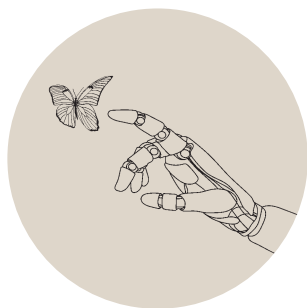




EL VERANO DE LA INTELIGENCIA ARTIFICIAL

Carol Hernández

*En ninguna otra especialidad necesitamos más
una brújula ética que en la inteligencia artificial*

Gabriela Ramos,
subdirectora general de Ciencias Sociales y Humanas, Unesco

La bioética, desde su origen en la segunda mitad del siglo XX, ha asumido la tarea de restablecer el diálogo entre las ciencias y las humanidades, para reflexionar sobre cómo interactuamos con todas las manifestaciones de la vida y crear relaciones más justas, incluyentes y sustentables. Entre otras preguntas centrales, esta disciplina se plantea qué tipo de relaciones establecemos con la ciencia y la tecnología, cuáles son los impactos de ambas en las sociedades y en los ecosistemas, y cómo estos avances transforman nuestros modos de comprender y valorar la vida.

Según la definición de la Unesco, las inteligencias artificiales son aquellos sistemas capaces de procesar datos como lo haría un ser de comportamiento inteligente. Dicha tecnología está revolucionando el mundo y nuestra comprensión sobre las capacidades y los límites de la mente humana. La IA y sus derivaciones ya se han vuelto cotidianas e incluso indispensables en diversos ámbitos, desde las redes sociales y los diagnósticos médicos hasta los *softwares* de vigilancia masiva y el armamento especializado.

Es paradójico que, pese a su omnipresencia, el sector de la IA se caracterice por su abstracción y opacidad. Salvo por las personas que diseñan

e implementan estos sistemas, los ciudadanos comunes no entendemos a profundidad qué es la IA, cuáles son sus aplicaciones y cómo interfieren en asuntos prácticos, emocionales y hasta ontológicos. La iniciativa privada de un puñado de países concentra la inversión y el desarrollo de esta tecnología; los polos más dinámicos son Estados Unidos, China y la Unión

entiendan las necesidades de otras entidades inteligentes, incluido el discernimiento de sus creencias, emociones y procesos de pensamiento (en la actualidad, esa tecnología ya está en desarrollo). Esto requeriría que las máquinas nos comprendieran y funcionaran como nosotros, para replicar nuestra multifuncionalidad. A esto se le denomina IA general.

Se ha demostrado que el uso indiscriminado de algoritmos en la justicia penal reproduce nuestros sesgos raciales y de clase.

Europea, en ese orden. Todo esto ha propiciado que prácticamente no exista regulación y que los debates entre diversos grupos sociales sobre las consecuencias sean escasos.

¿DEBEMOS CREAR TODO LO QUE PODEMOS CREAR?

El objetivo de la investigación en IA es crear máquinas que emulen la versatilidad de la inteligencia y el comportamiento de nuestra especie. El grado en el que una máquina puede replicar estas capacidades, incluida la habilidad de “pensar y sentir como los humanos”, es uno de los criterios para clasificar los tipos de IA.

Las IA reactivas son máquinas sin memoria que solo responden a estímulos y no pueden aprender. Se trata de las formas más antiguas de IA y su capacidad es muy reducida. En cambio, las IA de memoria limitada usan un sistema de memoria para aprender datos a lo largo del tiempo y mejorar sus respuestas. En la actualidad, la mayoría de las aplicaciones de IA se encuentran en esta categoría, pero siguen siendo limitadas porque solo pueden hacer ciertas tareas específicas de manera autónoma.

Según la inteligencia emocional artificial, en el futuro será posible diseñar máquinas que

La IA autoconsciente son máquinas hipotéticas conscientes de sí mismas y poseedoras de una inteligencia que emularía a la humana; por lo tanto, experimentarían emociones, necesidades, creencias y deseos propios. A esto se le conoce como superinteligencia artificial. Su desarrollo ha sido el objetivo de la programación en esta área desde sus orígenes, en la década de 1950.

Abundan las obras literarias y cinematográficas que problematizan o anticipan los riesgos de la IA. Se han escrito numerosos relatos sobre máquinas que desplazan a los trabajadores y transforman comunidades enteras. En varias películas y novelas encontramos robots que son amantes, niñeras, profesores y compañeros de trabajo, y los personajes humanos se relacionan con ellos de maneras más satisfactorias y profundas. Incluso hemos imaginado la *singularidad*, es decir, futuros distópicos en los que las máquinas amenazan, dañan y exterminan a nuestra especie.

