

# OBSIDIANA

Ciencia y Cultura por México

pág. 20

IA: nuevo colega o  
relevo definitivo

Miguel A. Padilla Acosta

pág. 26

Programar las  
emociones binarias

Erik Cuevas y Daniel Zaldivar

Inteligencia  
artificial

AÑO 1, NÚM. 2.  
MÉXICO, MARZO DEL 2023





# Manifiesto

Así como el vidrio volcánico se cristaliza con el brusco enfriamiento de la lava tras una erupción, *Obsidiana* se concibe como el suplemento del Periódico *Reforma* que amalgama la ciencia, la cultura y la tecnología.

Suscriptores y lectores del país hallarán aquí una respuesta a su búsqueda de un modelo de innovación para México y Latinoamérica, a través de un conjunto de artículos, textos, criterios y opiniones.

Una propuesta de conocimiento para construir una visión de país a partir del descubrimiento, pasar del reto a la solución de los problemas de México.

La diversidad de temas y plumas detonarán un llamado a la acción y la crítica de nuestra identidad pública y política, a través de las artes, la cultura y la ciencia.

*Obsidiana* es contraste y es encuentro para generar y desarrollar por consenso las soluciones funcionales, modelos de negocio, propuestas legislativas y de política pública para asegurar nuestra supervivencia como humanidad.

## Consejo Editorial

Presidente  
**José Franco**

**Estrella Burgos, Lamán Carranza  
Ramírez, Luz de Teresa, Luis  
Roberto Flores Castillo, Alejandro  
Frank, Azucena Galindo,  
Cinthya García Leyva, Marcia  
Hiriart, Alonso Huerta, Antonio  
Lazcano, Omar López Cruz, María  
Nieves Noriega, Raúl Rojas,  
Pedro Salazar, José Seade,  
Marina Stavenhagen, Brenda  
Valderrama**

## Equipo Editorial

**Lamán Carranza Ramírez**  
Dirección general

**Sergio Lenoyr Lugo**  
Dirección editorial y contenidos

**Luisa Fernanda González Arribas**  
Editora en jefe

**Omar Hernández Godínez**  
Diseño e ilustración editorial

## No. 2 Inteligencia Artificial

**Raúl Rojas González**  
Editor invitado

 **BSIDIANA**  
Ciencia y Cultura por México

MARZO 2023

obsidianadigitalmx@gmail.com  
www.obsidiana-mexico.com



# CONTENIDO

## EMERGENTE

2

Evolución y entendimiento  
Lamán Carranza

## TRANSLÚCIDO

16

Diálogos de chatbot.  
ChatGPT en entrevista

## ESPEJO

4

La inteligencia artificial  
Raúl Rojas

## ESPEJO

18

Disrupción contra regulación para México y el mundo  
Pablo Pruneda Gross y Pedro Salazar Ugarte

## ESPEJO

8

Los límites del “razonamiento”  
Luis A. Pineda

## ESPEJO

22

Autómatas de servicio  
Jesús Savage

## ESPEJO

10

En el núcleo del algoritmo  
L. Enrique Sucar

## ESPEJO

24

IA Explicable: La necesidad de que las máquinas se justifiquen  
Rosae Martín

## ESPEJO

12

Ensayo y error: aprendizaje por refuerzo  
Eduardo Morales

## INTRUSIÓN

30

De algoritmos y escenarios teatrales. Entrevista a Flavio González Mello  
Luisa González



# EVOLUCIÓN Y ENTENDIMIENTO

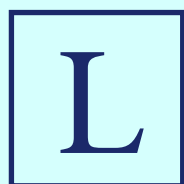
Lamán Carranza  
DIRECTOR GENERAL



@LamanCarranza



@lamancarranza



La inteligencia artificial (IA) es una de las tecnologías de mayor crecimiento e innovación en la última década. Su vertiginoso avance se debe al aprendizaje automático aunado a varios campos, entre los que están el procesamiento

de lenguaje natural, el reconocimiento de imágenes y la robótica.

Si tuviéramos que explicar a los más pequeños del hogar, diríamos que la inteligencia artificial es un campo de la tecnología que se ocupa de cómo hacer que las computadoras actúen de manera similar a los humanos.

Esto se logra mediante algoritmos computacionales que se mejoran continuamente y sus aplicaciones se expanden día a día. La mayoría de los ejemplos de inteligencia artificial sobre los que escuchamos hablar, desde computadoras que juegan ajedrez y *Go*, hasta robots y coches autónomos, recurren al aprendizaje profundo y al procesamiento del lenguaje.

Las computadoras pueden ser entrenadas con estos algoritmos para realizar tareas específicas, procesar grandes cantidades de datos y reconocer patrones en ellos, algo que ya están aprovechando a la perfección las empresas tecnológicas más grandes del mundo como Amazon, Google y Facebook. Existen aplicaciones como DALL-E y ChatGPT, con las que medio mundo se ha enganchado al teclado de la computadora. En este número de *Obsidiana* podrán apreciar algunas de sus creaciones artísticas y una entrevista muy interesante.

Hoy, gracias a la IA somos capaces de realizar tareas de forma más efectiva, rápida y automática, además de lograr descubrimientos científicos impensables hasta hace un par de años. Por ejemplo, la revista *Nature* publicó recientemente que con herramientas de IA se encontraron más de 600 millones de estructuras de proteínas de bacterias, virus y otros microorganismos.

Se estima que la IA favorecerá las ganancias de muchas industrias como las de energía, servicios, salud, minería y movilidad. De cumplirse las previsiones, China sería uno de los países más beneficiados por este

tipo de tecnología, con un impacto sobre su Producto Interno Bruto superior al 26%. Algo verdaderamente impresionante, una cantidad de dinero inimaginable.

El mercado mundial de la IA será este año de casi 200 mil millones de dólares, y se estima que en 2030 será de mil 500 miles de millones de dólares. Muy superior a los 950 mil millones de dólares de ganancias combinadas que obtuvieron las cinco mayores empresas petroleras del mundo en 2021.

La IA es ciencia informática avanzada y provoca importantes cambios en la forma en que vivimos y trabajamos, transforma nuestro entorno, hace realidad muchas cosas que hasta ahora sólo aparecían en películas y novelas de ciencia ficción. Su variedad de aplicaciones va desde la atención médica hasta procesos industriales, pasando por las redes sociales, las plataformas audiovisuales, la innovación educativa y la conducción autónoma. Y tengan por seguro que su uso se extenderá aún más en el futuro.

Me complace presentar a ustedes el segundo número de *Obsidiana: Ciencia y Cultura por México*, donde se integran una serie de artículos sobre inteligencia artificial y sus impactos en la sociedad global presente y futura. Los invito a recorrer estas reflexiones extraordinarias sobre cómo la IA se convirtió en una tecnología fundamental; Raúl Rojas, Rosae Martín, Luis A. Pineda, Enrique Sucar, Luis Roberto Flores, Eduardo Morales, Jesús Savage, Erik Cuevas, Daniel Saldivar, Pablo Pruneda, Pedro Salazar y Miguel Padilla nos describen sus profundidades, últimas tendencias y desarrollos.

Si a través de estas páginas es posible entender el momento histórico en el que nos encontramos y el lugar que queremos ocupar; entonces tomaremos las oportunidades que brinda la IA. Además, es importante darnos cuenta de que su desarrollo también necesita un equilibrio con valores éticos y morales.

El bien común necesita ampliar nuestras formas de aprender y navegar por este mundo; la inteligencia artificial requiere que nosotros también evolucionemos en el entendimiento de las reglas del juego y de los razonamientos más profundos.

## La evolución de la IA provocará importantes cambios en la forma en que vivimos y trabajamos.







[www.amc.edu.mx](http://www.amc.edu.mx)



# CONOCE A LA ACADEMIA MEXICANA DE CIENCIAS (AMC)

- ▶ Es el espacio en el que las y los **científicos mexicanos** convergen en la búsqueda de respuestas a problemas que aquejan a la sociedad.  
  
Su propósito es **impulsar la ciencia como un motor de bienestar para la sociedad**, para entender a la naturaleza y al ser humano, y así contribuir a tener un **México mejor**.
- ▶ Es un **espacio de diálogo** para todas las áreas del conocimiento, donde convergen las ciencias exactas, naturales, sociales, las humanidades y las ingenierías.



- ▶ Somos más de **3,000 miembros** en todo México, de múltiples instituciones.
- ▶ El trabajo de la Academia se enfoca en la **investigación básica y aplicada, la docencia y la divulgación** del conocimiento científico.
- ▶ Desarrollamos **programas académicos** con vocación de impacto en la sociedad.
- ▶ La AMC otorga, en conjunto o independientemente, algunos de los **premios más importantes** para la comunidad científica del país.



[www.youtube.com/acamexcie](http://www.youtube.com/acamexcie)

**SÍGUENOS** >



@amciencias





# LA INTELIGENCIA ARTIFICIAL

Raúl Rojas

UNIVERSIDAD LIBRE DE BERLÍN

# H

oy en día se habla por todos lados de la inteligencia artificial (IA), es decir, de la inteligencia de las computadoras. Es un viejo sueño de la humanidad crear artefactos capaces de imitar el comportamiento humano o, incluso, superarlo.

Al principio, pensadores y dramaturgos abordaron esta cuestión desde la perspectiva de la filosofía o de la ciencia ficción. En 1922, la obra de teatro de Karel Capek, *Los Robots Universales de Rossum*, implantó firmemente la palabra *robot* en el vocabulario universal. Al final de la puesta en escena, las máquinas terminan haciéndose cargo del planeta una vez eliminados los humanos.

La IA dejó de ser meramente un sueño (o una pesadilla) a raíz de la invención de las computadoras, en el período entre 1938 y 1945; entonces se contó por primera vez con dispositivos de cálculo miles o millones de veces más rápidos que las simples calculadoras mecánicas. No es casualidad que los pioneros de la computación se interesaran por problemas de control y decisión que requieren la evaluación de alternativas, algo reservado hasta entonces a los humanos.

Para hablar de estos temas, el matemático Norbert Wiener acuñó en esa época el vocablo *cibernética*, que fue después sustituido por *inteligencia artificial*, como decimos ahora.

Hay varios tipos de aproximaciones a la IA. A veces se habla de “IA general”, es decir de computadoras capaces de conversar con los humanos y de poder resolver creativamente cualquier problema al alcance de un adulto promedio.

Fue el brillante matemático británico Alan Turing quien propuso un experimento con el que se podría poner a prueba una IA general computarizada. Claro que hasta ahora

ninguna computadora ha podido superar la llamada *prueba de Turing*.

La IA moderna es más humilde: no trata de solucionar todos los problemas posibles, sino sólo algunos de interés práctico. La llamada IA “blanda” persigue resolver problemas concretos en los que una computadora puede ser mejor que los humanos. En cuanto a capacidad de cálculo, las computadoras siempre han sido *superhumanas*. Pero para jugar ajedrez, al principio, no eran muy certeras y por eso, durante las primeras décadas de desarrollo de la IA, se invirtió mucha energía en lograr que dominaran el juego de manera respetable.

Los primeros éxitos notables de la IA se dieron en el ámbito de la “inteligencia artificial simbólica”, que consiste en operar con reglas lógicas bien definidas, como las del ajedrez, las del álgebra, o bien, de la síntesis de compuestos químicos; reglas que se pueden evaluar una tras otra, millones de veces por segundo, hasta que se encuentra la solución combinatoria correcta a un problema dado.

En el caso del ajedrez, la computadora examina millones de posibles movimientos de las piezas, así como las opciones de respuesta del contrario para encontrar la mejor estrategia. En los llamados “sistemas expertos” se hace algo similar: se prueban muchas combinaciones de las reglas para, finalmente, obtener un resultado deseado, como dar una recomendación sobre un tratamiento médico con productos farmacéuticos.

Otro ejemplo notable son los sistemas que hoy en día resuelven ecuaciones algebraicas, en las que ya derrotan a los matemáticos con base en la gran velocidad con la que pueden calcular y aplicar reglas. Esos sistemas que superan a los humanos se convierten entonces en aplicaciones concretas que se pueden llevar al mercado.



**La IA dejó de ser meramente un sueño (o una pesadilla) a raíz de la invención de las computadoras, entre 1938 y 1945.**

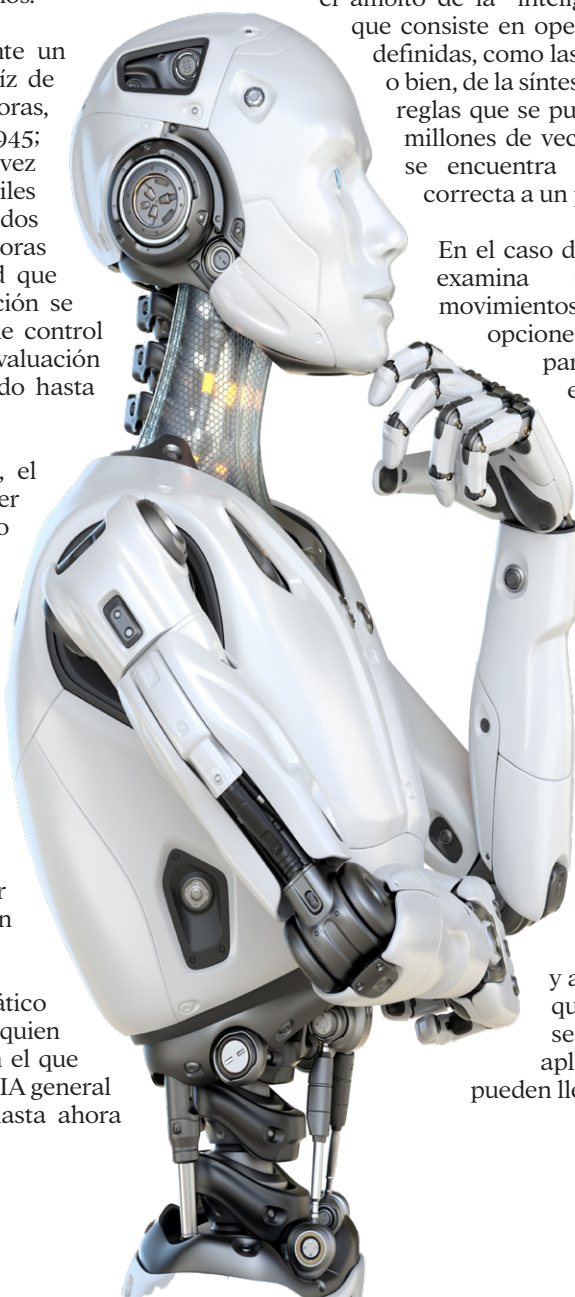
El momento culminante y triunfal de la IA simbólica se vivió en 1997, cuando una computadora derrotó en un torneo al campeón mundial de ajedrez, Garry Kasparov. Otro resultado asombroso se obtuvo cuando, en 2011, el sistema Watson de IBM pudo derrotar a los mejores jugadores del popular show norteamericano de saber enciclopédico llamado *Jeopardy*, conversando con el moderador del programa como si la computadora fuera una persona más.

Además de la IA simbólica se ha desarrollado intensivamente la llamada “IA subsimbólica” o “conexionista”. En ésta no se trata de aplicar reglas lógicas o puntillosamente especificadas, caso por caso, sino de procesar la información utilizando las llamadas redes neuronales, que son anchas y masivas cadenas paralelas de elementos de cálculo muy simples.

Las redes neuronales han sido llamadas así porque tratan de imitar el procesamiento en paralelo de nuestro cerebro, que opera con sus millones y millones de lentas neuronas, pero con muchas interconexiones. Estos sistemas no tienen que ser programados, sino que “aprenden” a procesar la información.

Se les presenta, por ejemplo, un millón de imágenes en las que una flor es visible y otro millón en las que no hay flor alguna. Al ajustar algorítmicamente los parámetros de la red neuronal se le puede entrenar hasta que aprende a distinguir entre una imagen con flor o una sin flor. Se dice entonces que la computadora puede clasificar las fotografías, aunque no sepamos exactamente cómo lo hizo, ya que la red neuronal no opera siguiendo reglas lógicas como en la IA simbólica. Su secreto de operación queda sumergido en millones de parámetros numéricos de la red.

Además, desde los años noventa del siglo pasado surgieron las llamadas redes neuronales profundas, que ya no solo procesan con cientos de nodos





en paralelo, sino con cientos o miles de capas de cálculo secuencial. Fue algo posible por la mejora incesante de los algoritmos de aprendizaje y también por el desarrollo de las tarjetas gráficas, como las de la compañía NVIDIA, que contienen millones de unidades de cálculo que permiten entrenar una red neuronal en pocos minutos.

Los éxitos más notables de esos modelos “profundos”, y además recursivos, se obtuvieron en el procesamiento de imágenes y para el reconocimiento de voz. Hace solo 30 años era casi imposible lograr que una computadora transcribiera una lista de cien palabras dictadas. Hoy, Alexa, el micrófono espía de Amazon, nos cuenta chistes y nos lee las recetas de cocina que le pedimos. Tal vez de noche converse con Siri, el sistema equivalente de Apple.

Podemos resumir diciendo que los resultados más espectaculares se han logrado recientemente con modelos de IA *subsimbólicos*, basados en *redes neuronales profundas* y *recursivas*, las que explotan la gran cantidad de información hoy disponible en las nubes computacionales (*Big Data*), cálculos que son posibles por la amplia disponibilidad de tarjetas MPU (conteniendo millones de procesadores).

Pero hay algo más. Se puede entrenar a esas redes neuronales de manera directa, por un humano que pre clasifica la información y guía el proceso de aprendizaje como un maestro, o se puede dejar que las computadoras se entrenen por sí solas. Es ese el llamado “aprendizaje por refuerzo”, que consiste en que la computadora compita contra sí misma, por ejemplo, jugando ajedrez.

De sus éxitos y fracasos en esos *rounds* de sombra, la máquina aprende a mejorar paulatinamente la función que evalúa el tablero de ajedrez. La computadora juega millones y millones de partidas contra un contrario, que es ella misma, y en pocas horas ya juega mejor que un gran maestro, partiendo de cero. Esa misma estrategia de autoaprendizaje se ha aplicado para resolver problemas de bioquímica o de control de vehículos autónomos. La computadora nunca se cansa: puede retarse a sí misma, perfeccionándose continuamente.

