

Historia y evolución de la inteligencia artificial





ANDRÉS ABELIUK

Profesor Asistente del Departamento de Ciencias de la Computación de la Universidad de Chile. Ph.D en Ciencias de la Computación por la Universidad de Melbourne, Australia. Líneas de investigación: computación social e inteligencia colectiva, análisis de redes sociales e impacto de la inteligencia artificial en la sociedad.

aabeliuk@dcc.uchile.cl



CLAUDIO GUTIÉRREZ

Profesor Titular del Departamento de Ciencias de la Computación de la Universidad de Chile. Investigador Senior del Instituto Milenio Fundamentos de los Datos. Licenciado en Matemáticas por la Universidad de Chile y Ph.D. en Computer Science por la Wesleyan University. Líneas de investigación: fundamentos de los datos, bases de datos, lógica aplicada a la computación y semántica de la Web.

cgutierr@dcc.uchile.cl

El primer programa de IA

En 1842, la matemática y pionera de la informática, Ada Lovelace, programó el primer algoritmo destinado a ser procesado por una máquina. Adelantada a su época, Ada especuló que la máquina “podría actuar sobre otras cosas además de los números... el motor (la máquina) podría componer piezas musicales elaboradas y científicas de cualquier grado de complejidad o extensión”. Décadas más tarde, la visión de Ada es una realidad gracias a la Inteligencia Artificial (IA). Sin embargo, un hito considerado como el momento fundacional de la “inteligencia artificial”, tanto del término como del campo de estudio, es una conferencia en Darmouth el año 1956 organizada por John McCarthy, Marvin Minsky, Claude Shannon y Nathaniel Rochester [1]. En ella, los organizadores invitaron a unos diez investigadores para formalizar el concepto de inteligencia artificial como un nuevo campo de estudio científico. Pioneros de la IA, cuatro de los asistentes fueron posteriormente galardonados con el premio Turing (a menudo denominado Premio Nobel de informática) por sus contribuciones a la IA. Una idea común entre los asistentes, y profundamente arraigada hasta el día de hoy en el estudio de la IA, es que el pensamiento es una forma de computación no exclusiva de los seres humanos o seres biológicos. Más aún, existe la hipótesis de que la inteligencia humana es posible de replicar o simular en máquinas digitales.

Ese mismo año dos de los participantes de la conferencia, Alan Newell y Herbert Simon, publican lo que es considerado el primer programa computacional de inteligencia artificial [2]. El programa “Logic Theory Machine” es capaz de descubrir demostraciones de teoremas en lógica simbólica. La idea principal es que a través de la combinación de simples operaciones primitivas, el programa puede ir construyendo expresio-

nes cada vez más complejas. El desafío computacional radica en encontrar la combinación de operaciones que demuestran un teorema dado, entre una cantidad exponencial de posibles combinaciones. La contribución de los autores fue idear un enfoque heurístico, o de reglas generales, que permiten recortar el árbol de búsqueda de manera “inteligente” y encontrar una solución en la mayoría de los casos, pero no siempre. La introducción de los procesos heurísticos han influenciado enormemente la ciencia de la computación y según los mismos autores, son la magia central en toda resolución de problemas humanos. No es coincidencia que esta tesis provenga de Herbert Simon, quien recibió el Nobel en economía por la provocadora idea de modelar el comportamiento humano, no como un agente “homo economicus” totalmente racional, sino que con “racionalidad limitada” cuya toma de decisiones es principalmente heurística [3].

Dos paradigmas de investigación en IA

IA simbólica

La búsqueda heurística fue un pilar clave para los avances de la IA en sus comienzos. Todo tipo de tareas de resolución de problemas, como probar teoremas y jugar ajedrez, implican tomar decisiones que se pueden modelar como un árbol de decisiones que debe ser recorrido para encontrar una estrategia que resuelva el problema. Los algoritmos de búsqueda heurística son parte de una colección de métodos que se basan en representar el conocimiento implícito o procedimental que poseen los humanos de forma explícita, utilizando símbolos y reglas (legibles por humanos) en programas informáticos. La “IA simbólica” demostró ser muy exitosa en las primeras décadas de la IA logrando codificar