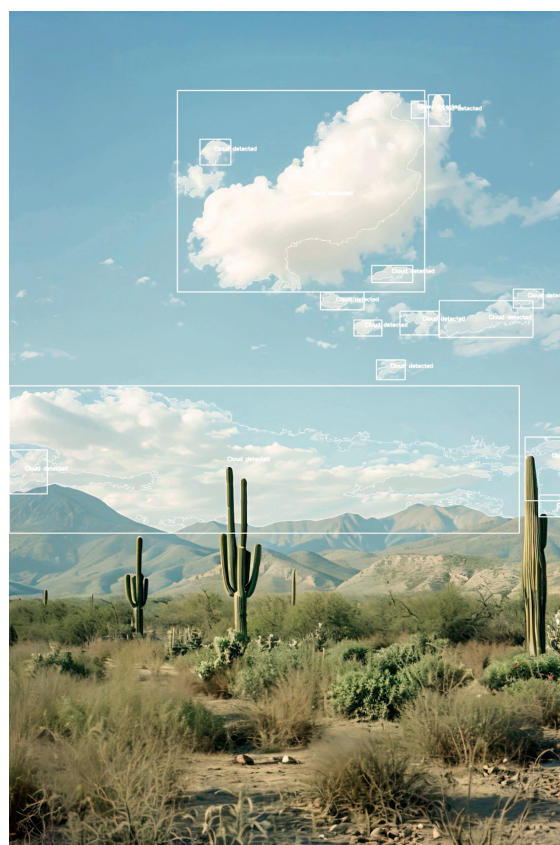
Si bien la mayoría de estos escenarios se inscriben en la ciencia ficción, o bien, son posibilidades de futuros lejanos, no debemos ignorar los llamados urgentes sobre la IA que se han acumulado desde 2023.

En marzo del año pasado, después de que la nueva versión de ChatGPT saliera al mercado, más de mil líderes e investigadores, incluido Elon Musk, firmaron una carta abierta pidiendo pausar el desarrollo de los sistemas más avanzados hasta que se introduzcan protocolos de seguridad compartidos por las entidades involucradas. Aun si no fuera posible conceder este paréntesis, agrega la carta, los gobiernos deben intervenir y establecer una moratoria.¹

Unos días más tarde, diecinueve líderes de la Asociación para el Avance de la Inteligencia Artificial² compartieron la misma preocupación en otra carta pública. A finales de abril, Geoffrey Hinton, “padrino de la IA”, tras renunciar a Google después de una década de trabajar en esa compañía, habló sobre la necesidad de regular el desarrollo de estas tecnologías y de mantener el control humano y democrático sobre máquinas que podrían ser usadas de forma equivocada o con fines dañinos.³ Hinton ha creado algunas de las herramientas más poderosas de los últimos cincuenta años, como las redes neuronales que son la base de ChatGPT, Siri y Alexa.

Un caso conocido sobre el uso inadecuado de la IA son las campañas masivas de desinformación que han puesto en riesgo las elecciones en varios países; también hay campañas similares que contradicen el conocimiento científico sobre el cambio climático. En el fu-

turo podrían fabricarse tecnologías bélicas (ya existen, por ejemplo, los soldados robot). En repetidas ocasiones se ha demostrado que el uso indiscriminado de algoritmos en la justicia penal reproduce nuestros sesgos raciales y de clase, por lo que la IA podría perpetuar las desigualdades sociales. Gary Marcus, académico y empresario del sector, caracteriza así la situación: “[Las tecnologías de IA] están dando forma a nuestro mundo. [...] Estamos ante una tormenta perfecta que es resultado de la irresponsabilidad corporativa, la adopción generalizada [de estas tecnologías], la falta



Helio Santos, *Máquinas que contemplan*, 2024. Paisaje sintético del desierto de la frontera mexicana visto por un sistema de detección computarizado

¹ “Pause giant AI experiments: An open letter”, Future of Life Institute, 22 de marzo de 2023. Disponible en bit.ly/3UbwnhA.

² En inglés, Association for the Advancement of Artificial Intelligence. Se trata de una sociedad científica internacional dedicada a la comprensión de los mecanismos del pensamiento y del comportamiento inteligente, y de su reproducción en máquinas.