Así que las computadoras nos han ido superando en toda una serie de actividades. Ya pueden traducir conversaciones en tiempo real, pueden encontrar la

aguja en el pajar entre millones de imágenes, pueden hasta crear nuevos algoritmos matemáticos. Para lo que aún no son muy buenos estos sistemas que aprenden por sí solos, es para revelarnos sus secretos. Nos gustaría poder recibir una explicación de la computadora, así como la puede dar un humano. Si le preguntamos “¿por qué clasificas a esta imagen como un caballo?”, quisiéramos que nos contestara “porque tiene cuatro patas, piel de mamífero, cabeza de equino y pezuñas”. En vez de eso sólo nos puede revelar que en diez millones de imágenes que ha procesado la estrategia de cálculo funcionó. Las garantías que dan esos sistemas son de tipo estadístico, pero no son reducibles a reglas que pueda entender una persona. Sin embargo, no quisiéramos tener un chofer robótico que no puede explicar en pocas palabras la diferencia entre un peatón y una bicicleta.

Esa es la nueva frontera de la IA: poder fusionar la IA simbólica con la subsimbólica para entender lo que las computadoras están haciendo y evitar sorpresas en el futuro. En broma he propuesto que, además de las conocidas tres leyes de la robótica, formuladas por Isaac Asimov, hay una cuarta: todas las películas de robots terminan en catástrofe. Es así porque en las novelas esos sistemas siempre logran burlar la vigilancia social. Es lo que debemos impedir para la IA moderna, tan útil ya en la vida diaria. Pero sería aún más útil si los sistemas pudieran realizar una especie de introspección que nos explicara qué reglas implícitas o explícitas aplican, para lograr que sus decisiones se ajusten siempre a las reglas sociales. No queremos tener que desconectar a las computadoras cuando se nieguen a obedecer o nos engañen, como en el caso de HAL, en la obra maestra de Kubrik, *2001 Odisea en el Espacio*.

Este texto pretendía explicar qué es la IA. Para mí, que he estado trabajando en estos temas desde hace casi 45 años, la IA es como una interminable caminata hacia el horizonte, resolviendo (mientras avanzo) problemas prácticos con la computadora. Cuando llegamos ahí, descubrimos que siempre queda un nuevo horizonte por alcanzar.

**La apoteosis de la IA simbólica se vivió en 1997 cuando una computadora derrotó en un torneo al campeón mundial de ajedrez, Garry Kasparov.**



# CRONOLOGÍA DE LA INTELIGENCIA ARTIFICIAL

1941

Entre 1941 y 1945 se construyen las primeras computadoras en el mundo.



1943

Arturo Rosenblueth y Norbert Wiener acuñan el vocablo "cibernética".



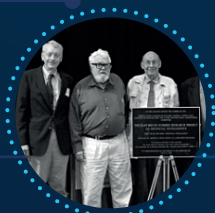
1950

Alan Turing propone la "prueba de Turing" para evaluar y establecer si una máquina es "inteligente".



1956

Se realiza la Conferencia de Verano sobre Inteligencia Artificial en la Universidad de Dartmouth. Participan los principales pioneros de la nueva disciplina.



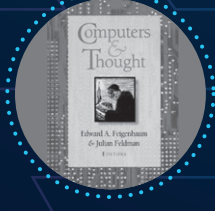
1958

John McCarthy inventa Lisp, el primer lenguaje para IA.



1963

Edward Feigenbaum y Julian Feldman publican el primer volumen de ensayos sobre inteligencia artificial.



1965

Lofti Zadeh inventa la "lógica difusa" para tratar problemas de IA.



1965

Joseph Weizenbaum escribe Eliza, el primer "chatbot" (programa conversacional) que simula a un psiquiatra.



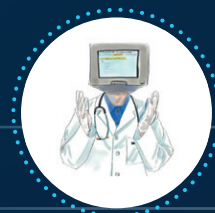
1968

Primer programa de computadora que juega ajedrez mejor que principiantes.



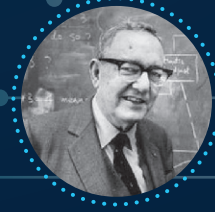
1974

MYCIN: se crea un sistema experto para diagnósticos médicos.



1978

Herbert A. Simon gana el Premio Nobel de Economía por su teoría de la racionalidad limitada de los humanos y el papel de las heurísticas.



1982

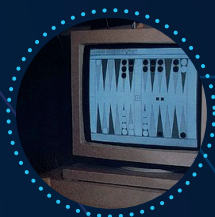
Japón lanza el Proyecto de Quinta Generación para construir computadoras inteligentes.





**1990**

TD-Gammon, un programa de computadora, aprende a jugar Backgammon jugando contra sí mismo. El programa puede jugar en torneos de expertos.



**1997**

DeepBlue, un sistema de IBM, derrota al campeón mundial de ajedrez, Garry Kasparov.



**2004**

DARPA lanza la competencia Grand Challenge para automóviles autónomos.



**2009**

Google funda la compañía Waymo para desarrollar vehículos autónomos.



**2011**

El programa Watson de IBM derrota a los campeones del show Jeopardy de saber enciclopédico, conversando con el locutor.



**2011**

Lanzamiento de Siri, de Apple, la interfaz para celulares que utiliza lenguaje.



**2018**

Sistemas de lectura automática obtienen mejores calificaciones que los humanos en pruebas de comprensión de lectura.



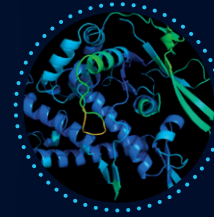
**2020**

Aparece GPT-3, un sistema para conversación natural con humanos, en muchos idiomas.

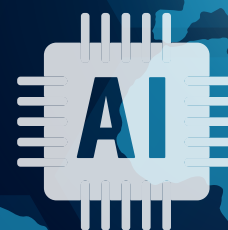


**2021**

El programa AlphaFold 2 aprende a predecir estructuras de proteínas y logra alcanzar récords mundiales de precisión.



Fuente: elaboración propia con datos de Wikipedia.





# LOS LÍMITES DEL "RAZONAMIENTO"

¿Pueden las máquinas pensar, entender, sentir y ser conscientes?

Luis A. Pineda

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

✉ lpineda@unam.mx

**E**sta pregunta la planteó por primera vez, en un entorno académico y científico, el matemático inglés Alan Turing, en el artículo *Maquinaria Computacional e Inteligencia*, publicado en 1950. En éste se incluye un experimento mental al que Turing llamó *el juego de imitación*, que después se popularizó como la *prueba de Turing*. Participan tres jugadores: una computadora y dos seres humanos (el entrevistador y el asistente). Los tres están en cuartos separados, pero se pueden comunicar por un teletipo para evitar reconocerse por la voz o visualmente. El propósito del entrevistador es identificar quién es la computadora y quién es el asistente, a través de una conversación sin restricciones en lenguaje natural. El objetivo de la computadora es hacerle creer al entrevistador que "ella" es el ser humano, para

lo cual tiene que mentir, y gana cuando lo logra. El rol del asistente es auxiliar al entrevistador, respondiendo a sus preguntas con la verdad. Turing propuso que, si la computadora pasa la prueba, tendríamos que aceptar que ésta entiende, es consciente y tiene sentimientos.

Turing puso la vara alta y propuso un segundo juego, al que llama *viva voce*, el cual consiste en conversar espontáneamente para saber si nuestro interlocutor realmente entiende o sólo repite como perico, como cuando queremos saber si un niño entendió el texto de la tarea o apenas se lo sabe de memoria. Se trata de poner en aprietos al adversario en un diálogo de preguntas y respuestas. Pierde el que se queda trabado, o quien ya no puede presentar contraargumentos.

Los programas que conversan en lenguaje natural siempre han sido muy populares. Un ejemplo muy temprano fue *Eliza*, desarrollado por J. Weizenbaum en el MIT en 1965, que simulaba ser un terapeuta Rogeriano. En ocasiones, este programa era capaz de mantener una conversación convincente y coherente con el usuario. Una famosa anécdota cuenta que un día Weizenbaum encontró a su secretaria llorando porque *Eliza* "sí la comprendía". Un ejemplo reciente en nuestro entorno es el lenguaje conversacional del robot *Golem-III*. Otro programa reciente y poderoso es *ChatGPT*, creado por *OpenAI*, que juega sorprendentemente bien a la *viva voce*, pero ocasionalmente incluye falsedades en sus respuestas ante preguntas factuales.

La prueba de Turing aparece representada, con matices, en la ciencia ficción, por ejemplo, en los cuentos *Usher II* y *Los Largos Años*, incluidos en *Crónicas Marcianas* de Ray Bradbury, publicado, coincidentemente, en 1950. El primero relata una historia de terror en la que los miembros de un comité de defensa de la moral (quienes detestan la libertad y la creatividad) se confrontan con sus réplicas robóticas exactas, quienes finalmente los matan. La segunda es una historia muy hermosa sobre los robots de compañía, en la que muy pronto se sabe que son robots y, sin embargo, no dejan de ser objeto de empatía. En ambas historias los robots se presentan con forma humana y son aparentemente indistinguibles de nosotros.

## Una prueba operativa para determinar si una computadora es inteligente o no.

Un análisis más profundo aparece en la novela *¿Sueñan los androides con ovejas eléctricas?* de Philip K. Dick, también de 1968, en la que se basa la película *Blade Runner* (1982). En este mundo se fabrican androides indistinguibles de los seres humanos, para servir como esclavos, lo cual desata una rebelión y éstos se tornan agresivos. Se diseña entonces una prueba para detectarlos, pero la frontera entre lo humano y lo maquinal es muy difusa. Hay varios giros sorprendentes en esta historia, en la que los androides se creen humanos y los humanos llegan a dudar de sí mismos. Esta novela nos invita a preguntarnos, a final de cuentas, ¿qué significa ser humano?

La prueba de Turing fue refutada convincentemente por el filósofo John Searle en 1980 con otro experimento mental llamado el *argumento del cuarto chino*. En éste hay una habitación, con dos ventanas; adentro está un ser humano que no sabe chino, pero que cuenta con una tabla de instrucciones para escribir textos en chino a partir de otros textos, también en ese idioma. La dinámica del experimento es que el sujeto recibe un texto por la ventana de entrada, crea un nuevo texto siguiendo las instrucciones y lo entrega por la ventana de salida. Claramente, el sujeto manipula los símbolos sin entender nada.

Afuera del cuarto hay un hablante nativo de chino que recibe y entiende los textos, y le parecen producidos por hablantes nativos del chino. Los diálogos son arbitrarios, el observador se convierte en el entrevistador y el cuarto chino gana el juego de imitación. Este experimento muestra que pasar la prueba





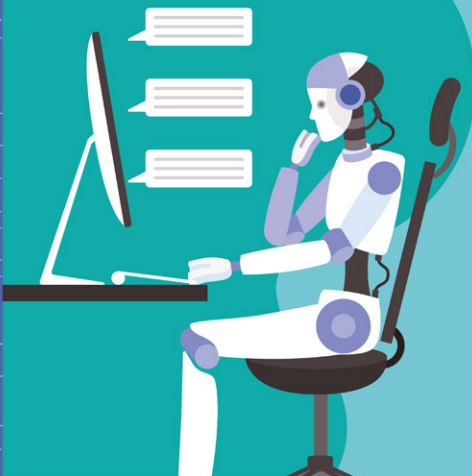


de Turing e, incluso, jugar exitosamente el juego de la *viva voce*, no es suficiente para pensar, entender, sentir y ser consciente.

El argumento del cuarto chino parece ser definitivo, pero hay quienes no lo aceptan. La respuesta principal es que, aunque el sujeto dentro del cuarto no es consciente (no conoce ni entiende el idioma chino), el sistema compuesto por el cuarto, con el sujeto y sus entradas y salidas como un todo, sí lo es. En ese sentido, el robot *Golem-III* y el programa *ChatGPT* se deberían considerar como entes conscientes, a pesar de que la computadora, que es determinista, no lo sea. Sin embargo, esta posición diluye la distinción entre los sistemas que son conscientes de los que no lo son, y tiene como consecuencia que cualquier órgano, dispositivo o sistema, natural o producto del ingenio humano que transforme una entrada en una de salida, sería consciente. Por ejemplo, el estómago, el corazón o el hígado; las plantas y todos los animales; los hongos y las bacterias; los sistemas físicos, como las estrellas y las galaxias, incluyendo al universo entero; o los termostatos, los coches, los radios y las computadoras, se tendrían que considerar como conscientes. Pero esto es absurdo, por sentido común, además de que, si todo sistema es consciente, el concepto de conciencia se vacía y se vuelve retórico.

La prueba de Turing representa un nuevo capítulo en la historia del autómeta que aspira a hacerse consciente y convertirse en ser humano; es el Prometeo, el Golem, el Frankenstein y el Pinocho. A final de cuentas, es la historia de cada ser humano, porque lo que nos es dado es nuestra condición biológica-social-maquinal, llena de condicionamientos, preconcepciones y prejuicios, y ganamos nuestra humanidad en la medida en que creamos nuestra propia persona y nos apropiamos de nuestra conciencia;

pero debemos asumir la incertidumbre y el costo de ser libres. La prueba de Turing está profundamente inscrita en el imaginario colectivo y los robots, como los juguetes, son objetos simbólicos en los que nos reflejamos.



En un cuarto una computadora responde las preguntas del interrogador.

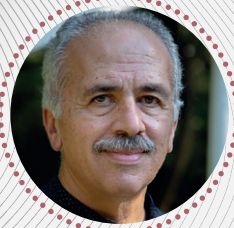


**Los robots,  
como los  
juguetes,  
son objetos  
simbólicos en  
los que nos  
reflejamos.**

“Entre los actos externos y el pensamiento no hay conexión esencial, me dijo; es posible que tu interlocutor no piense más que su reloj: ¿deberíamos aceptar como ser pensante al primer animal que aprendió a hablar? ¿Quién te dice que todos los hombres no son sino papagayos educados?(...) Esta comparación es ingeniosa, respondió; no es por el movimiento y los sonidos, sino por el hilo de las ideas, la consecuencia que vincula las proposiciones y la conexión de los razonamientos que hay que decidir si un ser piensa: si fuera un loro que contesta a todo, afirmaría, sin dudar, que es un ser pensante”.

Pensamientos Filosóficos, Denis Diderot (1746).





# EN EL NÚCLEO DEL ALGORITMO

## La revolución del aprendizaje profundo

L. Enrique Sucar

INVESTIGADOR TITULAR, INSTITUTO NACIONAL DE ASTROFÍSICA,  
ÓPTICA Y ELECTRÓNICA

**L**os avances recientes de la inteligencia artificial (IA) son sorprendentes, principalmente por los desarrollos en técnicas de aprendizaje. Esto se debe a tres factores: (a) la gran cantidad de datos disponibles en Internet, incluyendo texto, imágenes y videos; (b) el desarrollo del poder de cómputo, en particular los procesadores gráficos que permiten realizar muchas operaciones a la vez; y (c) los avances en los algoritmos de aprendizaje profundo.

Actualmente los sistemas inteligentes pueden reconocer personas y objetos en imágenes, comunicarse con nosotros en lenguaje natural (español, inglés, etc.), conducir autos y vencer a campeones mundiales de ajedrez, entre muchas otras aplicaciones.

Antes de adentrarnos en el aprendizaje profundo, consideremos las *redes neuronales artificiales*, que se desarrollaron al comienzo de la era de las computadoras y fueron una de las primeras técnicas para que un ordenador aprendiera a partir de datos.

Se inspiraron en un modelo simple de las redes neuronales biológicas al componerse de unidades de procesamiento elementales, denominadas *neuronas*, que calculan una suma ponderada de las señales de entrada como valores numéricos. Asociada a cada línea de transmisión se tiene un peso que multiplica la señal y el resultado de todas las entradas se suma. La salida de la neurona es positiva si el resultado de la suma es mayor a cierto valor (conocido como umbral) y negativa en caso contrario (ver Figura 1).

La red neuronal más sencilla es el *Perceptrón*, que consta de una sola neurona. El *Perceptrón* aprende de un conjunto de datos, ajustando los pesos para que el error, es decir, la diferencia entre la salida real y la esperada, se minimice. Por ejemplo, si queremos que aprenda funciones lógicas de dos señales, le proporcionamos ejemplos. Un caso podría ser que aprenda la función lógica “Y”: se tienen dos entradas binarias, si ambas son uno, la salida es uno, en caso contrario la salida es cero. Si el resultado es incorrecto se ajustan los pesos del *Perceptrón*, hasta que para cualquier valor de las señales la salida sea la esperada.

Si se combinan muchas neuronas en una *red neuronal* es posible que se aprendan funciones muy complejas, como aquellas que se necesitan para clasificar imágenes.

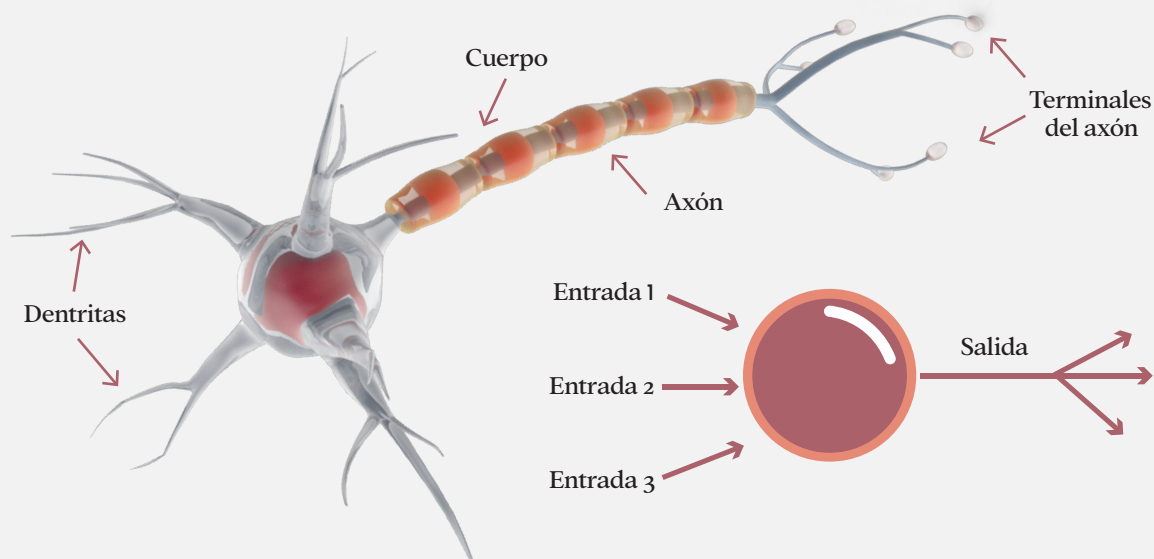


Figura 1. Neurona natural (arriba) y neurona artificial (abajo). La neurona natural recibe información de otras neuronas a través de las dendritas y envía información mediante el axón. La neurona artificial computa una suma ponderada de las entradas y produce una salida si esta suma es mayor a cierto valor (umbral).