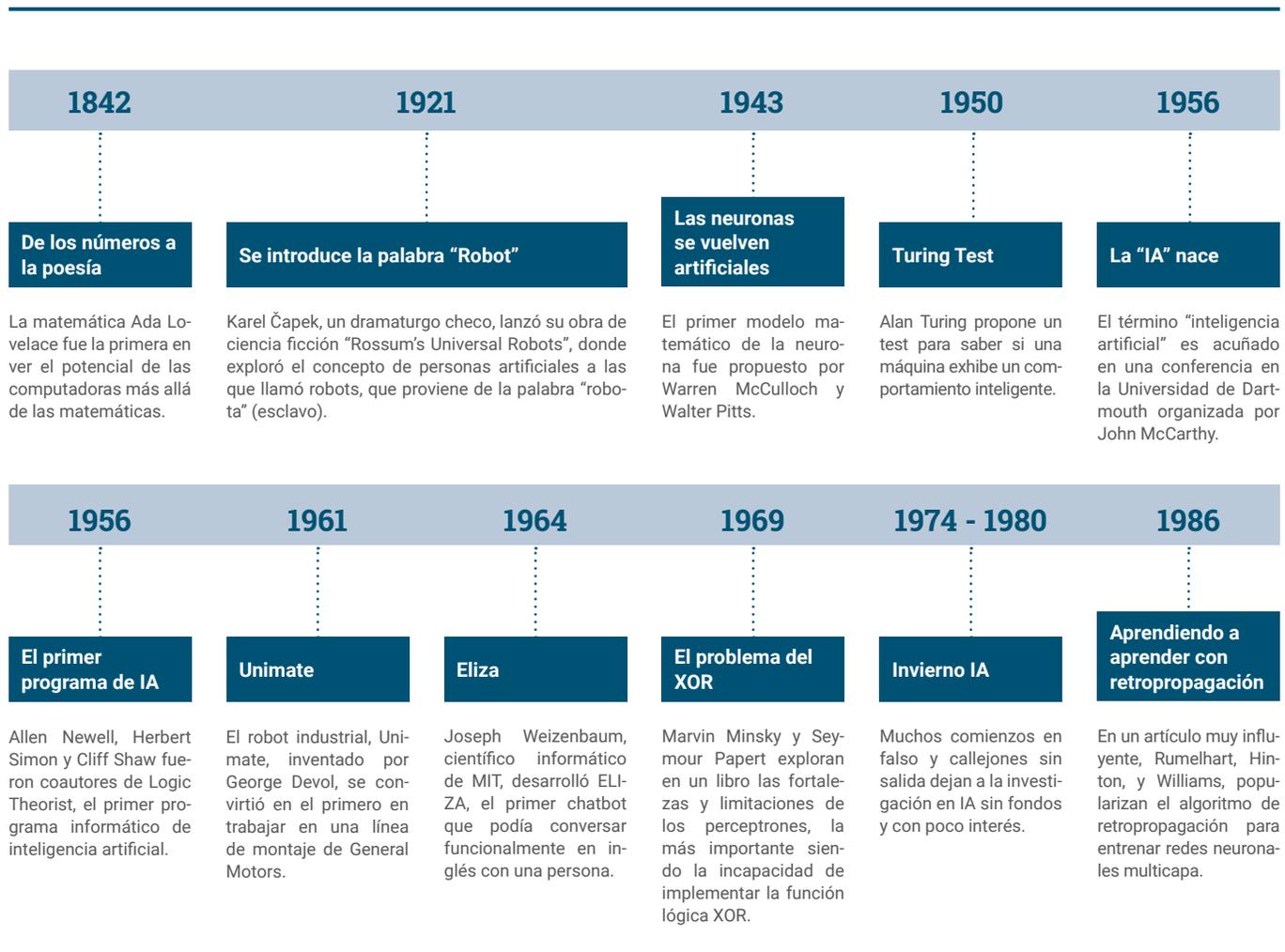
en “sistemas expertos” el razonamiento humano en dominios de conocimiento específico. Un ejemplo son los sistemas de apoyo de diagnóstico médico a través de motores de inferencia y bases de conocimientos que resumen el conocimiento médico basado en evidencia. Uno de los logros más populares de la IA simbólica culmina con la derrota del campeón mundial de ajedrez en 1997, Garry Kasparov, por el computador Deep Blue de IBM [4] (ver infografía de línea de tiempo en la Figura 1).