³ Cade Metz, “‘The godfather of A.I.’ leaves Google and warns of danger ahead”, *The New York Times*, 1 de mayo de 2023. Disponible en nyti.ms/4aRnOxT.

de regulación y una enorme cantidad de incertidumbre”.⁴

La suma de estos exhortos nos advierte que no contamos con el tiempo suficiente para discernir las implicaciones de las IA —cada vez más poderosas, autónomas y capaces de procesar cantidades inmensas de información—, ni para acordar consensos sobre si dichas tecnologías deben ser creadas y bajo qué condiciones pueden salir al mercado. Parece complicado retrasar su desarrollo. Las empresas están inmersas en un mercado muy competitivo y en una carrera permanente por crear tecnologías cada vez más sofisticadas y renta-

bles en el menor tiempo posible. Tan solo entre 2013 y 2023, la inversión global en este sector pasó de menos de 20 mil millones de dólares a más de 160 mil millones, y se espera que en 2025 este monto supere los 200 mil millones, con Estados Unidos a la vanguardia.⁵

AIRES ÉTICOS

Como ha sucedido en otros periodos de desarrollo tecnológico acelerado, hoy comienzan a surgir iniciativas para delinear una ética de la IA. Embedded Ethics, creado por la Universidad de Harvard en 2017, es un proyecto pionero dentro del sector. Su propósito es educar

⁴ C. Metz y Gregory Schmidt, “Elon Musk and Others Call for Pause on A.I., Citing ‘Profound Risks to Society’”, *The New York Times*, 29 de marzo de 2023.

⁵ “AI Investment Forecast to Approach \$200 Billion globally by 2025”, Goldman Sachs, 1 de agosto de 2023. Disponible en bit.ly/3PYUYWw.



Alonso Cedillo, *Muro de hielo*, 2024. Video generado con Gen-2 + Tortoise

a estudiantes de ingeniería y ciencias aplicadas sobre las implicaciones éticas y sociales de los sistemas inteligentes que ellos mismos diseñan. En vez de impartir cursos aislados sobre el asunto, la intención es diseñar un modelo pedagógico que tenga a la ética como eje transversal de la educación de los futuros programadores. El reto, de acuerdo con las cofundadoras del programa, Barbara Grosz y Alison Simmons, es que los alumnos aprovechen este conocimiento para dar forma a una práctica ética que apliquen al incorporarse a un mercado laboral que se caracteriza por la hipercompetitividad y la exigencia permanente de innovación. Como bien dice Grosz, los estudiantes no solo deben preguntarse si pueden construir un sistema, también es indispensable que se pregunten si deben construirlo y cómo pueden hacerlo de forma ética.⁶

Fuera de dicho sector, el logro más importante es la *Recomendación sobre la ética de la inteligencia artificial* de la Unesco, que en 2021 fue adoptada por sus 193 Estados miembros. Este marco establece cuatro valores fundamentales y diez principios para formular un enfoque ético centrado en derechos humanos que permite a los gobiernos avanzar hacia una regulación responsable. Del debate sobre este documento ha surgido una observación crucial: pese a la progresiva automatización de los algoritmos de la IA, estos siguen siendo resultado de acciones humanas (como el diseño, el financiamiento, la programación y la implementación), las cuales pueden y deben ser reguladas por la ley. En este sentido, la vicepresidenta de Estados Unidos, Kamala Harris,

apuntó que “el sector privado tiene la responsabilidad ética, moral y legal de garantizar la seguridad de sus productos y cada compañía debe cumplir con las leyes existentes”.⁷

¿QUIÉN DEBE DECIDIR NUESTRO FUTURO?

Los algoritmos de la IA no son inteligibles para la mayoría de las personas. Para revertir esta situación, se debe promover la educación de la sociedad y los gobiernos, de modo que se comprendan estas tecnologías, sus usos y sus implicaciones. Esto resultará imposible si las autoridades no obligan a las corporaciones a transparentar sus operaciones.

La IA no es neutral, de ahí que sea necesario limitar su capacidad de reproducir prejuicios que profundicen las desigualdades de clase, género, raza y etnia. Vale la pena recordar que el sector aún se encuentra dominado por hombres jóvenes, blancos, de clase media y alta que residen en los países del norte global.

Por si fuera poco, los costos ambientales de la IA son elevados pero pasan desapercibidos porque creemos erróneamente que los datos intangibles se concentran en internet o en “nubes” como iCloud que imaginamos igual de inmateriales.

La revolución tecnológica actual seguirá su curso, pero debemos discutir cómo regular su desarrollo y prevenir sus riesgos, cómo distribuir sus beneficios y costos de forma equitativa, y cómo usarlas para promover la paz y la justicia. Si la sociedad civil y los gobiernos no nos involucramos lo antes posible, dejaremos en manos de unos cuantos empresarios e ingenieros la decisión sobre el tipo de sociedades en las que viviremos. **U**

⁶ Christina Pazzanese, “Trailblazing Initiative Marries Ethics, tech”, *The Harvard Gazette*, 16 de octubre de 2020. Disponible en bit.ly/43KyCsf.

⁷ C. Metz, “The godfather of A.I.”, *op. cit.*, 1 de mayo de 2023.