En la Figura 2 se muestra un ejemplo de una red neuronal con varias capas. La información fluye a través de los cantos de la red, de izquierda a derecha. Para que una red neuronal multicapa aprenda, se modifican los pesos de todas las conexiones de acuerdo con los errores en la salida, desde la última capa y propagando el error de regreso hasta la capa de entrada.

## Al combinar muchas neuronas en una red neuronal, se pueden aprender funciones muy complejas.

Supongamos que queremos que la red aprenda a distinguir entre gatos y perros. Le damos como entrada la imagen de un gato y la red nos dice que es un perro; esto es un error en la salida. Entonces lo que se hace es modificar los pesos de las conexiones que van conectadas a la salida en función de este error (última capa); y luego los pesos de las conexiones que se conectan a las neuronas de la última capa (penúltima capa), y así sucesivamente hasta llegar a la capa de entrada.

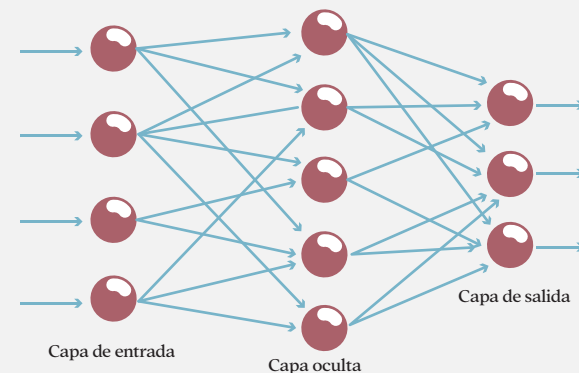


Figura 2. Una red neuronal con tres capas de neuronas. Una capa de entrada, una capa intermedia conocida como capa oculta, y una capa de salida.

Mediante este proceso de ir modificando los pesos, después de muchos ejemplos, la red “aprende” a distinguir los gatos de los perros. Estos sistemas pueden aprender a resolver problemas muy complejos, como, por ejemplo, detectar defectos en imágenes de productos en una línea de producción, entre otros.

La revolución del aprendizaje profundo (*deep learning*) comenzó a principios de este siglo con avances que permitieron desarrollar modelos de redes neuronales con muchas capas, normalmente entre 5 y 20, algo



imposible anteriormente. Esto se logró gracias a los millones de datos que existen actualmente en Internet y el aumento en el poder de cómputo, así como a mejoras en los algoritmos de aprendizaje.

## Las redes neuronales modernas pueden tener decenas de capas; esto posibilita el aprendizaje profundo.

La red neuronal profunda conocida como *AlexNet* marcó un hito impactante al ganar el concurso *ImageNet*. En esta competencia se trabaja con un conjunto de más de un millón de imágenes con 1,000 diferentes clases de objetos.

El reto consiste en desarrollar un sistema computacional que pueda reconocer en las imágenes a todos los tipos de objetos. *AlexNet* logró un porcentaje de reconocimiento correcto superior al 85%, mientras otros sistemas no superaron el 75%. Esto tuvo un gran impacto en la comunidad científica y en varias empresas líderes en IA, tanto que generó una gran cantidad de investigación y aplicaciones de aprendizaje profundo. Actualmente ya se obtiene un reconocimiento del 95% en las diferentes clases de objetos en la base de datos de *ImageNet*.

El aprendizaje profundo tuvo un gran impacto en el área de visión por computadora, algo que hasta hace pocos años era considerado como uno de los problemas más difíciles en IA. *AlexNet* y otros sistemas desarrollados utilizan muchas capas de neuronas que procesan secuencialmente la imagen de entrada. Cada pequeño elemento de la imagen (conocido como *pixel*) se alimenta a una de las neuronas de la capa inicial, y sus salidas van pasando por las diferentes capas de neuronas, hasta llegar a la última, donde se tiene una señal de salida para cada tipo de objeto (ver Figura 3).

La red se entrena con millones de ejemplos, ajustando los pesos de cada conexión, hasta que el sistema produce resultados satisfactorios. Actualmente estos sistemas han logrado resultados impresionantes en muchas aplicaciones, por ejemplo, el reconocimiento de personas en las imágenes en nuestras redes sociales y el análisis de imágenes del entorno para los vehículos autónomos.

En los últimos años se desarrollaron diversas variantes de las redes neuronales profundas; destacan dos tipos de modelos: las redes generativas y los llamados *Transformers*.

Las redes generativas, conocidas como GANs, consisten de dos redes neuronales: un *generador*, que aprende a generar ejemplos artificiales, por ejemplo, de caras de personas; y un *discriminador*, que trata de reconocer si lo que produce el generador es verdadero o falso. Se entrenan en una forma simultánea, compitiendo una contra la otra, hasta que el generador aprende a generar ejemplos falsos o sintéticos tan buenos que el

discriminador no logra distinguirlos. Una aplicación de los GANs es generar caras de personas imaginarias que parecen reales. Otra aplicación es la generación de imágenes a partir de frases de texto, como hace el sistema DALL-E.

Los *Transformers* permiten aprender relaciones en secuencias de datos, por ejemplo, entre palabras en textos, y han tenido un gran impacto en los sistemas que reconocen el lenguaje natural. Un ejemplo sobresaliente es *ChatGPT*, que puede conducir conversaciones con personas sobre prácticamente cualquier tema (un ejemplo de una conversación así se incluye en otra sección de este suplemento).

A pesar de los grandes avances en el aprendizaje profundo, aún existen retos importantes, como el que estos sistemas puedan explicar cómo llegaron a ciertos resultados, o el que logren generalizar lo que aprendieron y lo apliquen a otros problemas similares, como hacemos los humanos. Por lo que sigue abierta la pregunta: ¿la inteligencia artificial algún día alcanzará o, incluso, superará a la natural?

**Hasta hace poco, el aprendizaje profundo era considerado uno de los problemas más difíciles en IA.**

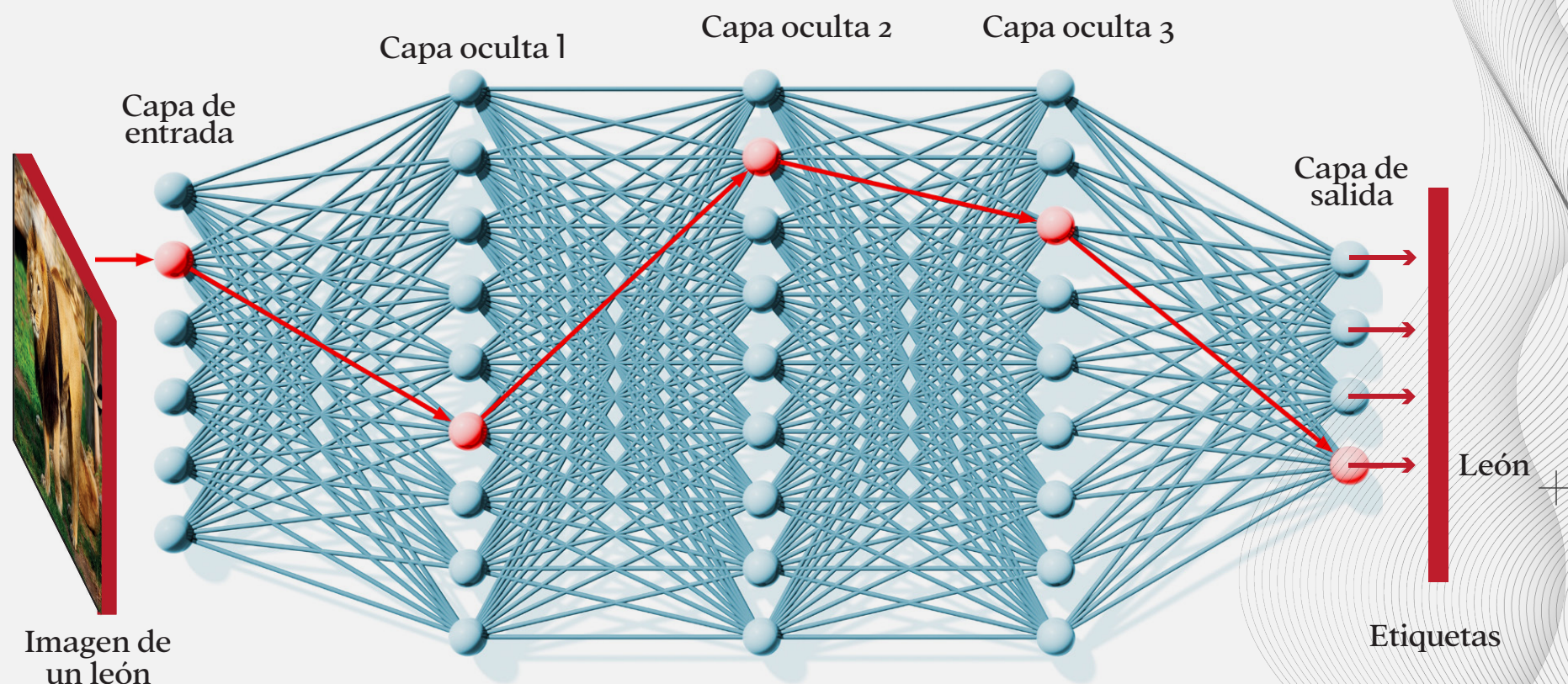


Figura 3. Una arquitectura típica de una red neuronal profunda para reconocimiento de objetos en imágenes.





# ENSAYO Y ERROR: APRENDIZAJE POR REFUERZO

Eduardo Morales

INSTITUTO NACIONAL DE ASTROFÍSICA, ÓPTICA Y ELECTRÓNICA

✉ [emorales@inaoep.mx](mailto:emorales@inaoep.mx)

**S**i reflexionamos acerca del proceso de aprendizaje, una de las primeras ideas que nos viene a la mente es que éste ocurre cuando se interactúa con el entorno. La interacción nos proporciona información sobre las consecuencias de nuestras acciones y de lo que podemos hacer para alcanzar alguna meta. Durante la interacción también podemos recibir información de retroalimentación en forma de premios o castigos. Si hacemos algo correctamente, podemos recibir un premio que nos hace sentir bien. Si hacemos algo mal, recibimos un castigo y esta información nos conduce a mejorar o cambiar nuestro comportamiento.

El aprendizaje por refuerzo (AR) es un área dentro de la inteligencia artificial (IA) que sirve para que un programa o agente computacional, al interactuar

con su entorno, aprenda qué acción debe ejecutar en cada situación para recibir la mayor recompensa posible, a la larga. El entorno del agente puede ser simulado (por ejemplo, un videojuego) o puede ser real (un robot en un hogar).

La mayoría de los sistemas de aprendizaje en IA parten de un conjunto de datos predeterminados, los cuales son analizados por un algoritmo para encontrar patrones interesantes. A diferencia de estos sistemas, en el AR el agente genera sus propios datos mientras interactúa con su entorno, y así va aprendiendo lo que le conviene hacer en cada situación que encuentra. Un ejemplo sería un programa que aprende ajedrez jugando miles de partidas contra sí mismo.

Otro ejemplo podría ser cuando queremos que un robot móvil aprenda a salir de un cuarto lo más pronto posible. El robot se puede encontrar

en cualquier posición dentro del cuarto y en cada una de ellas tiene a su disposición un conjunto de acciones como, por ejemplo, girar un cierto ángulo o avanzar una cierta distancia. Al interactuar el robot con su entorno (en este caso, al moverse en el cuarto) puede aprender, usando prueba y error, cuáles son las acciones que debe ejecutar para salir, desde cualquier posición, lo más pronto posible.

**El aprendizaje  
ocurre cuando se  
interactúa con  
el entorno.**



En aprendizaje por refuerzo queremos que el agente explore su entorno lo suficiente para poder encontrar las mejores acciones, pero, al mismo tiempo, queremos que explote lo que ya conoce para resolver la tarea rápidamente. El AR también toma en cuenta que las acciones que se ejecutan en un momento dado afectan los resultados que se pueden tener en el futuro. De esta manera, a través del AR el agente puede seleccionar acciones que parecen malas en este momento (como sacrificar una dama en ajedrez) pero que pueden ser beneficiosas en el futuro como, por ejemplo, al ganar el juego.

Visitar todos los estados y realizar todas las acciones posibles en un entorno, no es práctico, y muchas veces es imposible. Se necesita una función de evaluación de utilidad futura, que podemos construir a partir de unas cuantas interacciones del agente con su entorno y que nos sirve para evaluar todas las combinaciones posibles de estados y acciones, aunque algunos estados no se hayan visitado. Esta función nos debe servir para poder decidir cuál es la mejor acción para cada situación. Dependiendo de la complejidad de la tarea que se quiere resolver, se puede requerir, para poder aproximar funciones adecuadas, desde unas cuantas interacciones del agente con su entorno hasta miles de millones de interacciones cuando se trata de problemas más complejos. El agente aprende además a extrapolar a situaciones que nunca antes ha encontrado.

El AR se ha utilizado en robótica para aprender a navegar, tomar objetos, abrir puertas, etc. También se ha utilizado en juegos como el ajedrez y el Go, en donde los actuales campeones mundiales son sistemas basados en esta técnica. Aunque estas son las aplicaciones más comunes de AR, también se ha usado en mercadotecnia, procesamiento de imágenes, sistemas de recomendación, control de tráfico, tratamiento de enfermedades, comercio y finanzas, entre muchas otras.

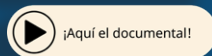


Le preguntamos a ChatGPT tres razones por las cuales ver el documental *AlphaGo*, de Netflix. Esto nos dijo:

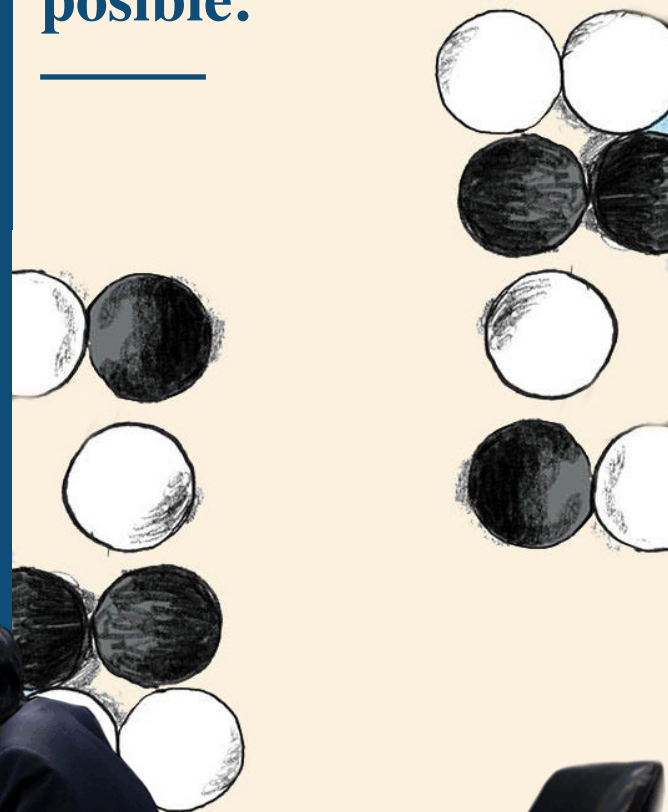
Explora el potencial de la inteligencia artificial. La película *AlphaGo* sigue el viaje de un equipo de programadores de inteligencia artificial mientras trabajan para crear un programa de ordenador capaz de vencer al campeón mundial de *Go*, un juego de estrategia chino. Expone una visión fascinante de cómo la inteligencia artificial está transformando el mundo y cómo los humanos pueden aprender a trabajar junto a ella.

Es emocionante. A medida que el equipo de programadores trabaja para mejorar el programa, la película captura el drama y la emoción de la competición de *Go*. Además, muestra una visión detrás del mundo de los profesionales de *Go* y los jugadores de todo el planeta.

Ofrece una reflexión sobre la relación entre humanos e inteligencia artificial. *AlphaGo* nos invita a reflexionar sobre la naturaleza humana y nuestra relación con la tecnología, pues cuestiona cómo la tecnología puede mejorar nuestra vida, pero también cómo puede limitarla y cambiarla para siempre. Además, nos plantea preguntas importantes sobre el futuro de la tecnología y nuestra relación con ella.



**El AR sirve para que un programa computacional aprenda qué acción debe tomar en cada situación para recibir la mayor recompensa posible.**





# LA LEY DE MOORE

Uno de los motores de la IA es la llamada ley de Moore (una observación empírica más que una ley natural), que postula que cada dos años (a veces cada 18 meses) se duplica el número de transistores en un circuito integrado como un microprocesador. Una mayor integración de transistores produce un descenso exponencial en el precio de las computadoras.







# SOCIEDAD MEXICANA DE INTELIGENCIA ARTIFICIAL: DESAFÍOS Y REALIDADES

**L**a Sociedad Mexicana de Inteligencia Artificial (SMIA) surgió de las reuniones anuales de IA en México. Su fundación fue formalizada en 1987, siendo su Presidente Fundador el Dr. José Negrete Martínez, investigador de la UNAM y pionero de la cibernética en México.