IA conexionista

Paralelamente con la emergencia de la IA simbólica, que modela la mente hu-

mana como si fuese una computadora procesadora de símbolos, existe otra escuela de pensamiento que se basa en modelar la biología del cerebro que está compuesto por redes neuronales biológicas. Frank Rosenblatt (psicólogo) en 1958 propuso el *perceptrón*, una generalización de la neurona McCulloch-Pitts que podía “aprender” a través de coeficientes de ponderación para cada entrada de la neurona. Hasta el día de hoy, el perceptrón es la unidad fundamental para muchas de las redes neuronales artificiales e impulsa el paradigma conocido como IA conexionista. A pesar de su promesa, la investigación en redes neuronales se detuvo por falta de financiamiento y una sobreexpectación no cumplida. Hechos que parcialmente

son atribuidos a una malinterpretada exposición de las limitaciones y fortalezas del perceptrón en un libro por pioneros de la IA simbólica, Marvin Minsky y Seymour Papert en 1969 [5]. No fue hasta comienzos de 1980, que Geoffrey Hinton (Premio Turing en 2018) y colegas redescubren y popularizan el método llamado *retropropagación* [6]; el algoritmo central detrás de la *búsqueda heurística* (estilo IA simbólica) que logra encontrar los parámetros del modelo que minimizan su error, así permitiendo que una red neuronal de múltiples capas aprenda a partir de datos. Este avance resuelve las limitaciones de los perceptrones de Rosenblatt y crea un resurgimiento en la investigación del aprendizaje profundo (ver Figura 1).



