobre la ética de la inteligencia artificial. Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Promulgada en París, Francia, noviembre de 2021.
10. Cossitt A. Why Health Care Organizations Need Technology Ethics Committees. *Bioethics Forum Essay* [Internet]. 2020 April 10, 2022. Disponible en: <https://www.thehastingscenter.org/why-health-care-organizations-need-technology-ethics-committees/>.