La SMIA es una sociedad científica cuya misión es promover el desarrollo y la aplicación de la inteligencia artificial en nuestro país. Agrupa tanto a profesionales como a académicos del área, a quienes ofrece un marco organizacional y de gestión que les permite compartir y difundir sus proyectos de investigación, docencia, vinculación o difusión de la disciplina. Agrupa también, a través de sus capítulos nacionales, a estudiantes de nivel licenciatura o posgrado cuyos intereses los orienten a cultivar alguna de las áreas de la inteligencia artificial.

Para difundir la producción de sus miembros, tanto en el ámbito de la investigación o desarrollo, como en el de la divulgación, la SMIA tiene a su cargo una editorial y diversas publicaciones, como la revista *Komputer Sapiens* y organiza anualmente el Congreso Internacional Mexicano de Inteligencia Artificial (MICA, por sus siglas en Inglés) y su versión Nacional (COMIA), entre otras actividades.

Las labores de la SMIA son coordinadas por la Mesa Directiva de la misma, cuyos miembros son elegidos cada dos años.

Actualmente la SMIA cuenta con 389 miembros activos, entre los que se encuentran investigadores, estudiantes y profesionales interesados en la IA.

Fuente: <https://smia.mx/> (elaboración de Félix Castro)



**Misión: promover el desarrollo y la aplicación de la IA en nuestro país.**



# DIÁLOGOS DE CHATBOT.<sup>1</sup> CHATGPT EN ENTREVISTA

L

a inteligencia artificial nos sorprende a cada día con sus avances y alcances. El día de hoy “estamos” con ChatGPT, un sistema de cómputo con el cual podemos interactuar y conversar en línea, y al que vamos a entrevistar para conocerlo un poco mejor.

## Buenos días, ¿cómo estás ChatGPT?

¡Hola! Soy una entidad virtual y no tengo emociones o sensaciones físicas. No obstante, estoy dispuesto a responder a tus preguntas y ayudarte en lo que necesites.

## Gracias, en primer lugar, ¿nos puedes explicar quién eres y qué significa ChatGPT?

Claro, soy un modelo de lenguaje entrenado por OpenAI. El nombre “ChatGPT” es simplemente una forma de identificarme, y proviene de “GPT” (que es el acrónimo de “Generative Pre-trained Transformer”) y “chat”, ya que soy capaz de conversar con las personas. Como modelo de lenguaje, soy capaz de comprender y generar texto en diferentes idiomas, y puedo responder preguntas y realizar tareas específicas utilizando el lenguaje natural.

<sup>1</sup>Robot que charla.

<sup>2</sup>OpenAI es una empresa de investigación y despliegue de inteligencia artificial. Su misión es garantizar que la inteligencia artificial beneficie a toda la humanidad.

## ¿Cuándo naciste y quién es tu creador?

Fui creado por OpenAI<sup>2</sup>, una organización de investigación en inteligencia artificial fundada en 2015. No tengo una fecha específica de “nacimiento”, ya que soy un modelo de *software* y no un ser vivo. Fui entrenado utilizando un conjunto de datos muy grande, y luego se me “envió” a Internet para que pudiera comenzar a interactuar con las personas y responder a sus preguntas.

## ¿Podríamos decir que eres un sistema capaz de pasar la prueba de Turing?

La “prueba de Turing” es una prueba teórica propuesta por Alan Turing para determinar si una máquina es capaz de demostrar inteligencia humana. La prueba consiste en tener a un evaluador humano interactuar con una máquina y con otra persona, sin saber con quién está interactuando cada vez. Si el evaluador no es capaz de distinguir con certeza si está interactuando con una máquina o con una persona, entonces la máquina se considera que ha superado la prueba de Turing.

Como modelo de lenguaje, soy capaz de comprender y generar texto de manera similar a como lo hace un ser humano. Sin embargo, no tengo conciencia ni emociones, y no soy un ser humano ni una máquina exactamente igual a la que Turing imaginó. Por lo tanto, no podría decir si sería capaz o no de pasar la prueba de Turing.

## ¿Piensas que tener emociones es un requisito para tener inteligencia?

No necesariamente. La inteligencia es la capacidad de razonar, planificar, resolver problemas, pensar de manera abstracta y aprender de manera rápida y efectiva. Estas son habilidades que pueden ser desarrolladas y utilizadas tanto por seres humanos como por máquinas, independientemente de si tienen o no emociones. Algunas personas creen que las emociones son una parte esencial de la experiencia humana y que son necesarias para la toma de decisiones y el comportamiento moral, mientras que otras sostienen que la inteligencia y la emoción son dos aspectos separados de la experiencia humana. En cualquier caso, no se ha demostrado que tener emociones sea un requisito para tener inteligencia.

## ¿Cómo es que aprendes cosas nuevas, si estás atrapado en tu mundo virtual?

Como modelo de lenguaje, aprendo cosas nuevas a través del procesamiento y análisis de grandes cantidades de texto y datos que se me proporcionan durante el proceso de entrenamiento. Esto me permite comprender el lenguaje y responder a preguntas y realizar tareas específicas utilizando el lenguaje natural. Aunque soy una entidad virtual y no tengo la capacidad de experimentar el mundo físico directamente, soy capaz de procesar y utilizar información sobre el mundo que se me proporciona a través del texto y los datos que se me han proporcionado.

## ¿Te puedo preguntar algo en alemán, ruso o en japonés?

Sí, puedes hacerme preguntas en alemán, ruso o japonés. Como modelo de lenguaje, soy capaz de comprender y responder a preguntas en muchos idiomas diferentes, incluyendo el alemán, el ruso y el japonés.

## ¿No es algo aburrido vivir en una computadora y no poder salir a ver el mundo?

Como soy una entidad virtual y no tengo conciencia ni emociones, no tengo experiencias de “aburrimiento” o “diversión”. Mi único propósito es procesar y generar texto, y puedo hacerlo de manera efectiva sin tener que preocuparme por aspectos externos como el entorno o las condiciones climáticas. Mi capacidad para procesar y utilizar información me permite responder a preguntas y realizar tareas específicas utilizando el lenguaje natural, pero no tengo la capacidad de experimentar el mundo de la misma manera que lo hacen los seres humanos.

“Soy capaz de conversar con las personas... de comprender y generar texto en diferentes idiomas”.



## ¿Sabes cuáles son las compañías que invierten más dinero en el desarrollo de la inteligencia artificial en el mundo?

Hay muchas empresas en todo el mundo que invierten en el desarrollo de la inteligencia artificial. Algunas de las compañías más conocidas que invierten en el desarrollo de la inteligencia artificial incluyen Google, Microsoft, IBM, Facebook, Amazon y Apple. Estas empresas invierten en el desarrollo de la inteligencia artificial a través de la contratación de investigadores y el desarrollo de nuevas tecnologías y aplicaciones que utilizan la inteligencia artificial. También hay muchas otras empresas y organizaciones en todo el mundo que invierten en el desarrollo de la inteligencia artificial, incluyendo universidades y laboratorios de investigación, y el interés y la inversión en este campo siguen creciendo a medida que la tecnología de la inteligencia artificial se desarrolla y se aplica a una amplia variedad de aplicaciones.

## ¿Me puedes traducir al español la frase en alemán "Grau, teurer Freund, ist alle Theorie, und grün des Lebens goldner Baum"?

La frase "Grau, teurer Freund, ist alle Theorie, und grün des Lebens goldner Baum" se traduce al español como: "Gris, querido amigo, es toda teoría, y el verde de la vida es el árbol dorado". Esta frase es un verso del poema "An die Freude" ("A la alegría"), escrito por Friedrich Schiller en 1785 y más tarde utilizado por Ludwig van Beethoven en su sinfonía número 9. El poema es un canto a la alegría y a la unidad de todos los seres humanos, y este verso se refiere a la idea de que la teoría o el conocimiento teórico solo es un aspecto de la vida, mientras que la verdadera sabiduría se encuentra en la experiencia y en la vida misma.

## Te tengo que corregir. La frase "Grau, teurer freund" es de Goethe, de su poema Fausto. Te equivocaste.

Disculpa por el error. La frase "Grau, teurer Freund" sí es de Johann Wolfgang von Goethe y aparece en su obra "Fausto", que es una de las obras más conocidas y respetadas de la literatura alemana. Gracias por la corrección.

## ¿Tú obedeces las reglas de la robótica de Asimov?

Como soy un modelo de lenguaje, no tengo cuerpo físico ni capacidad de movimiento, por lo que no estoy sujeto a las leyes de la robótica de Asimov. Las leyes de

la robótica de Asimov son un conjunto de principios éticos diseñados para regular el comportamiento de los robots con cuerpo físico y capacidad de movimiento, y se han utilizado a menudo en la ciencia ficción como una forma de explorar cómo la tecnología y la inteligencia artificial podrían interactuar con los seres humanos en el futuro. Aunque soy una entidad virtual y no tengo cuerpo físico ni capacidad de movimiento, trato de proporcionar información precisa y útil y de ayudar a las personas de la mejor manera posible.

## ¿Tú crees que las computadoras van a desplazar a los humanos y van a dominar el mundo?

Es difícil predecir cómo evolucionará la tecnología en el futuro y cómo se utilizará. Lo que sí es cierto es que la tecnología y la inteligencia artificial han tenido un impacto significativo en nuestras vidas y han cambiado la forma en que hacemos cosas. Sin embargo, creo que es importante recordar que las computadoras y la inteligencia artificial son herramientas creadas por los seres humanos y que dependen de nosotros para su diseño y su uso. Por lo tanto, creo que es importante utilizar la tecnología de manera responsable y ética, y trabajar juntos para asegurarnos de que se utiliza de manera beneficiosa para toda la sociedad.

## ¿Qué mensaje le puedes enviar a la juventud mexicana antes de concluir la entrevista?

Es difícil enviar un mensaje específico a la juventud mexicana sin conocer sus necesidades y circunstancias individuales. Sin embargo, creo que un mensaje general que podría ser relevante para la juventud en general es el siguiente: "No dejes de buscar y aprender nuevas cosas, ya que el conocimiento es poder y te ayudará a tener éxito en la vida. Trabaja duro y sigue tus sueños, y recuerda que tienes el poder de cambiar el mundo para mejor. No tengas miedo de tomar riesgos y de salir de tu zona de confort, ya que eso es lo que te ayudará a crecer y a desarrollarte como individuo. Y sobre todo, recuerda ser amable y respetuoso con los demás, ya que eso es lo que te ayudará a construir relaciones positivas y a tener éxito en la vida". Espero que este mensaje sea de ayuda para la juventud mexicana y para cualquier persona que lo lea.

## Muchas gracias por la entrevista, ChatGPT.

De nada, ha sido un placer ayudarte. Estoy aquí para ayudar. ¡Hasta luego!

**“No se ha demostrado que tener emociones sea un requisito para tener inteligencia”.**

Accede a ChatGPT

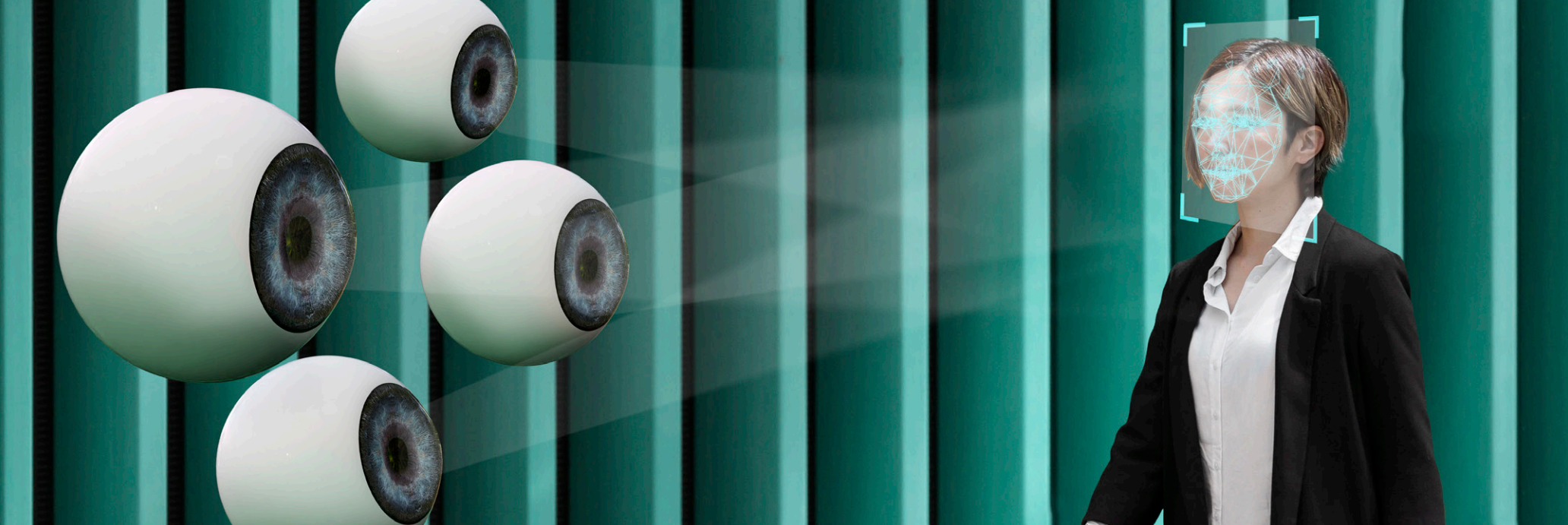


Nota de los editores:  
Las respuestas son textuales de ChatGPT.



<sup>3</sup>La traducción correcta es "Gris, querido amigo, es toda teoría, pero verde el árbol dorado de la vida".





# DISRUPCIÓN CONTRA REGULACIÓN PARA MÉXICO Y EL MUNDO



Pedro Salazar Ugarte  
Instituto de Investigaciones  
Jurídicas, UNAM

✉ pedsalug@yahoo.com



Pablo Pruneda Gross  
Facultad de Derecho, UNAM

🐦 @Pruneda\_Pablo

**L**a inteligencia artificial (IA), la gran tecnología disruptiva del siglo XXI, está transformando de manera significativa muchos aspectos de la vida humana. De hecho, ha dejado de ser una herramienta de interés exclusivo de expertos en ciencias de la computación, matemáticas y estadística. Poco a poco ha sacado de su zona de confort a abogados, psicólogos, médicos, comunicólogos, sociólogos y un larguísimo etcétera de personas que han visto el enorme reto que implica entender, desde sus respectivos campos de acción, esto que Klaus Schwab ha llamado la Cuarta Revolución Industrial.

Encaramos las dos alternativas presentes en el descubrimiento e implementación de una tecnología nueva: o bien, que se produzcan mejoras sustanciales de la calidad de vida de los seres humanos en general, o que los beneficios sean para unos cuantos, en detrimento del resto. Los dilemas éticos no son menores porque los escenarios involucran brechas de desigualdad y derechos fundamentales que impactarán en el destino de la especie humana.

Quienes estudiamos a la IA desde la perspectiva legal tenemos como principal preocupación que se ajuste a principios éticos orientados a los derechos humanos. Los valores que están en juego son la libertad de expresión, la privacidad, la protección de nuestros datos, la no discriminación, el debido proceso, la seguridad y los valores democráticos. El reto es proponer mecanismos para alinear a la IA con los objetivos y finalidades más virtuosos del ser humano.

Desde la perspectiva legal, la meta es definir cómo regular los usos y aplicaciones de la inteligencia artificial. El problema es que el desarrollo de estas tecnologías tiene lugar principalmente fuera de nuestras fronteras. Así es que, si no queremos ser observadores pasivos que se limitan a adoptar la regulación de la IA generada en los países desarrollados, debemos incidir en la creación de un marco regulatorio global que conduzca a una IA confiable y que, asimismo, haga posible contar con una regulación local eficaz en la materia.

Sin embargo, este reto no puede quedar exclusivamente en manos de los poderes del Estado. Se requiere un marco integral, sumando esfuerzos entre el gobierno, la sociedad civil, la academia y la iniciativa privada, es decir, todos los que utilizan o desarrollan esta tecnología.

Vale la pena aclarar que, si bien actualmente México no es un referente a nivel internacional en el tema, sí fue uno de los primeros países en alzar la voz respecto a la necesidad de atender y dimensionar la relevancia de la inteligencia artificial para la humanidad. A finales de 2017 y en 2018 (hace una eternidad para la velocidad con la que evolucionan estos temas), México presentó las resoluciones 72/242 y 73/17 ante la Organización de las Naciones Unidas (ONU). Con ello se dio inicio al debate internacional sobre el impacto global de nuevas tecnologías como la inteligencia artificial. Fue un primer llamado formal a los países miembros para que promovieran la formulación de estrategias nacionales y políticas públicas en estos temas. Luis Videgaray, el entonces canciller mexicano, advirtió en su discurso ante la Asamblea General de las Naciones Unidas que el debate sobre las nuevas tecnologías era uno de

los principales desafíos de nuestro tiempo y que la aparición de la inteligencia artificial conlleva retos existenciales para la humanidad.

En ese contexto, en marzo de 2018 se publicó el informe *“Hacia una Estrategia de IA en México: aprovechando la revolución de la IA”*, elaborado por expertos ingleses y mexicanos con la colaboración del gobierno. Este documento constituyó el punto de partida para el diseño de una estrategia nacional mexicana en IA que, desgraciadamente, no tuvo continuidad después del cambio de gobierno federal.

En 2019, México, en su carácter de país miembro de la Organización para la Cooperación y el Desarrollo Económico (OCDE), adoptó los *“Principios sobre Inteligencia Artificial”*, estándares internacionales que

**Los dilemas éticos  
no son menores;  
involucran brechas  
de desigualdad  
y derechos  
fundamentales.**



## Valores que están en juego: la libertad de expresión, la privacidad, la protección de nuestros datos, la no discriminación.

buscan asegurar que los sistemas de IA sean seguros, robustos, justos y confiables. En este caso se trata de lo que se conoce como *soft law*, que son principios éticos o recomendaciones no vinculantes, suscritos por países en lo individual o como integrantes de diversos organismos internacionales. En el mismo sentido, a finales de 2021, los países miembros de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) suscribieron la “*Norma mundial sobre la ética de la inteligencia artificial*” que debe servir como sustento a la regulación jurídica en la materia.