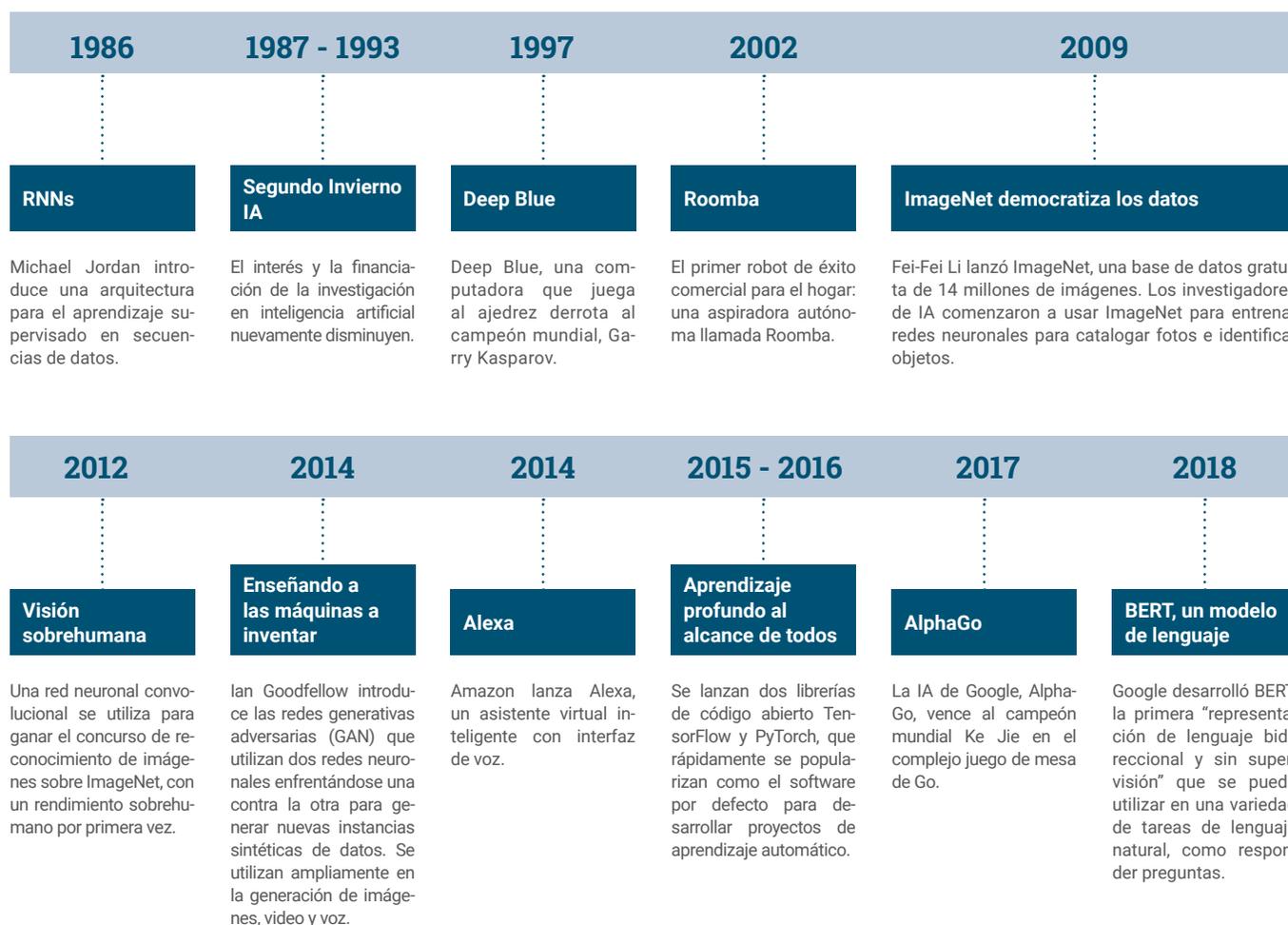


Figura 1. Historia de la inteligencia artificial.

La revolución del aprendizaje profundo

En la década del 2010, dos cosas harían posible la revolución de aplicaciones de redes neuronales y algoritmos de aprendizaje profundo. Primero, los avances de hardware especializado han acelerado drásticamente el entrenamiento y el rendimiento de las redes neuronales y reducido su consumo de energía. Segundo, el aumento de datos abiertos disponibles online y servicios de bajo costo para etiquetar datos vía *crowdsourcing*

impulsan el desarrollo de la IA. La Figura 2 muestra cómo los conceptos de IA, aprendizaje automático (*machine learning*) y aprendizaje profundo (*deep learning*) se relacionan el uno con el otro.

Como consecuencia de estos avances, se desarrollaron aplicaciones basadas en las redes neuronales donde la IA simbólica no tuvo éxito. Por ejemplo en aplicaciones de visión, como reconocimiento facial y detección de cáncer, y en aplicaciones de lenguaje, como la traducción de idiomas y asistentes virtuales. En 2015, Microsoft Research utiliza una arquitectura de red neuronal para catego-

rizar imágenes con una mayor precisión que el humano promedio [7]. Al siguiente año, el sistema AlphaGo de DeepMind se corona maestro de Go tras vencer al campeón mundial, Lee Sedol [8]. Este suceso es impactante ya que en el Go hay en promedio alrededor de 300 movimientos posibles que se pueden hacer en cada turno, mientras que en el ajedrez es cercano a 30 movimientos. En otras palabras, el árbol de búsqueda del Go tiene un factor de ramificación de un orden de magnitud mayor al ajedrez, razón principal por la cual la IA simbólica, por sí sola, falló en desarrollar un programa para jugar Go.



Limitaciones de la IA

Un aspecto clave y poderoso de las redes neuronales es que no requieren que se especifiquen las reglas del dominio a modelar; las reglas se aprenden a partir de los datos de entrenamiento. La falta de conocimiento de alto nivel embebido en el sistema por expertos humanos, como es el caso de la IA simbólica, se contrarresta con la capacidad de inferir estadísticamente un modelo del dominio a partir de suficientes datos. Sin embargo, una desventaja importante de las redes neuronales es que requieren grandes recursos computacionales y cantidades enormes de datos. Por ejemplo, se estima que replicar los experimentos de AlphaGo costaría alrededor de 35 millones de dólares sólo en poder computacional [9]. Por otro lado, los datos deben ser cuidadosamente “curados” para ser representativos y así poder generalizar correctamente y no producir resultados sesgados, como ha sido el caso en textos sexistas y racistas generados a partir de modelos de lenguaje [10]. Por otro lado, mientras que programas de software basados en reglas explícitas son fáciles de rastrear y comprender cómo llegaron a tomar ciertas decisiones, no se puede decir lo mismo de los algoritmos de aprendizaje profundo que debido a su alta complejidad son difíciles de interpretar y comunicar por humanos. Estas limitaciones son uno de los grandes desafíos en la IA y hay mucha investigación activa en estas direcciones [11,12].