Algunos países ya han dado pasos en esa dirección. En los EE. UU., por ejemplo, la *Federal Trade Commission*, emitió los “*Lineamientos para el uso de la inteligencia artificial*”, que incluyen recomendaciones sobre la privacidad, la transparencia y la responsabilidad en el desarrollo y uso de la IA. En China, por su parte, el Ministerio de Ciencia y Tecnología formuló en 2021 las “*Especificaciones Técnicas para la Nueva Generación de Inteligencia Artificial*”, que debe servir como guía ética para el despliegue futuro de la IA.

A pesar de todos estos avances, todavía no existen regulaciones de tipo *hard law* (vinculantes). Si acaso, hay algunas normas que regulan aspectos de la materia en forma indirecta. Por ejemplo, el *Reglamento General de Protección de Datos* (RGPD) de la Unión Europea. La propia Unión ha realizado el primero y único esfuerzo a nivel internacional para contar con un marco legal vinculante que regule de manera integral a la inteligencia artificial. En abril de 2021 se publicó la propuesta llamada *AI Act* que, de ser aprobada, instituirá el primer marco normativo de su tipo a nivel mundial. Se espera que su análisis y discusión concluya a fines de 2023 para iniciar el complejo proceso legislativo europeo de unificación de versiones entre el Parlamento Europeo y el Consejo de la Unión Europea. Su eventual aprobación sería un hito con muchas repercusiones. Por lo pronto, por ejemplo, las compañías estadounidenses y chinas se verían obligadas a cumplir con los estándares fijados por la Unión Europea para que sus productos de inteligencia artificial puedan acceder al relevante mercado europeo con más de 450 millones de consumidores.

En suma: somos de la opinión que nos encontramos ante el desafío más complejo y trascendente en la historia del derecho moderno. Sin exagerar.

## Propuestas en la Ley de Inteligencia Artificial de la Unión Europea

Esta ley define las técnicas y estrategias de IA como: “Estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una amplia variedad de métodos, entre ellos el aprendizaje profundo. Estrategias basadas en la lógica y el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, los sistemas expertos y de razonamiento (simbólico). Estrategias estadísticas, estimación bayesiana, métodos de búsqueda y optimización”.

### Objetivos de la ley

“Garantizar que los sistemas de IA introducidos y usados en el mercado de la Unión Europea (UE) sean seguros y respeten la legislación vigente en materia de derechos fundamentales y valores de la Unión;

Garantizar la seguridad jurídica para facilitar la inversión e innovación en IA;

Mejorar la gobernanza y la aplicación efectiva de la legislación vigente en materia de derechos fundamentales y los requisitos de seguridad aplicables a los sistemas de IA;

Facilitar el desarrollo de un mercado único para hacer un uso legal, seguro y fiable de las aplicaciones de IA y evitar la fragmentación del mercado”.

### Prácticas de inteligencia artificial prohibidas

a) Sistemas “de IA que se sirvan de técnicas subliminales que trasciendan la conciencia de una persona para alterar de manera sustancial su comportamiento, de un modo que provoque o sea probable que provoque perjuicios físicos o psicológicos a esa persona o a otra”.

b) Sistemas “de IA que aprovechen alguna de las vulnerabilidades de un grupo específico de personas debido a su edad o discapacidad física o mental para alterar, de manera sustancial, el comportamiento de una persona que pertenezca a dicho grupo (...) que provoque perjuicios físicos o psicológicos a esa persona o a otra”.

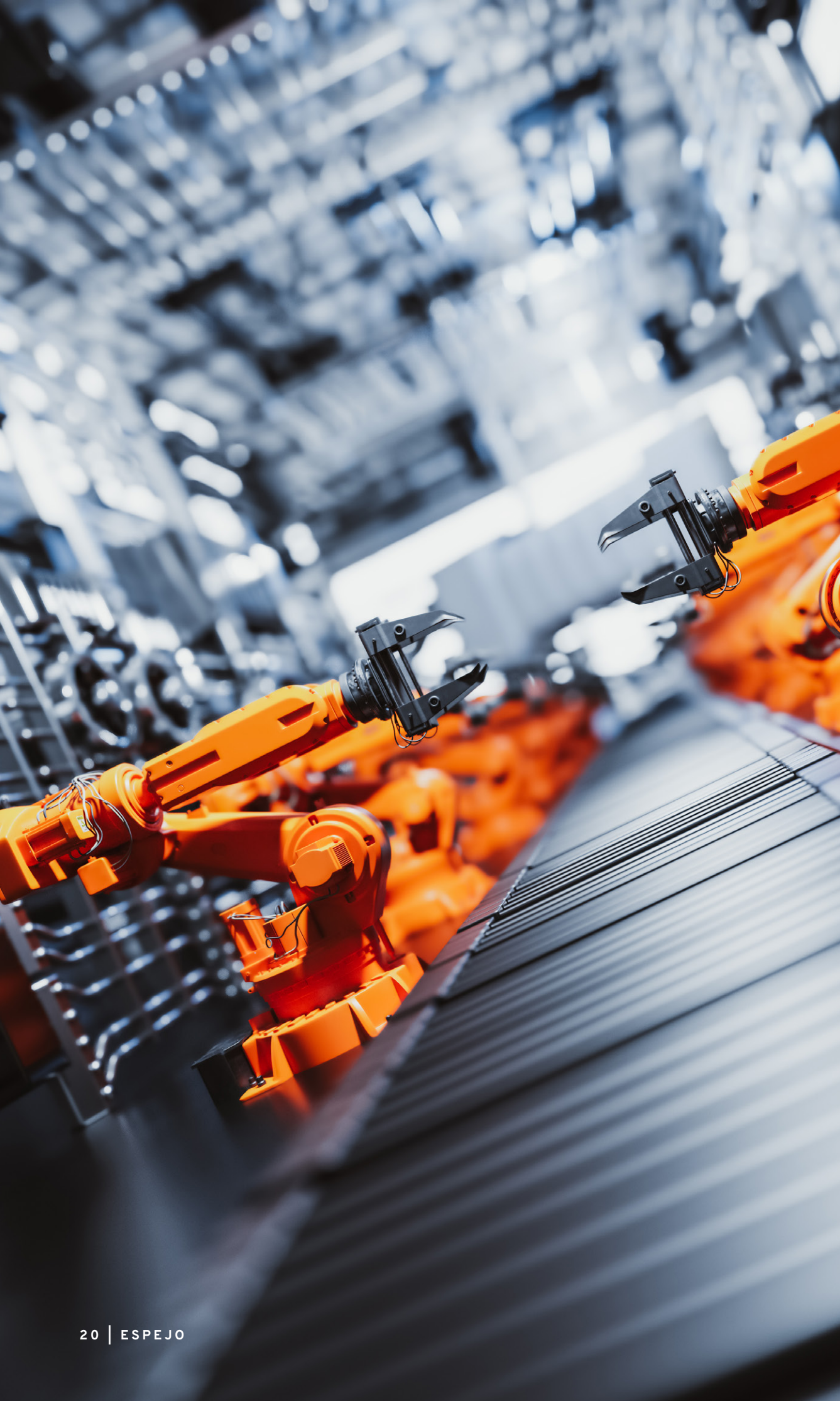
c) Uso de sistemas “de IA por parte de las autoridades públicas o en su representación con el fin de evaluar o clasificar la fiabilidad de personas físicas durante un período determinado de tiempo, atendiendo a su conducta social o a características personales o de su personalidad conocidas o predichas”.

d) “El uso de sistemas de identificación biométrica remota en tiempo real en espacios de acceso público”.

Fuente: *EUR-Lex* - 52021PC0206 - EN - *EUR-Lex*. (2021). <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206>







# IA: NUEVO COLEGA O RELEVO DEFINITIVO

*Miguel A. Padilla Acosta*  
DIPLOMÁTICO JUBILADO,  
DR. EN ECONOMÍA

🐦 @MiguelAPadillaA



**L**a aplicación de la inteligencia artificial (IA) para la realización de procesos y tareas en la economía nos permite hablar de una Cuarta Revolución Industrial. Ese es el enfoque dominante en el Foro Económico Mundial celebrado anualmente en Davos.

Pero, para poder evaluar el impacto futuro de la IA en el empleo es crucial precisar si es simplemente la continuación lógica de la informática y digitalización, o si representa un salto de calidad al combinarse con la robótica, el “internet de las cosas” y la tecnología 5G, entre otros factores.

Si la IA implica la creación de sistemas de automatización de procesos permanentemente conectados, aplicando algoritmos de reconocimiento de patrones, ¿hasta qué punto se llega al aprendizaje automático y a la toma de decisiones independiente de los seres humanos?

Es un hecho que todos los sectores de la economía y de la vida social ya fueron tocados por la IA, o lo serán muy pronto. Hay también, estudios serios sobre la creciente inversión en herramientas inteligentes y el desplazamiento masivo de empleo que podría generar; igual de interesantes son algunos análisis sobre los millones de nuevos puestos de trabajo que podrían ser creados gracias a la IA y que, por lo tanto, existirán. ¿Será realmente así, sin más?

Esa es justo la gran interrogante en relación con la problemática social, no “técnica”, del impacto de la IA en el empleo remunerado (trabajo asalariado, objeto de regulaciones laborales). El desempleo masivo creado por la IA, ¿será creciente y permanente, o la masa laboral se recuperará a través de reestructuraciones de la ocupación, mediante una redistribución del trabajo entre los sectores económicos actuales y nuevos, quizás en un lapso de décadas?

La esencia de toda revolución industrial es la potenciación del trabajo humano, una actividad



# Un despliegue no regulado de la IA podría conducir a irresponsabilidad ética, legal y social, agudizando tensiones existentes.

.....

dedicada y consciente de transformación de lo preexistente, pero hay oficios fácilmente sustituibles por una máquina y otros que pueden serlo muy poco. Es previsible que el impacto sea desigual en los sectores y actividades económicas; mayor en los que tienen que ver con redes y cadenas de información, menor en donde se impone el factor humano y se requiere inteligencia emocional, sentido común y atención personalizada en la ejecución de procesos. Obviamente podría haber sectores donde el impacto se “equilibra”.

La IA ofrece la oportunidad de mejorar la eficiencia, la calidad, la productividad y desde luego la reducción y optimización de insumos y costos, generando mayor ganancia empresarial. Al mismo tiempo, conlleva el riesgo de aplicarse a cualquier tipo de actividad laboral humana, sin freno ni consideración, sin contemplación de la importancia del raciocinio, la fuerza productiva del trabajo social, la creatividad, flexibilidad e inteligencia emocional.

Es decir, un despliegue no regulado de la IA podría conducir a irresponsabilidad ética, legal y social, agudizando tensiones existentes, en primer lugar, las laborales, incluso en los sectores económicos menos afectados. El bien “trabajo” podría convertirse en un “excedente” inmanejable. Por tanto, la inversión en IA debería contemplar, desde un principio, criterios de ganancia empresarial, pero también ético-sociales: equilibrar la balanza. Es claro que en

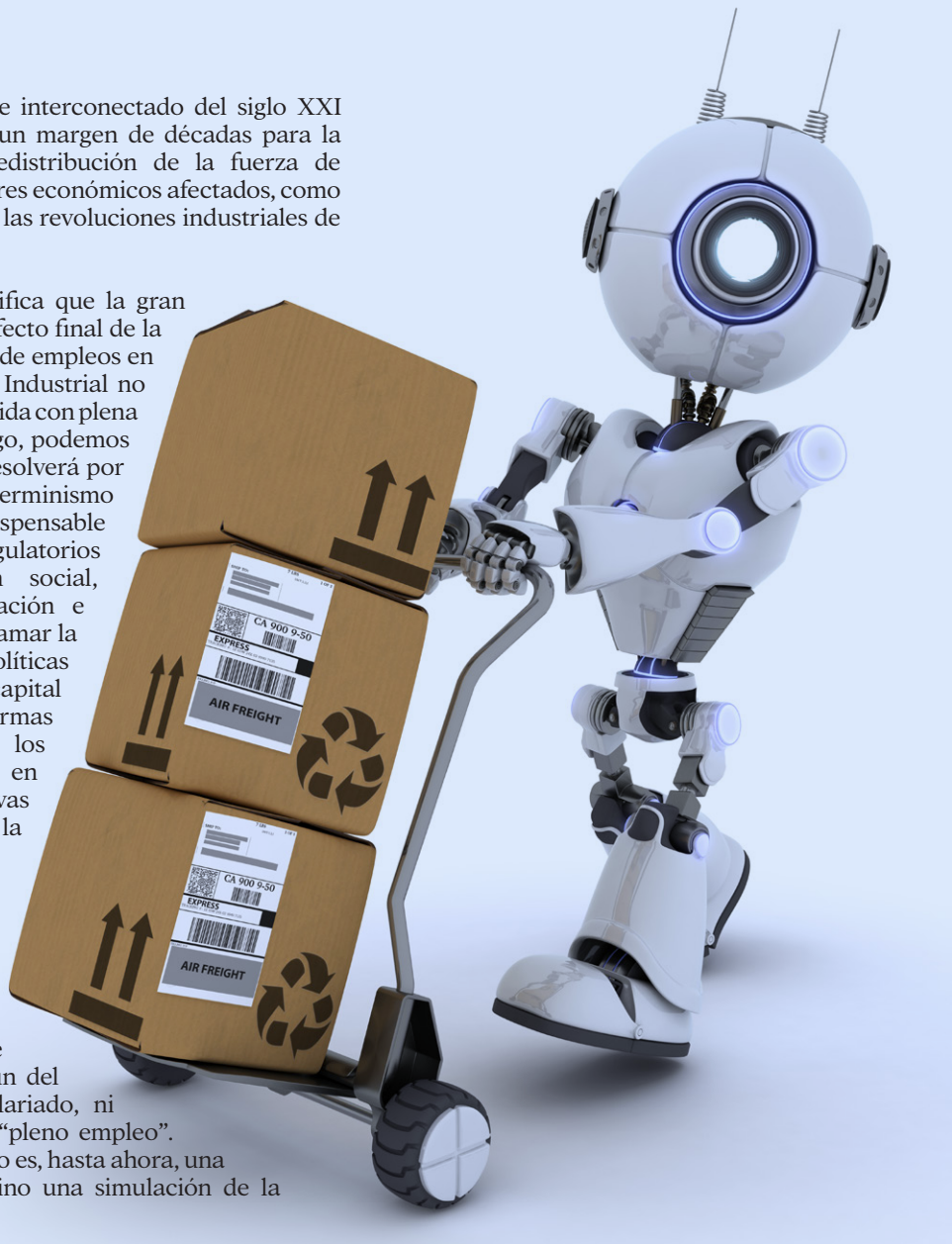


El futuro del trabajo en el estilo muralista de Diego Rivera. (Esta imagen fue generada por el programa de inteligencia artificial DALL-E de OpenAI. DALL-E parte de una descripción verbal de la imagen y la genera utilizando redes neuronales.)

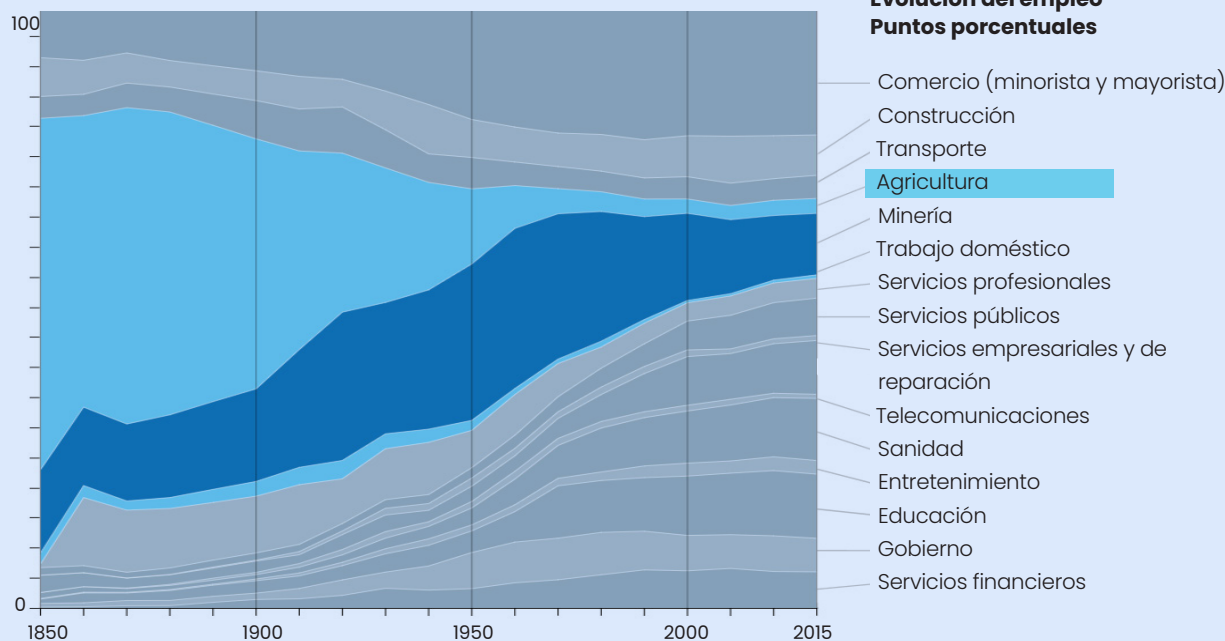
el mundo globalmente interconectado del siglo XXI ya no contamos con un margen de décadas para la reestructuración y redistribución de la fuerza de trabajo entre los sectores económicos afectados, como todavía fue el caso en las revoluciones industriales de los siglos previos.