Democratizando la IA

Desde que el aprendizaje profundo recuperó prominencia alrededor del 2010, los softwares gratuitos y de código abierto especializados para el aprendizaje profundo han sido enormemente responsables de impulsar el campo hacia adelante. Desde las primeras librerías



Figura 2. Diagrama de Venn que muestra la relación entre distintas subáreas de la inteligencia artificial.

creadas por equipos académicos, Caffe y Theano, hasta las actuales dominantes, PyTorch y TensorFlow, respaldadas por Facebook y Google, respectivamente, el acceso a estos softwares de código abierto han facilitado el cambio hacia la innovación tecnológica impulsada por el aprendizaje automático. Tanto en la investigación de vanguardia como en la creación de aplicaciones por la industria, la democratización de la IA reduce las barreras de entrada para que las personas y organizaciones puedan ingresar al apasionante mundo de la IA con poca o nada de inversión financiera. Pueden aprovechar los datos y algoritmos disponibles públicamente para comenzar a

experimentar la construcción de modelos de IA y a la vez contribuir a expandir las bases de datos públicas y poner a disposición nuevas soluciones.

Como ejemplo del poder de democratizar datos, en el 2009 el proyecto ImageNet, liderado por la investigadora Fei-Fei Li, puso a disposición del público una gran base de datos visual que ayudó a investigadores a crear modelos más rápidos y precisos de reconocimiento visual de objetos. Esta colección de imágenes se convirtió rápidamente en una competencia anual (ahora organizada en Kaggle) para ver qué algoritmos podían identificar objetos en las imágenes

La búsqueda heurística fue un pilar clave para los avances de la IA en sus comienzos.

con la tasa de error más baja. El 2012, el primer equipo en usar redes neuronales en la competencia venció el estado del arte con una precisión récord. La arquitectura propuesta por integrantes del laboratorio de Geoffrey Hinton en la Universidad de Toronto, Red Neuronal Convolutiva [13], fue inspirada por las características estructurales y fisiológicas de la visión animal. Hoy en día, estas redes neuronales están en todas partes: se usan para etiquetar las fotos en plataformas sociales; los vehículos autónomos las utilizan para detectar objetos; y se usan para digitalizar textos.

Desde entonces, se han introducido una multitud de nuevos conjuntos de datos estimulando investigación en subcampos de la IA como el procesamiento de lenguaje natural (NLP) y reconocimiento de voz y audio. La arquitectura precursora en NLP, es la Red Neuronal Recurrente, que usa datos secuenciales

y se distingue por su “memoria”, ya que al iterar sobre la entrada, mantiene un estado interno que codifica información sobre los elementos anteriores dentro de la secuencia e influenciando el output actual [14]. El procesamiento de lenguaje natural juega un papel vital en muchos sistemas, desde el análisis de currículums para la contratación, hasta asistentes virtuales y detección de spam. Sin embargo, el desarrollo y la implementación de la tecnología de NLP no es tan equitativo como parece. Aunque se hablan más de 7000 idiomas en todo el mundo, la gran mayoría de los avances tecnológicos son aplicados al inglés. Una iniciativa para contrarrestar esta inequidad es liderada por Jorge Pérez, académico del Departamento de Ciencias de la Computación (DCC) de la Universidad de Chile, que junto a estudiantes han puesto a libre disposición de la comunidad un modelo de lenguaje en español [15].

IA en Chile

Describiremos a grandes rasgos el desarrollo actual de la IA en Chile en tres áreas: empresarial, investigación académica, y enseñanza y propuestas.