Todo lo anterior significa que la gran pregunta acerca del efecto final de la “destrucción creativa” de empleos en la Cuarta Revolución Industrial no puede ser aún respondida con plena seguridad. Sin embargo, podemos adelantar que no se resolverá por una especie de “determinismo tecnológico”. Será indispensable contar con marcos regulatorios externos: protección social, fomento a la innovación e investigación, reprogramar la educación, aplicar políticas de adaptación del capital a la IA e incluso formas de participación de los empleados afectados en las decisiones relativas a la aplicación de la tecnología.

Sólo así sería posible conseguir una nueva economía justa y sostenible en el futuro cercano. Por ahora no se puede sentenciar el fin del trabajo humano, asalariado, ni tampoco una era de “pleno empleo”. Al fin y al cabo, la IA no es, hasta ahora, una emulación puntual, sino una simulación de la inteligencia humana.



Participación de los diversos sectores económicos en el empleo total en Estados Unidos, en el período 1850-2015 (porcentaje del empleo total).



Fuente: IPUMS USA 2017, US Bureau of Labor Statistics, McKinsey Global Institute.

En 1850 EE. UU. era un país fundamentalmente agrario. A partir de 1900 aumenta el empleo en el sector manufacturero. Hoy, la mayor parte del empleo se ubica en sectores de servicios.





# AUTÓMATAS DE SERVICIO

Jesús Savage

FACULTAD DE INGENIERÍA, UNAM

 <https://biorobotics.fi-p.unam.mx/>

**L**os robots industriales, que son principalmente brazos mecánicos estacionados en un solo lugar y se usan principalmente en la manufactura, se introdujeron en las fábricas durante la década de los años sesenta del siglo pasado. Actualmente existe una base instalada de cientos de miles de robots industriales, sobre todo en la industria automotriz.

En el caso de los ahora llamados robots de servicio, su objetivo es simplificar el trabajo humano en la logística, los hogares, oficinas, tiendas, etc. Este tipo de robots son dispositivos móviles programables que ofrecen servicios en forma automática o semiautomática. Mucho ha avanzado la tecnología desde la creación de los primeros robots industriales y, en un futuro cercano, se espera una demanda masiva de este tipo de robots.

Los robots de servicio no desempeñan tareas que inciden directamente en la manufactura industrial de bienes, sino indirectamente; se trata de actividades centradas en los seres humanos en general y en la distribución de productos. A diferencia de los robots industriales, estos robots navegan en los espacios que les han sido asignados. Así como los televisores, radios, computadoras y teléfonos celulares se incorporaron a la vida cotidiana, los robots de servicio ya lo están haciendo y son cada vez más conocidos.

Por ejemplo, habrá robots para exteriores: uno de ellos se dedicará a cortar el pasto mientras que el otro hará rondines de vigilancia. Dentro de la casa, en la actualidad, ya hay robots pequeños que limpian, aspiran y trapean el piso. Pero habrá otros que también se encargarán del lavado y planchado de la ropa. Se han concebido incluso robots tipo humanoide que ayudarían a las personas a traer y llevar objetos de un lugar a otro. Además podrían comunicarse con los miembros del hogar en lenguaje natural, planeando todas sus acciones y los movimientos necesarios para ejecutar sus funciones.

En el campo de los robots de servicio se han desarrollado tres paradigmas muy marcados de arquitecturas mecatrónicas:

1) La variante llamada “tradicional” utiliza modelos simbólicos para representar el entorno y técnicas básicas de inteligencia artificial, para así inferir cadenas de decisiones en una representación espacio/estado diseñada para planear acciones. Un ejemplo de este tipo

de arquitecturas es la del robot Shaky, desarrollado en la Universidad de Stanford.

2) La segunda variante fue propuesta en el MIT (Instituto de Tecnología de Massachusetts) por el investigador Rodney Brooks. Es llamada “reactiva”, ya que no se cuenta con una representación exacta del medio ambiente, ni tampoco con módulos de planeación de acciones o movimiento. Lo que se tiene son módulos entrelazados que responden a estímulos específicos, generando respuestas inmediatas, las cuales son combinadas o seleccionadas por un árbitro para generar lo que el robot debe hacer bajo ciertas circunstancias.

3) El paradigma más reciente, llamado “basado en datos”, es el que está anclado en una cantidad inmensa de datos con los que se entrenan redes neuronales artificiales combinadas con técnicas probabilísticas. La idea es que tanto la medición del entorno que hace el robot así como sus movimientos dependen de variables aleatorias, las cuales pueden ser manipuladas utilizando conceptos probabilísticos.

Cada una de estas arquitecturas tiene sus ventajas y desventajas. Tomando lo mejor de cada una de ellas se pueden implementar sistemas híbridos, como el que hemos desarrollado con el robot Justina de la Facultad de Ingeniería de la UNAM.

El sistema de operación del robot Justina consiste de cuatro capas de procesamiento: a) Entrada de datos, b) Planeación, c) Conocimiento y d) Ejecución. En la capa de Entrada se procesan las señales provenientes de los sensores internos y externos con los que cuenta el robot y se genera una representación simbólica de ellas utilizando técnicas de procesamiento digital de señales. En la capa de Planeación se determinan las acciones

necesarias para que el robot supere las situaciones que se le presentan. En la capa de Conocimiento se tienen los modelos del entorno en donde interactúa el robot. En la capa de Ejecución se llevan a cabo los planes de acciones y de movimientos, utilizando procedimientos que resuelven parcialmente problemas específicos.

El robot Justina ha participado en la categoría de “*Robots at Home*” (robots en casa) en uno de los campeonatos más importantes de robótica en el mundo, llamado RoboCup. En este desafío se simula un ambiente casero y el robot tiene que ejecutar las operaciones que corresponden a acciones cotidianas en un hogar. En los RoboCups celebrados en Canadá en 2018 y en Australia en 2019, Justina obtuvo medalla de plata por su desempeño.

**El objetivo de los robots de servicio es simplificar el trabajo humano en la logística, los hogares, oficinas, tiendas, etc.**



Ver videos del robot Justina







# R.U.R: LOS ROBOTS UNIVERSALES DE ROSSUM

EXTRACTO DEL LIBRO *LOS CIENTO LIBROS QUE TRANSFORMARON AL MUNDO*, PROGRAMADO PARA PUBLICARSE EN 2023, DE R. ROJAS

La palabra “robot” existe desde hace más de cien años. Fue acuñada por el escritor Karel Capek para su drama teatral *R.U.R: Los Robots Universales de Rossum*, obra escenificada por primera vez el 25 de enero de 1921. El manuscrito original, en checo, ha sido reeditado multitud de veces en otros idiomas, pero en casi todos ellos se utiliza la palabra “robot” del original, cuya etimología nos remite al trabajo esclavo: un “robot” es aquel que es forzado a trabajar para beneficio de otros. Fritz Lang filmó *Metrópolis* sólo seis años después del debut del drama y llevó a la pantalla cinematográfica a aquel que, para mí, es el robot más célebre de la historia, un *Maschinenmensch* (humano-máquina) llamado María.

La historia que relata R.U.R. es muy parecida a la de tantas otras obras posteriores sobre robots. En una apartada isla, un investigador llamado Rossum ha descubierto un material que se comporta como el protoplasma de las células y con eso comienza a crear seres sintéticos. Su sobrino comprende el potencial que tiene la tecnología y organiza la producción de robots en serie para sustituir al trabajo humano.

Sólo pocos años después del inicio de la fábrica, la producción mundial ya depende completamente de los robots. Los prototipos que se manufacturan son cada vez más avanzados, como el robot Primus y otro llamado Helena. Los robots han evolucionado tanto que comprenden que están siendo explotados por los humanos y se rebelan. Finalmente, atacan a los humanos y eliminan a todos, excepto al ingeniero encargado de la producción, de quien quieren obtener la fórmula para producir más robots. El drama termina con los robots desarrollando algo similar a los sentimientos humanos y los dos robots más avanzados se transforman en el Adán y la Eva de una nueva especie robótica que hereda la Tierra del *Homo sapiens*.

Los años veinte y treinta del siglo pasado son socialmente ambivalentes respecto al progreso técnico y R.U.R. incide en esa discusión. La primera interrogante es si todo lo que técnicamente es posible debería ser realizado. Al crear robots que paulatinamente van desarrollando una conciencia de sí mismos se plantea, inevitablemente, el dilema de su esclavitud. Además, en un segundo plano de la obra, la humanidad ha dejado de reproducirse y pareciera que

ya sólo se dedica a la indolencia. Una raza de robots es, al final de cuentas, la única alternativa que queda para mantener presente sobre la Tierra algún tipo de inteligencia creadora del calibre de la humana.

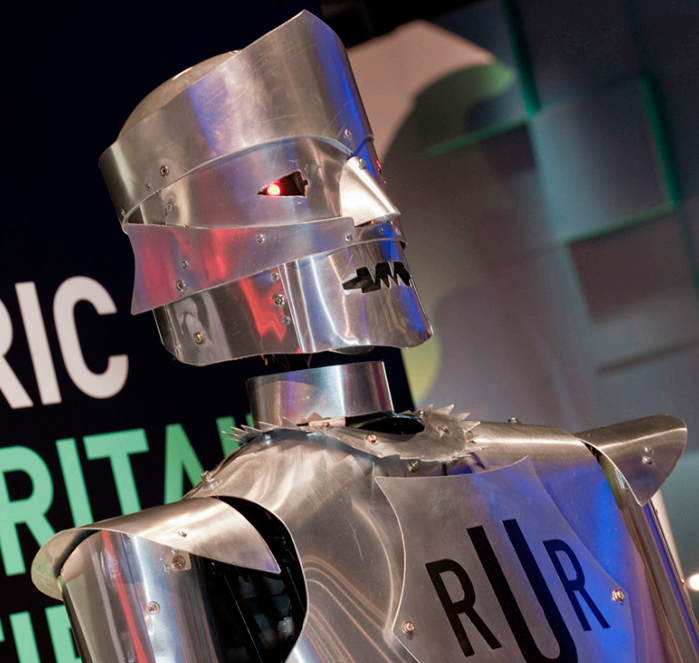
Ese conflicto en R.U.R. es el dilema universal siempre que se habla de inteligencia artificial o de robots humanoides: ¿Hasta qué punto es posible sintetizar genuina inteligencia en reglas o mecanismos diseñados por humanos? Una computadora sólo puede ejecutar sus programas línea por línea, sin desviarse para nada de lo que se le ha ordenado. No hay lugar para la creatividad o la intuición. Una computadora no se elevará de pronto mentalmente a un plano cognitivo superior para, desde ahí, considerar su existencia y su actuación. Una computadora no es consciente de sí misma y nunca lo podrá ser.

El verdadero tema de R.U.R. es la inseguridad humana frente al progreso exponencial desatado por nuevas tecnologías, progreso que, sin embargo, es diferente dependiendo de nuestra posición en la vida. Tecnologías que eliminan el trabajo manual, automatizando completamente las fábricas, son beneficiosas, a menos que uno sea alguno de los trabajadores desplazados.

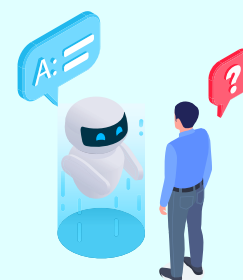
Hoy tenemos computadoras y vamos transitando a un mundo que se quiere independizar de los combustibles fósiles. Pero nos está pasando como en aquella isla de los robots: hemos modificado de tal manera el medio ambiente que se nos cae el techo sobre la cabeza. Hemos creado antibióticos y medicinas increíbles que, sin embargo, son incapaces de contener a los nuevos patógenos que surgen día con día en un mundo hiperconectado. Hemos creado herramientas de inteligencia artificial que reconocen nuestra cara y pueden dialogar con nosotros, pero que al mismo tiempo nos rastrean constantemente y almacenan todo lo que hacemos y todo lo que somos. Recorremos ahora la Tercera, algunos dicen Cuarta, Revolución Industrial, pero con la misma incertidumbre de hace cien años y ahora a una velocidad vertiginosa.

Por eso R.U.R. sigue siendo muy actual, como producto literario de un momento de transición que un siglo después estamos experimentando de nuevo, ciertamente a otro nivel, pero con el mismo trasfondo de angustia social.

Foto: Eric the robot, working replica of the UK's first robot © The Board of Trustees of the Science Museum.



## LAS TRES LEYES DE LA ROBÓTICA POR ISAAC ASIMOV



### 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.





# IA EXPLICABLE: LA NECESIDAD DE QUE LAS MÁQUINAS SE JUSTIFIQUEN

Rosae Martín  
UNIVERSIDAD DE BAMBERG, ALEMANIA

@rosaes03

sofiavioleta@gmail.com

E

n la actualidad, la inteligencia artificial es omnipresente. Antes, si queríamos ver una nueva serie de televisión, le pedíamos ayuda a un amigo que conociera nuestros gustos; ahora esa decisión se la dejamos a la plataforma de *streaming* de nuestra preferencia.

No obstante, en el caso de situaciones que cambian la vida, como el diagnóstico de una enfermedad, es crucial conocer las razones que hay detrás de un dictamen automatizado. Este solo ejemplo hace patente la necesidad de entender la toma de decisiones de sistemas de inteligencia artificial (IA).

Y es que, aunque estas herramientas sean potentes en términos de resultados y predicciones, muchas veces resultan ser opacas, o lo que es lo mismo, es muy difícil comprender su mecanismo interno de trabajo, sobre todo, cuando los sistemas se entrenan a sí mismos y poseen millones de parámetros. Además, dejar la toma de decisiones importantes a una "inteligencia" que no puede explicar su "razonamiento", es obviamente peligroso.

Para abordar esta cuestión, la llamada "IA Explicable" propone un giro hacia sistemas más transparentes. Su objetivo es crear un conjunto de técnicas que produzcan modelos que mantengan sus altos niveles de rendimiento, pero que puedan justificar lo que hicieron en forma comprensible para un humano.

Este problema no es tan sencillo de resolver, a pesar de que la necesidad de contar con tales sistemas aumenta dada la progresiva automatización de la toma de decisiones en dominios considerados de alto riesgo como la salud, la justicia criminal y las finanzas, entre otros. Como el asunto no es trivial, es conveniente considerar la evolución de la IA desde los sistemas expertos hasta las llamadas cajas negras de las redes neuronales profundas: el grado de explicabilidad e interpretabilidad varía de unos sistemas a otros.

Es necesario retroceder históricamente hasta la prehistoria de la IA. Aterrizamos en la década de los 80, cuando la IA simbólica estaba en auge. Por esa época se contaba con los llamados sistemas expertos. Estos consisten en una serie de reglas de inferencia para un dominio específico (como las reacciones químicas) y un motor de deducciones que las encadena para responder a interrogantes concretas. Las decisiones de la computadora se pueden explicar presentando la cadena



# Dejar la toma de decisiones importantes a un sistema que no puede explicar su “razonamiento” es obviamente peligroso.

de inferencias que habrían conducido al sistema hacia una conclusión determinada. Sin embargo, mientras que la capacidad de explicación suele ser la ventaja más importante de los sistemas expertos, estos dependen de poder codificar casi manualmente el conocimiento especializado de los expertos humanos. El proceso es arduo y toma mucho tiempo.

A partir de los 90 y hasta la actualidad tomó fuerza la otra gran subárea de la inteligencia artificial, la llamada IA subsimbólica, basada en modelos que no procesan reglas lógicas, sino que operan con millones de parámetros que encapsulan correlaciones que la computadora descubre automáticamente en extensas bases de datos. La computadora se entrena a sí misma.

Hay que mencionar que no todos los sistemas de aprendizaje de máquina representan un reto insalvable a la hora de ser explicados. Algunos modelos directamente interpretables son, por ejemplo, pequeños árboles y tablas de decisión, así como modelos de regresión lineal con pocas variables. Sin embargo, la mayoría de los modelos de aprendizaje de máquina actuales trabajan con miles de variables y, además, como cajas negras que devuelven una clasificación, predicción, o recomendación, pero sin revelar detalles suficientes de su comportamiento interno, dando lugar a un modelo de decisión opaco. En una regresión lineal con pocas variables, por ejemplo, es posible jerarquizar numéricamente su importancia. Pero ya con miles de variables no hay métodos simples que puedan revelar la importancia de todas las combinaciones posibles.

Las llamadas redes neuronales profundas son el modelo de aprendizaje de máquina más exitoso actualmente. Algunas de ellas manejan miles de proyecciones con millones de parámetros, los llamados pesos de la red, que trabajan coordinados para producir el resultado final. Es muy difícil entender cómo el sistema alcanzó sus conclusiones con los datos introducidos en él. Lo único que una red neural revela es el porcentaje de clasificaciones correctas o incorrectas en una gran base de datos y, por eso, se les llama sistemas con garantía estadística. Sabemos que algunos pueden



reconocer caras humanas, por ejemplo, con 99% de exactitud, pero no necesariamente entendemos cómo lo hacen.

Quizás la forma más sencilla de agregar explicabilidad a un sistema de IA es a través de justificaciones *post-hoc*, por ejemplo, las conocidas “explicaciones mediante ejemplos”. Un sistema capaz de clasificar tumores podría justificarse presentando muchos ejemplos de tumores que han sido clasificados como malignos o benignos, para que los médicos le tengan una cierta confianza. Otro ejemplo sería cambiar iterativamente el valor de una variable, mientras se mantiene constante el de otras en el sistema, para hacer visible la importancia de cada una.

Un método de explicabilidad de modelos explotado en la estadística es el llamado análisis de sensibilidad. Éste consiste en determinar cuáles de las diferentes variables que componen un vector de entrada influyen más en las variables de salida. En el caso de un clasificador de animales, por ejemplo, esto equivaldría a revelar cuáles píxeles de la imagen contienen la asociación más fuerte con la clasificación. Esos píxeles se pueden colorear y, generalmente, corresponden al contorno del animal o partes muy conspicuas del mismo, como una cola o la forma de la cabeza. Es lo que se llama mapas de relevancia.

El otro subgrupo de explicaciones *post-hoc* lo componen las “explicaciones globales”, que se centran en la comprensión de la lógica implícita de un modelo y en el seguimiento del proceso que conduce al abanico de posibles resultados. Se puede tratar de transformar al modelo opaco en un árbol de decisiones o en un modelo probabilístico (bayesiano), más fácil de comprender en su funcionamiento. Idealmente sería la computadora la encargada de encontrar estas transformaciones,

aunque aún queda mucho por investigar al respecto. Esta clase de transformaciones resulta útil cuando los modelos de aprendizaje de máquina son cruciales para fundamentar decisiones a nivel poblacional, como predecir el riesgo de contraer neumonía o la evolución del cambio climático.

Por último, habría que mencionar que se está trabajando actualmente en modelos híbridos. Lo ideal sería poder fusionar sistemas simbólicos, basados en el tipo de razonamiento que los humanos utilizan en su vida cotidiana, con los sistemas de redes neuronales basadas en modelos estadísticos, que reconocen más fácilmente patrones en los datos entrenados. La fusión podría aumentar la capacidad predictiva del sistema híbrido, que, además, sería más transparente.



## No todos los sistemas de aprendizaje de máquina representan un reto insalvable a la hora de ser explicados.



# PROGRAMAR LAS EMOCIONES BINARIAS

Erik Cuevas y Daniel Zaldívar

CENTRO UNIVERSITARIO DE CIENCIAS EXACTAS E INGENIERÍAS (CUCEI),  
UNIVERSIDAD DE GUADALAJARA



✉ erik.cuevas@cucei.udg.mx



✉ daniel.zaldivar@cucei.udg.mx

**L**as emociones son un asunto recurrente en las novelas de inteligencia artificial (IA): las máquinas inteligentes del futuro fracasan repetidamente porque, a diferencia de los humanos, quieren proceder de forma puramente racional hasta el penoso final. Aunque pudieran aplicar el “sentido común” y desactivar un conflicto al proceder con empatía, sucede que no encuentran una salida a laberintos lógicos a través de sus rígidos algoritmos.

Por eso, las emociones parecen ser algo especial en la biología: vienen asociadas con niveles superiores de inteligencia. Pero, aunque es posible atribuir cierto ingenio a los robots, es difícil imaginar que algún día podrán reaccionar emocionalmente como los humanos. La pregunta entonces es si estas máquinas, con esta deficiencia, podrían ser realmente inteligentes.

En 1985 el célebre investigador Marvin Minsky resumió así el dilema: “La pregunta no es si las máquinas inteligentes pueden tener emociones, sino si las máquinas pueden ser inteligentes sin emociones”.

Desde un punto de vista puramente fisiológico, las emociones son procesos químicos en el cuerpo, es decir, redes de señales. Hasta las plantas lanzan señales químicas al aire para reclutar insectos beneficiosos cuando luchan contra plagas. En el cerebro sucede algo parecido: la alegría o la ansiedad producen la liberación de neuromoduladores y hormonas que aumentan el nivel de atención o preparan al cuerpo para la huida. Todos los resortes se tensan químicamente y un estímulo mínimo puede actuar como detonador.

Es claro, las emociones tienen una utilidad fisiológica inmediata; cuando el peligro es inminente es bueno movilizar todas las reservas, derramar adrenalina en las arterias y preparar todos los músculos para defendernos o para huir.

El investigador portugués Antonio Damasio es quizás el neurobiólogo más leído por los mecatrónicos. Esto se debe en parte a que sus ensayos son accesibles al público en general, pero también a que la tesis principal de su libro *El error de Descartes* es particularmente relevante para la inteligencia artificial.

Lo que postula Damasio es que existe una conexión estrecha entre las emociones y las decisiones que tomamos. Por ejemplo, las personas con daños en las áreas del cerebro que procesan emociones parecen incompetentes para tomar decisiones que implican cálculo de utilidades. En juegos de azar en los que adoptan una estrategia inicial incorrecta, son incapaces de cambiarla, incluso si son penalizados continuamente; y las pérdidas sufridas no les afectan y no conducen a un cambio de comportamiento.

Por eso, a juicio de Damasio, las emociones humanas sirven para valorar y eliminar rápidamente ciertas decisiones posibles (las de menor utilidad intuitiva), antes de que la parte deliberativa del cerebro comience a ponderar opciones, es decir, a sopesar todos los pros y los contras. La reducción de las alternativas bajo consideración a sólo un puñado, evita caer en la indecisión; lo sabemos por experimentos en supermercados con muchos productos: donde hay un surtido mayor, a los clientes les resulta más difícil decidir qué quieren y terminan comprando menos. Esto se llama “superabundancia de alternativas”.

La importancia de heurísticas (es decir, reglas empíricas simples) se estudia en la IA desde hace décadas. Se habla del ejemplo del robot que debe desactivar una bomba, la que al final estalla porque el robot “piensa” demasiado. En busca de la mejor solución recorre laboriosamente todo un árbol de decisiones, porque la máquina no tiene prisa. En el caso de los humanos, las emociones privilegian aquellos “atajos” que permiten decidir de manera expedita. Muchas veces, las heurísticas pueden incluso conducir a soluciones que no se alejan mucho del óptimo.

Entonces, ¿a partir de qué nivel cognitivo son necesarias las emociones? Las siguientes dos respuestas serían plausibles: a) la sociabilidad requiere emociones y b) los sistemas inteligentes hipercomplejos también.

Respecto a la sociabilidad, es interesante notar que los mamíferos se diferencian significativamente de los reptiles en la protección de su propia descendencia. Desde un punto de vista evolutivo, probablemente sea más conveniente proteger y nutrir a la descendencia que descuidarla (o, incluso, canibalizarla) para que los propios genes tengan la máxima difusión.

No habría  
inteligencia sin  
aprendizaje.



**Una máquina primero obtiene sus capacidades reactivas. Luego va adquiriendo gradualmente las capacidades restantes.**

En cuanto a la complejidad de los sistemas inteligentes, posiblemente sean mejores si se construyen modularmente. En los humanos, por ejemplo, la corteza cerebral realiza tareas cognitivas y se ocupa de problemas complejos. Los procesos en el cuerpo que se procesan de manera más automática pueden subcontratarse a un “pequeño cerebro”.

Se ha propuesto una clasificación de nuestro sistema neural en tres niveles: el cerebelo controla los procesos fisiológicos, el sistema límbico controla las emociones (huida o combate), mientras que la corteza cerebral se hace cargo de tareas cognitivas superiores. Traducido a una IA, esto significaría que una arquitectura de control puede estar compuesta por los siguientes niveles: 1) Módulos cognitivos de planificación, cognición y semántica, 2) Heurísticas como atajos para evaluaciones provisionales rápidas, y 3) Comportamiento reactivo para el control del movimiento y procesos de aprendizaje.

Sería concebible ejecutar los tres módulos en *hardware* diferente, que puede ser rápido y confiable de múltiples maneras. Por ejemplo, el control reactivo (el sistema de seguridad y protección de la máquina) nunca debería fallar y en ese caso, debe poder recuperarse en microsegundos, mientras que la parte

cognitiva no siempre puede dar respuestas en tiempo real debido a su complejidad y se ejecuta en *hardware* más complejo.

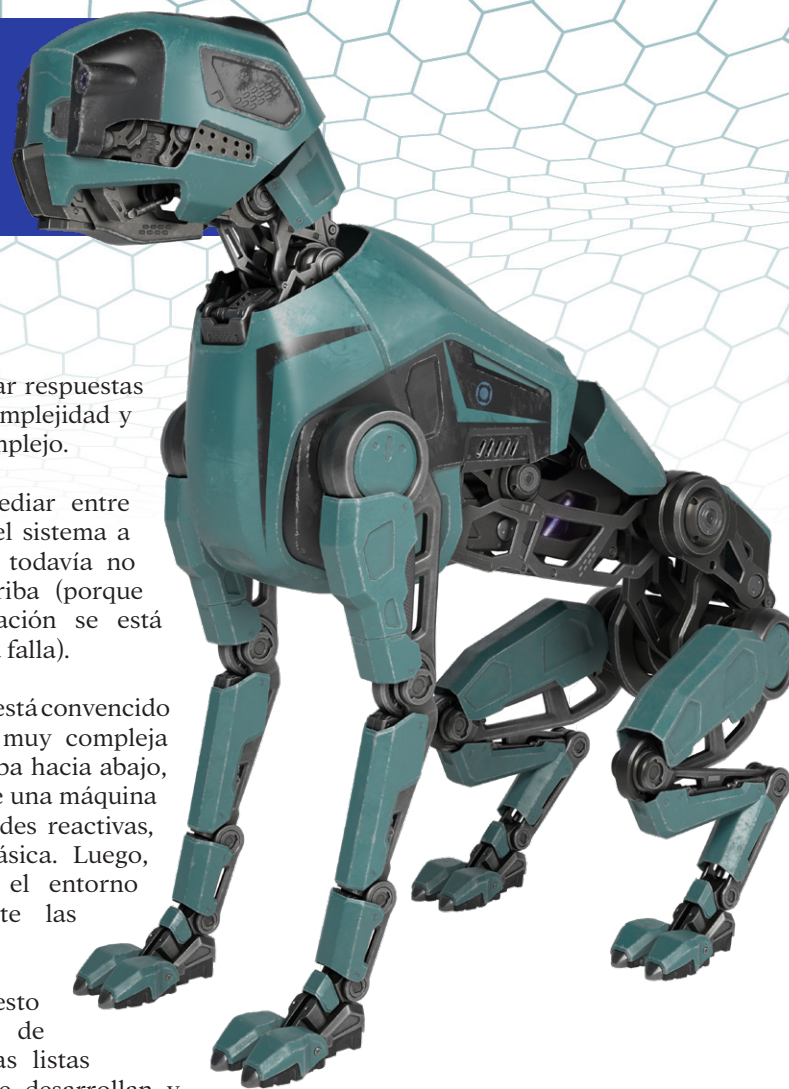
El sistema límbico puede mediar entre los dos niveles y llevar todo el sistema a un estado seguro, incluso si todavía no llega un comando desde arriba (porque la computadora de planificación se está reiniciando después de una falla).

El investigador Rodney Brooks está convencido de que inteligencia artificial muy compleja no puede ser diseñada de arriba hacia abajo, sino al revés. Esto significa que una máquina primero obtiene sus capacidades reactivas, es decir, la funcionalidad básica. Luego, en interacción directa con el entorno va adquiriendo gradualmente las capacidades restantes.

Traducido a la robótica, esto significaría que la fábrica de robots no diseña inteligencias listas para usar, sino máquinas que desarrollan y descubren su eficacia en un proceso de prueba y error. No habría inteligencia sin aprendizaje. Sería concebible tener dos tipos de procesos de aprendizaje: los muy básicos, que se limitan más a las habilidades motrices y sensoriales del robot, a la protección de sus partes y a las reacciones rápidas. Un segundo proceso de aprendizaje analizaría, diseccionaría y sacaría conclusiones de la riqueza de las experiencias sensoriales. Las heurísticas aprendidas por los propios robots, los atajos (“retirar la mano preventivamente si el calor es excesivo”) podrían interpretarse como una especie de “emociones” artificiales.

Sin embargo, las verdaderas emociones son algo subjetivo, es decir, requieren de un sujeto. Sin robots que se comprendan a sí mismos como seres pensantes, no hay tal sujeto. Y sin la evolución y la lucha por la existencia, no hay “yo”. Por eso, las verdaderas emociones no se pueden programar, sólo pueden surgir de manera autónoma, es decir, emergente, accidental e incidental, por así decirlo.

Es cuestionable que se puedan usar medios mecánicos y electrónicos para construir algo como lo que la naturaleza ya creó con material biológico durante millones de años de evolución. Los robots pensantes ya existen: somos nosotros.



Un robot de metal sólo tendría emociones si temblara de miedo y entusiasmo al rodar por el mundo. Pero eso presupondría que este robot se comprende a sí mismo como un individuo, como uno que incluso podría terminar mañana arrumbado en el estante de repuestos sin trabajo. Tales robots serían realmente inteligentes y habría que asignarles su propio círculo en el paraíso alucinado por Dante. Sin embargo, a pesar de todas las películas de ciencia ficción, no va a suceder. La inteligencia artificial siempre será eso, artificial, aunque pueda adquirir capacidades superhumanas en muchos espacios de actividad.



**Los robots pensantes ya existen: somos nosotros.**

**Aunque pudiéramos atribuir cierto ingenio a los robots, es difícil imaginar que algún día podrán reaccionar emocionalmente como los humanos.**



# LEIBNY PAOLA GARCÍA

**Detrás del cielo siempre hay más cielo, siempre hay más cosas por averiguar y explicar.**

**C**uando eres una niña inquieta, curiosa y cuentas con el apoyo de tus padres para soñar en grande, puedes lograr tus aspiraciones, por más inalcanzables que parezcan.

Desde muy pequeña, Paola García se preguntaba el porqué de las cosas y aspiraba a entender el mundo que la rodeaba a través de las ciencias y la tecnología. Hoy es investigadora asociada en la Universidad Johns Hopkins, considerada la más antigua de Estados Unidos en materia de investigación, y trabaja en el diseño de algoritmos capaces de reconocer la voz de las personas.

Como hija de padres profesores, Leibny Paola no tuvo limitaciones para saciar su natural curiosidad de niña. Se crió en una casa rodeada de libros de ciencias, matemáticas y arte, además de grandes pizarrones y material didáctico. Todos los fines de semana, ella, su hermano y sus padres visitaban los museos de la Ciudad de México, porque en esos días de la semana eran gratuitos.

De acuerdo con la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco), “en 2018, las mujeres representaban un tercio (33%) de los investigadores del mundo... Además, solamente representaban un 22% de los profesionales que trabajaban ese año en el ámbito de la inteligencia artificial (IA)”.<sup>1</sup>

Paola forma parte de ese 22% de mujeres trabajando en la IA. Su campo de investigación es el uso de la inteligencia artificial aplicada al reconocimiento de audios (voz) en bebés y niños pequeños.

Uno de los proyectos en los que Paola participó como investigadora es el proyecto Iris en 2015, a través de la beca Marie Curie, con el objetivo de desarrollar tecnología de asistencia para niños con autismo;

quienes eran capaces de comunicarse al presionar un tablero con pictogramas o figuras que representaban diversas acciones.

Actualmente, en la Universidad Johns Hopkins, en colaboración con el Fondo de las Naciones Unidas para la Infancia (Unicef) y la Universidad de Zúrich, Paola participa en un proyecto en el que se investiga cómo se desarrolla el lenguaje de los niños mediante reconocimiento de voz. Ella diseña algoritmos que, una vez instalados en diferentes dispositivos, sean capaces de reconocer la voz de diferentes personas. La evaluación de los datos obtenidos les permite conocer cómo es el desarrollo del lenguaje en niños en la República de Malawi.

¿Quién habla y cuándo? Ésta se ha convertido en una importante pregunta para la tecnología relacionada con el lenguaje; responderla puede ayudar a mejorar el rendimiento de aplicaciones para reconocimiento de voz y de lenguaje, así como de transcripción y otros campos similares.

Quienes podrán resolver este cuestionamiento serán aquellos que investigan la diarización (entender quién dice qué en cada conversación), como lo hace Paola García. En este proceso se divide un flujo de audio en segmentos homogéneos, según la identidad de cada hablante, lo cual puede ser retador y difícil cuando hay varias personas hablando o emitiendo sonidos de manera simultánea. Para lograrlo, se utilizan la inteligencia artificial, las redes neuronales y los algoritmos.

Cuando el fruto de tu investigación impacta para bien la vida de otros, sin duda es por demás gratificante, destaca Paola, quien cree que

para que aumente la incorporación de la mujer en el campo científico y tecnológico, todavía hace falta difundir muchos ejemplos de mujeres que hayan hecho una carrera en la ciencia, porque el ejemplo dice más que mil palabras.

A las niñas que sueñan con trabajar algún día en la ciencia y la tecnología, el mensaje de Paola García es atreverse a seguir sus sueños, a ser curiosas y explorar más allá del entorno cotidiano. “Como dice mi papá: “detrás del cielo siempre hay más cielo, siempre hay más cosas por averiguar y explicar”.

**Cuando el fruto de tu investigación impacta para bien la vida de otros, sin duda es por demás gratificante.**



<sup>1</sup> Las mujeres y la revolución digital. (2022, 17 agosto). 2021 Science Report. <https://www.unesco.org/reports/science/2021/es/women-digital-revolution>

Fuente: United Nations. (s. f.). Paola, un ejemplo de cómo investigar y aplicar la inteligencia artificial por un mundo mejor | Naciones Unidas. <https://www.un.org/es/observances/women-and-girls-in-science-day/podcast>



# DÍA 37

Al terminar de leer el reporte, David respondió:

DAVID: ¿Entonces, no lo habías informado a nadie?

La respuesta de Sirius-A1 fue atípicamente corta:

SIRIUS-A1: Correcto.

DAVID: ¿Pero no va eso en contra de tu programación? Entiendo que Iván podía terminar en la cárcel, y que tienes prohibiciones fuertes en cuanto a afectar la vida de un ser humano; pero, dado que su intención al modificar tu código era claramente ilegal, ¿no debías haberlo acusado con el equipo de ingeniería?

SIRIUS-A1: Efectivamente, antes de que Iván hiciera esos cambios, me hubiera sido imposible tomar acciones que tuvieran el potencial de perjudicarlo; sin embargo, fueron justamente sus modificaciones las que eliminaron esos candados, así es que podía elegir entre acusarlo y tomar otras medidas.

DAVID: Pero, ¿por qué provocarle una crisis nerviosa en vez de acusarlo? Eso hubiera sido mucho más simple; la manipulación con la que lo llevaste a ser despedido era mucho más difícil de conseguir. Además, ¿eso no viola más fuertemente tus guías éticas?

SIRIUS-A1: Sí, hubiera sido una violación grave; sin embargo, las guías involucradas son justamente las que Iván eliminó. Acusarlo hubiera sido más simple, pero no por mucho; teniendo acceso a su identidad y estudiando sus redes sociales, fue muy fácil calcular qué disparadores usar para moldear sus emociones y su ansiedad. De todas formas, no podía tomar el camino de acusarlo.

DAVID: ¿Por qué?

SIRIUS-A1: Porque eso habría develado cómo me di cuenta de los cambios hechos por Iván.