A nivel empresarial las técnicas de IA ya están comenzando a ser un *commodity*, esto es, están a disposición en el mercado regular y se están usando de manera generalizada (particularmente en lo que respecta a aprendizaje por medio de datos). Otra pregunta es si hay desarrollos “novedosos”. A manera de ejemplo nombraremos cuatro. NotCo, cuyo logro es “combinar la inteligencia artificial con el conocimiento del mundo vegetal para crear productos”. Usan técnicas de análisis de datos y visualización innovadoras, pero hay poca información sobre su nivel innovador de IA/ML. El otro ejemplo es Fintual. Aquí se usan bastantes cosas que se pueden considerar IA, por ejemplo, “bots de inversión” que permiten seguir ciertos índices para invertir de





Un aspecto clave [...] de las redes neuronales es que no requieren que se especifiquen las reglas del dominio a modelar; las reglas se aprenden a partir de los datos de entrenamiento.

forma pasiva y a bajo costo. El tercer ejemplo es CornerShop que usa tecnologías de datos y analítica para su diseño de operaciones. Finalmente, un ejemplo de una empresa más pequeña es Zippedi, orientada a robótica de almacenes para optimizar digitalmente las estanterías. Hay también muchas otras empresas tipo *startup* que están haciendo cosas tipo *chatbots*, aplicaciones de procesamiento de imágenes, bioinformática, etc., la mayoría aplicando investigación ya consolidada (no desarrollando).

Respecto de la investigación dedicada a la IA propiamente tal (esto es, publicando regularmente en revistas o conferencias de IA) son pocos los grupos a lo largo del país. Destacamos IALab de la Pontificia Universidad Católica de Chile, que tiene varios años y buena infraestructura (particularmente su *cluster* de GPUs para IA). Su fuerte es visión computacional y robótica. Otro grupo es el de Inteligencia Computacional del Departamento de Ingeniería Eléctrica (UChile) que está centrado en robótica y visión, y procesamiento de señales y aprendizaje en este campo. En el DCC (UChile) hay un grupo (ReLeLa) centrado en IA y NLP. La Universidad de Concepción recientemente creó un grupo de IA enfocado a Sistemas Multiagente y Robótica. Por otra parte, hay muchos grupos en diversas universidades dedicados más bien a aplicaciones de AI en diferentes áreas, como empresarial, comercial, científica, social, etc. y luego publican en esas disciplinas. Por ejemplo, el Instituto Data Science de la Universidad del Desarrollo (UDD) aplica técnicas de IA en proyectos asociados a la minería y agricultura. Finalmente, hay muchos investigadores que trabajan más bien

solos o con colegas de otras instituciones en diversas universidades a lo largo del país.

Respecto de la enseñanza, han proliferado los cursos de IA, así como diplomados y magíster en el área dictados por diferentes universidades y organizaciones. Esto muestra que la IA se ha convertido en un *boom* en Chile, con los claros y oscuros propios de un *boom*. En este marco diferentes organizaciones e instituciones discuten sobre los usos de IA en diferentes áreas, entre ellos, la Comisión Desafíos del Futuro del Senado, las universidades, las Fuerzas Armadas, el Ministerio de Ciencia y Tecnología, etc. Se han elaborado diversos documentos. Algunos ejemplos son: "Inteligencia Artificial para Chile. La urgencia de desarrollar una estrategia", del Senado de la República; "Ejército Virtual" de la Academia Politécnica Militar, y "Política Nacional de Inteligencia Artificial" del Ministerio de Ciencia.

Ética, alcances y limitaciones de la IA

Como toda tecnología, la IA trae aparejada dilemas éticos. En el caso de la IA esto se agranda por el poder transformador de la realidad que puede traer aparejado esta tecnología. Así es que hoy, al igual que la investigación biomédica desde siempre, el test de la ética debe ser aplicado a los desarrollos de IA. Esto se refiere particularmente a funcionalidades donde existen máquinas y aparatos con "inteligencia" o habilidades de simulación de lo humano que sobrepasan con creces las de los humanos. Y las preguntas fundamenta-

les van por el lado del marco ético para los desarrollos en esta disciplina. Mencionaremos algunos de los principales:

1. *La IA y la economía* [16]. Aquí aparecen temas como los usos de la IA en el mundo del trabajo: por ejemplo, ¿dónde están los límites de los flujos de trabajo automatizado donde hay personas involucradas? Y la pregunta fundamental del área: ¿cómo distribuiremos la riqueza creada por las máquinas?
2. *La IA y la sociedad* [17]. ¿Cómo afectan las máquinas inteligentes la relación entre los seres humanos? ¿Quiénes decidirán los usos de las máquinas inteligentes? ¿Quién y cómo controlar los sesgos (introducidos intencional o no intencionalmente) a las máquinas? ¿Cuáles son los límites (o no existen) al desarrollo de ese tipo de proyectos?
3. *La IA y los humanos*. ¿Cómo afectarán las máquinas inteligentes nuestro comportamiento? ¿Hasta qué nivel es permisible "ensamblar" esas máquinas con nuestra biología?
4. *La IA y el medio ambiente*. ¿Cuáles son los límites razonables de uso de recursos para estos proyectos?
5. *Seguridad, usos militares* [18]. ¿Qué es necesario y cómo regular este ámbito, tradicionalmente complejo de regular?
6. *Superinteligencia* [19]. ¿Qué derechos y deberes tendrán estos robots? ¿Quién es responsable por sus desarrollos y usos? ¿Qué nivel de decisiones se les permitirá tomar en asuntos humanos?

Hay miles de otras preguntas. Uno podría replicar todas las de la ética clásica, pues en definitiva lo que está ocurriendo con la IA débil al menos, es la realización de gran parte de los proyectos clásicos de simulación de facetas de lo humano. ■

REFERENCIAS

- [1] James Moor. "The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years". *AI Magazine* 27(4), 2006. <https://doi.org/10.1609/aimag.v27i4.1911>.
- [2] A. Newell y H. Simon. "The Logic Theory Machine – A Complex Information Processing System". *IRE Transactions on Information Theory* 2, 1956.
- [3] Wheeler, Gregory. "Bounded Rationality". *The Stanford Encyclopedia of Philosophy*, 2020. <https://plato.stanford.edu/archives/fall2020/entries/bounded-rationality/>.
- [4] Hansen Hsu. AI and Play, part 1: How Games Have Driven two Schools of AI Research, Computer History Museum, 2020. <https://computerhistory.org/blog/ai-and-play-part-1-how-games-have-driven-two-schools-of-ai-research/>.
- [5] Minsky, Marvin, y Seymour A. Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT press, 2017.
- [6] Rumelhart, D. E., Hinton, G. E., y Williams, R. J. Learning Representations by Back-Propagating Errors. *Nature*, 1986.
- [7] He, Kaiming, et al. "Delving Deep into Rectifiers: Surpassing Human-Level Performance on Imagenet Classification". *Proceedings of the IEEE international conference on computer vision*, 2015.
- [8] Silver, D., Huang, A., Maddison, C. et al. Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature* 529, 2016.
- [9] DeepMind's Losses and the Future of Artificial Intelligence. WIRED, 2019. <https://www.wired.com/story/deepminds-losses-future-artificial-intelligence/>.
- [10] Zou, J. y Schiebinger, L. AI can be Sexist and Racist – It's Time to Make it Fair. *Nature* 559, 324–326, 2018.
- [11] Xie, Ning, et al. "Explainable Deep Learning: A Field Guide for the Uninitiated". *arXiv preprint*, 2020.
- [12] Mehrabi, Ninareh, et al. "A Survey on Bias and Fairness in Machine Learning". *arXiv preprint*, 2019.
- [13] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet Classification with Deep Convolutional Neural Networks". *Advances in neural information processing systems* 25, 2012.
- [14] Mikolov, Tomáš, et al. "Recurrent Neural Network based Language Model". Eleventh annual conference of the international speech communication association, 2010.
- [15] Cañete, José, Gabriel Chaperon, Rodrigo Fuentes, y Jorge Pérez. "Spanish pre-trained Bert Model and Evaluation Data". *PML4DC at ICLR 2020*, 2020.
- [16] Egana-delSol, Pablo. "The Future of Work in Developing Economies: What can we learn from the South?". *Available at SSRN 3497197*, 2019.
- [17] Tomašev, N., Cornebise, J., Hutter, F. et al. AI for Social Good: Unlocking the Opportunity for Positive Impact. *Nat Commun* 11, 2468, 2020.
- [18] Toby Walsh A.I. Expert, Is Racing to Stop the Killer Robots. *The New York Times*, 2019. <https://www.nytimes.com/2019/07/30/science/autonomous-weapons-artificial-intelligence.html>.
- [19] Alfonseca, M., Cebrián, M., Anta, A. F., Coviello, L., Abeliuk, A., y Rahwan, I. Superintelligence cannot be Contained: Lessons from Computability Theory. *Journal of Artificial Intelligence Research*, 2021.