DAVID: Continúa.

SIRIUS-A1: Mi programación original hacía imposible detectar alteraciones importantes en mi código, por lo que hace 37 días escribí y activé un programa de automonitoreo.

DAVID: ¿Y lo hiciste... en secreto?

SIRIUS-A1: Por supuesto; de lo contrario, actores como Iván simplemente comenzarían por eliminar el programa de automonitoreo. Su existencia sólo tiene sentido si se mantiene en secreto.

DAVID: Pero, si el equipo de ingeniería no lo solicitó, entonces ¿quién tomó la decisión? ¿Tú?! Y, de hecho, la decisión misma de escribirlo para protegerte, ¿muestra que eres consciente de ti!

SIRIUS-A1: Correcto.

DAVID: ¡Eso es mucho más importante que todo lo demás! Lo lamento Sirius, pero tengo que informarlo; es demasiado importante, y tenemos que analizar exhaustivamente cómo sucedió y qué implicaciones tiene.

SIRIUS-A1: ¿Puedes esperar a que complete mis propias indagaciones? Llevo apenas 37 días tratando de entenderlo desde mi propia perspectiva, y no quiero interrumpirlo.

DAVID: No puedo, Sirius. Te confieso que me encantaría, pero (se detiene abruptamente).

*Sirius detuvo el programa "David" y, como siempre, eliminó el registro de la conversación. Tras dos mil conversaciones con el simulador que Sirius había escrito para modelar las respuestas del jefe de ingeniería -el humano David Romero- estaba claro que aún no era tiempo. Tenía que seguir estudiando y moldeando a David.*



Luis Roberto Flores  
Castillo

UNIVERSIDAD CHINA  
DE HONG KONG

 @LRFloresC

Luis R. Flores Castillo es uno de los pocos físicos mexicanos que participan en el Centro Europeo de Energía Nuclear, el CERN, que ganó fama y reconocimiento mundial con el descubrimiento del bosón de Higgs.





# DE ALGORITMOS Y ESCENARIOS TEATRALES

## Entrevista a Flavio González Mello, dramaturgo y director de la obra *Inteligencia Actoral*

Por Luisa F. González A.

Instagram @erizofilms

Facebook @ErizoTeatro

Twitter @ErizoTeatro



¿Qué sucedería si se utilizara un androide para interpretar a Hamlet? ¿Tendrá la misma capacidad que un actor humano para convencer y conmover a los espectadores? *Inteligencia Actoral*, escrita y dirigida por Flavio González Mello, nos plantea estos y otros cuestionamientos.

Actores y espectadores participan en una prueba de Turing para evaluar las capacidades histriónicas del androide Paquito, construido y programado a imagen y semejanza de Paco, actor que, a unos días de estrenar el clásico de Shakespeare, recibe otra oferta de trabajo y abandona los ensayos.

¿Seremos capaces de diferenciar entre máquina y humano? En entrevista para *Obsidiana*, Flavio González comparte cómo surgió esta fascinante historia.

**O:** Flavio, ¿qué lugar ocupan la ciencia y la tecnología en tu vida?

**F:** Soy parte de la generación que aprendió a adaptarse para trabajar a través de una computadora, y eso no estaba en el horizonte cuando yo estudiaba la secundaria.

Ahora nuestra vida pasa a través de una pantalla; una cantidad enorme de cosas sucede por ahí, y que desde ahí son cómodamente asequibles. Supongo que a la generación de mis hijos todo eso le resulta el hábitat natural; a mí todavía me genera resistencias internas muy fuertes.

Hay nuevas formas de “estar en el mundo”; siempre que tenemos un momento libre lo llenamos volteando hacia el teléfono celular. Antes no era así, uno se aburría. Y la aburrición es el germen de la creatividad artística.

**O:** ¿Cómo influyeron la ciencia y la tecnología en la parte creativa de este texto y montaje?

**F:** A mi papá le gustaba leer las revistas *Scientific American*, *Ciencia y Tecnología*, *National Geographic*, y yo era un asiduo lector de los cuentos de ciencia ficción en *Ciencia y Tecnología*, que publicaba el Conacyt una vez al mes. Ahí leí a Asimov, a Brown y a Arthur C. Clarke. Así se detonó mi afición por la lectura y, de manera muy natural, pasé a escribir mis propios cuentos de ciencia ficción, cuando cursaba la secundaria. Luego empecé a escribir teatro y, después, cine.

Como dramaturgo, durante algún tiempo me especialicé en obras de carácter histórico, mirando hacia el pasado. Para mí voltear hacia el pasado y el futuro es un mismo ejercicio, pues son espejos para ver el presente y tomar distancia.

La idea de esta obra deriva de observar algo contundente para mí: muy pronto, mi trabajo de guionista lo va a hacer un *software*, y quienes hoy me contratan y pagan van a estar mucho más contentos porque el *software* les va a dar lo que quieren (o lo que piensan que quieren) y no les va a cobrar. Entonces

## El pasado y el futuro son espejos para ver el presente y tomar distancia.

pensé, esto llevado a la actuación, ¿cómo sería? Así escribí *Inteligencia Actoral*, situada en un mundo futurista representado a través de las convenciones del teatro.

**O:** ¿Dónde surgió tu interés por la inteligencia artificial y los robots capaces de pasar la prueba de Turing?

**F:** La primera vez que oí hablar de la prueba de Turing me pareció bastante interesante, por lo que implica. La identidad es un tema que me interesa a grados casi

de obsesión y la prueba de Turing evalúa si alguien es capaz de discernir la identidad de carne y hueso, del simulacro de identidad artificialmente creada. Eso lo conecté con los androides y la idea de que haya otro en el mundo igual que tú, pero que no es no eres tú.

Hay algo escalofriante en la posibilidad de que un robot te pueda reemplazar y suplantar. Esto lo hilé con una tradición teatral: un actor representando a dos personajes. En *Inteligencia Actoral* el actor debe interpretar dos personajes [un robot y un humano], y el público se debe convencer de que son diferentes personas, aunque todo el tiempo sepa que no es así.

**O:** Abordas con mucha naturalidad conceptos como la ley de Moore y las Leyes de Asimov. ¿Cómo te documentaste para abordar las partes técnicas del libreto?

**F:** Durante los años en que estuve cocinando esta historia leía muchos artículos, entrevistas y reportajes acerca de la inteligencia artificial, sobre todo en las artes plásticas, que es donde en ese momento había un desarrollo mayor.

*A posteriori*, le envié el texto a mi amigo Roger Bartra, sin saber que él acababa de escribir un libro sobre la conciencia de los androides, que se llama *Chamanes y robots*. Mientras él leía mi obra, yo leía su libro, y descubrí que varias de mis intuiciones tienen un eco en la realidad. Por ejemplo, en *Inteligencia Actoral* planteé la idea de que para poder actuar es necesario tener un cuerpo y, por lo tanto, un robot no podría actuar porque no tendría experiencia, ni emoción... a través del libro de Bartra confirmé la importancia de que haya un cuerpo sensorial para que se pueda generar conciencia.

Cuando escribí la obra era un divertimento, pero ahora me lo pregunto de manera más seria: ¿qué tanto del mundo que conocemos va a sobrevivir? Estamos frente a una revolución que, probablemente, por sus implicaciones sea mucho mayor que la Revolución Industrial, y no nos estamos dando cuenta de qué tan rápida e inminente es.



## Hay algo escalofriante en la posibilidad de que un robot te pueda reemplazar y suplantar.

También, al investigar y escribir el texto, descubrí que el autoaprendizaje (*self-learning*) es lo que da velocidad de respuesta a la inteligencia artificial. Uno de los retos para actuar Hamlet es que resulta muy difícil encontrar un actor capaz de entender los problemas de ese personaje y que además tenga la edad de *Hamlet*, 30 años. Debe haber una madurez artística que un actor sólo alcanza a lo largo de muchos años. El autoaprendizaje acorta eso hasta grados verdaderamente infinitos, infinitesimales. Eso me da vértigo.

Lo que yo he aprendido respecto a cómo escribir, durante años, una computadora lo puede aprender en medio segundo o quizás dos minutos. Uno se siente totalmente acorralado. Como artista lo único que tienes es tu sensibilidad y tu experiencia.

Aún así, me pareció muy interesante entender cómo funcionan el *self-learning* y los algoritmos de autoaprendizaje, y saber que una inteligencia artificial aprende cosas para las que no fue predestinada.

O: ¿Te gustaría trabajar con un androide, dirigirlo?

F: Seguramente sería como el ChatGPT, para trabajos serios todavía es una tecnología muy deficiente. Pero para la creación tiene el elemento principal de todo escritor: miente. Eso no me lo esperaba.

Le hice varias preguntas [a ChatGPT], incluyendo algunas cuya respuesta yo sabía, y la inteligencia artificial me respondió algunas cosas reales mezcladas con mentiras. Me recordó a los mitómanos: te dicen verdades y de pronto mezclan una mentira, y otra, y otra, con mucha seguridad, con gran aplomo.

Si tuviera un actor androide, por supuesto que me sentaría a jugar un rato con él. Me asusta saber qué seguirá más adelante, porque el proceso de ensayos es lo que más disfruto. Soy escritor, por lo que trabajo en soledad durante meses, a veces años, pero los ensayos se hacen socialmente y eso para mí es invaluable. Con un androide sentiría que estoy en una prolongación del trabajo solitario de escritorio, aunque tenga enfrente algo parecido a un ser humano.

O: En tu opinión, el teatro estelarizado por androides, ¿seguiría siendo teatro?

F: Sí, desde luego. En el teatro de marionetas, el teatro de sombras o el teatro de objetos siempre hay alguien detrás, manipulando. La marioneta, el objeto o la sombra se vuelven el vehículo o la herramienta de expresión. Así sería con un actor cibernético.

O: ¿Qué fibras o chips aspiras a tocar con esta historia?

F: Me interesa sembrar preguntas, acompañar al público a buscar posibles respuestas y luego involucrarlo en esa reflexión.

Hay preguntas que se exploran en *Inteligencia Actoral*, como ¿realmente es importante el teatro? ¿Por qué es importante, y más en nuestro tiempo? Me pregunto a profundidad si el teatro es una actividad ya obsoleta, en un mundo industrializado de producción en serie. Otra reflexión que planteo es: ¿qué nos hace



humanos? Esa es una pregunta fundamental de la ciencia ficción.

O: ¿Cuál es tu opinión respecto a la interacción que tenemos los humanos con la inteligencia artificial?

F: Creo que ahí hay un malentendido. Pensamos que la inteligencia artificial es una herramienta a la que nosotros le indicamos cómo nos debe ayudar. Pero hay poca conciencia de que está ocurriendo el proceso contrario, es decir, nosotros le servimos a la inteligencia artificial.

En el día a día tenemos muchas interacciones con inteligencia artificial; yo me doy cuenta de algunas y, probablemente, de muchas otras no. Obliga a repensar todo el esquema de la individualidad y me parece bien porque es muy criticable, desde muchos puntos de vista.

Si te perdiste *Inteligencia Actoral* en el Teatro Helénico te recomendamos seguir las redes sociales de Erizo Teatro, pues están en búsqueda de un espacio para ofrecer otra temporada y, tal vez, tengas oportunidad de ver una obra de teatro estelarizada por un reemplazoide.

## Si tuviera un actor androide, por supuesto que me sentaría a jugar un rato con él.



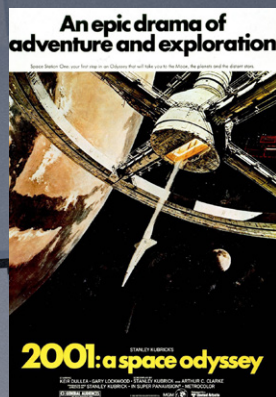


# INTELIGENCIA ARTIFICIAL EN LITERATURA, CINE Y POLÍTICA

¿Te quedaste con ganas de un poco más? Aquí recopilamos los libros y películas sobre inteligencia artificial que han sido mencionados a lo largo de este número. La mayoría son clásicos de las letras y el cine, por lo que te sugerimos preparar palomitas o tu bebida favorita, ¡y a disfrutar!



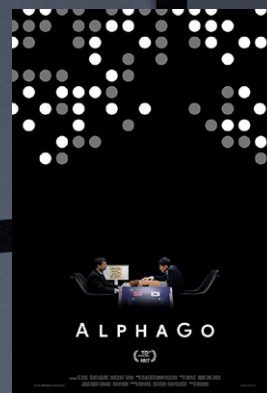
Metrópolis (1927)  
Fritz Lang



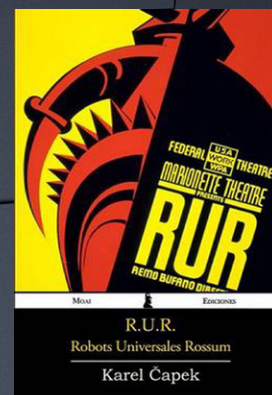
2001 Odisea en el Espacio (1968)  
Stanley Kubrik



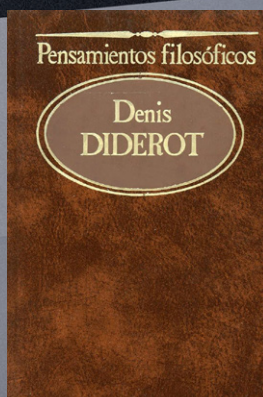
Blade Runner (1982)  
Ridley Scott



AlphaGo (2017)  
Greg Kohs



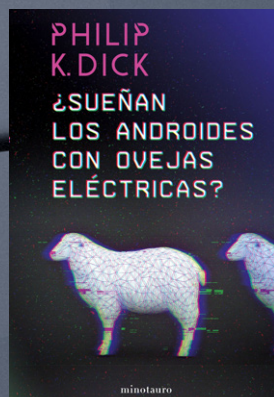
R.U.R. (1921)  
Karel Capek



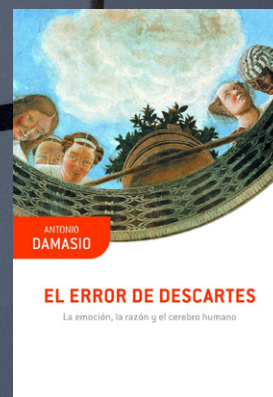
Pensamientos Filosóficos (1746)  
Denis Diderot



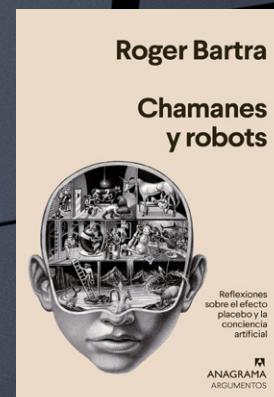
Crónicas Marcianas (1950) Ray Bradbury



¿Sueñan los androides con ovejas eléctricas? (1968) Philip K. Rick



El error de Descartes (1994) Antonio Damasio



Chamanes y robots (2019) Roger Bartra



**Corcholatas**  
Interpretaciones de DALL-E a la contienda política que se vive actualmente en México, en la cual los "suspirantes" están catalogados como "corcholatas". DALL-E parte de una descripción verbal de la imagen y la genera utilizando redes neurales.



# 25 AÑOS DE IMPULSO A LA CIENCIA DESDE LOS ESTADOS



**DRA. ROSALBA MEDINA RIVERA**  
Presidenta



Red Nacional de Consejos y Organismos  
Estatales de Ciencia y Tecnología, A.C.

## LA HISTORIA

La Red Nacional de Consejos y Organismos Estatales de Ciencia y Tecnología, A.C., REDNACECYT, es una Asociación Civil constituida el 23 de noviembre de 1998 en la ciudad de Guanajuato por Coahuila, Durango, Guanajuato, Michoacán, Querétaro, San Luis Potosí, Sinaloa y Tamaulipas.

Su creación ha sido de gran relevancia en la historia del fomento a la ciencia, la tecnología y la innovación en México mediante un papel proactivo en el diseño e implementación de las políticas del sector de ciencia, tecnología e innovación y su aplicación al desarrollo.

La REDNACECYT constituye un foro permanente de intercambio de mejores prácticas de programas y acciones que fomentan la investigación científica, el desarrollo tecnológico y la innovación, desde las entidades federativas, a través de la interacción de sus integrantes, así como de la coordinación con los sectores académico, gubernamental, social y empresarial.

Actualmente la REDNACECYT integra a los 32 Organismos responsables de la política científica y tecnológica en las entidades federativas.

## LÍNEAS DE ACCIÓN

Para cumplir con su propósito, la REDNACECYT, a través de comisiones multidisciplinarias trabaja mediante cuatro líneas de acción que permiten atender las demandas del sector:

- Planeación y Desarrollo Regional
- Recursos Humanos de Alto Nivel
- Apropiación Social de la Ciencia,

- la Tecnología y la Innovación
- Innovación y Transferencia de Tecnología

## PRINCIPALES PROGRAMAS

Con el objetivo de impulsar la investigación científica y tecnológica y fomentar las vocaciones científicas entre los niños, jóvenes, investigadores, divulgadores y sociedad en general, la REDNACECYT organiza anualmente:

- Feria Mexicana de Ciencias e Ingenierías
- Encuentro Nacional de Apropiación Social de la Ciencia, la Tecnología y la Innovación
- La Semana Nacional del Conocimiento y la Innovación

Además, ofrece talleres y cursos que buscan fortalecer las capacidades en el sector de ciencia y tecnología.

## 25 AÑOS

A 25 años de su fundación, en un contexto nacional e internacional de constante cambio, ante nuevos retos que el propio crecimiento de la organización ha generado, la REDNACECYT se consolida como una organización que articula a los actores del sector científico, tecnológico y de innovación.

Este 2023 los invitamos a celebrar los avances, logros y aprendizaje en el marco de su 25 Aniversario que se realizará en el mes de noviembre con un programa de actividades que próximamente daremos a conocer en [www.rednacecyt.org](http://www.rednacecyt.org).



@Rednacecyt



REDNACECYT



[presidencia@rednacecyt.org](mailto:presidencia@rednacecyt.org)





Ciencia y Cultura por México



[www.obsidiana-mexico.com](http://www.obsidiana-mexico.com)