

# ATLAS DE IA



# Kate Crawford

## ATLAS DE IA

PODER, POLÍTICA Y COSTES PLANETARIOS  
DE LA INTELIGENCIA ARTIFICIAL



Título original en inglés: *Atlas of AI*  
© 2021 by Kate Crawford  
Originally published by Yale University Press

© De la traducción: Francisco Diaz Klaassen  
De la corrección: Marta Beltrán Vahón  
Montaje de cubierta: Juan Pablo Venditti

Primera edición: abril 2023  
© Ned Ediciones, 2023

Depósito legal: B 6552-2023  
ISBN: 978-84-19407-02-3

Preimpresión: Moelmo, SCP  
[www.moelmo.com](http://www.moelmo.com)

Impreso en Sagrafic  
*Printed in Spain*

Derechos reservados para todas las ediciones en castellano

**Ned Ediciones**  
[www.nedediciones.com](http://www.nedediciones.com)

# Índice

<i>Agradecimientos</i> .....	11
<i>Introducción</i> .....	17
I. <i>La Tierra</i> .....	47
II. <i>El trabajo</i> .....	89
III. <i>Los datos</i> .....	139
IV. <i>La clasificación</i> .....	187
V. <i>Las emociones</i> .....	229
VI. <i>El Estado</i> .....	273
<i>Conclusión. El poder</i> .....	319
<i>Coda. El espacio</i> .....	345
<i>Bibliografía</i> .....	361
<i>Índice de nombres</i> .....	437

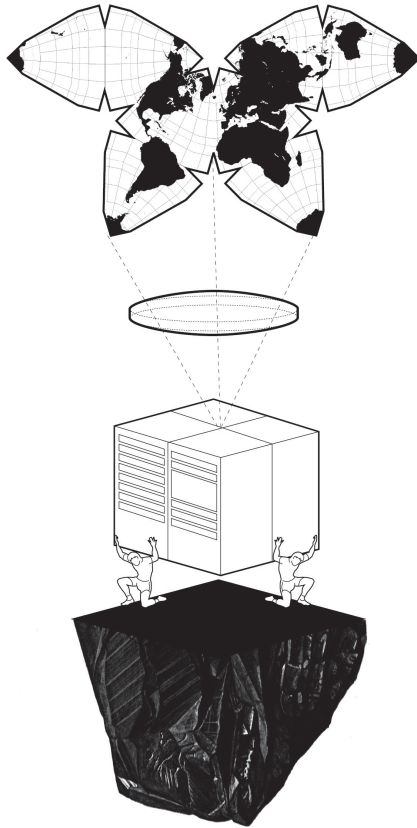


*Para Elliott y Margaret*





# *Introducción*





## EL CABALLO MÁS INTELIGENTE DEL MUNDO

A fines del siglo XIX, Europa se encontraba embelesada por un caballo llamado Hans. "Clever Hans"\* era una verdadera maravilla: podía resolver problemas matemáticos, decir la hora, identificar días en el calendario, diferenciar tonos musicales y deletrear palabras y frases. La gente acudía en masa a ver al semental alemán golpeteando con el casco las respuestas a problemas complejos y dando consistentemente las respuestas correctas. "¿Cuánto es dos más tres?" De manera diligente, Hans chocaba un casco contra el suelo cinco veces. "¿Qué día de la semana es?" El caballo entonces se acercaba a un tablero especialmente diseñado y deletreaba la respuesta correcta usando el mismo método. Hans llegó incluso a dominar preguntas más complejas: "Estoy pensando en un número. Si le quito nueve, me quedan tres. ¿De qué número se trata?". Para 1904, Clever Hans ya era famoso en el mundo entero, y *The New York Times*

\* "Hans el astuto", en castellano. A pesar de que el apelativo original era alemán ("Hans der Kluge"), mundialmente se lo conoció con la versión en inglés. [N. del T.]

promovía sus habilidades con el siguiente titular: “El maravilloso caballo de Berlín. ¡Solo le falta hablar!”.<sup>1</sup>

El entrenador de Hans, Wilhelm von Osten —un profesor de matemáticas retirado—, había estado fascinado por la inteligencia animal durante mucho tiempo. Intentó sin éxito enseñarles a gatitos y crías de oso los números cardinales, y solo triunfó cuando empezó a trabajar con su propio caballo. Primero, le enseñó a contar sujetándole la pata, mostrándole un número y luego golpeteando con el casco el número correcto de veces. Pronto Hans fue capaz de realizar sumas simples. A continuación, Von Osten incorporó un pizarrón que tenía escrito el abecedario, de modo que el caballo pudiera hacer chocar un casco sobre un número para cada letra. Después de dos años de entrenamiento, Von Osten estaba asombrado por la capacidad del animal para comprender conceptos intelectuales avanzados. Así que se lo llevó de gira para probarle al mundo que los animales podían razonar. Hans se viralizó durante la *Belle Époque*.

Pero había muchos escépticos, y el consejo educativo alemán creó una comisión de investigación para comprobar la validez de las declaraciones científicas hechas por Von Osten. La Comisión Hans estaba dirigida por el psicólogo y filósofo Carl Stumpf y su asistente Oskar Pfungst, e incluía a un director de circo, un profesor retirado, un zoólogo, un veterinario y un oficial de caballería. Sin embargo, después de que interrogaran largamente a Hans, tanto con su entrenador presente como sin él, el caballo mantuvo su registro de respuestas correctas, y la comisión no fue capaz de encontrar evidencia alguna de fraude. Como escribiría más tarde Pfungst, Hans actuó frente a “miles de espectador-

<sup>1</sup> Edward T. Heyn, “Berlin’s Wonderful Horse; He Can Do Almost Everything but Talk - How He Was Taught”, en *The New York Times*, 4 de septiembre de 1904, disponible en línea: <<https://timesmachine.nytimes.com/timesmachine/1904/09/04/101396572.pdf>>.



*Wilhelm von Osten y Clever Hans.*

res, aficionados a los caballos, entrenadores de calibre, y ninguno de ellos, durante el transcurso de observaciones que duraron muchos meses, descubrió algún tipo de señal secreta regular” entre el interrogador y el caballo.<sup>2</sup>

La comisión descubrió que los métodos enseñados a Hans se asemejaban más a “la enseñanza en la educación básica” que al entrenamiento de animales, y eran “dignos de estudiarse científicamente”.<sup>3</sup> Pero Stumpf y Pfungst todavía tenían algunas dudas. Un hallazgo en particular los preocupaba: cuando el mismo interrogador no sabía la respuesta a la pregunta que estaba for-

<sup>2</sup> Oskar Pfungst, *Clever Hans (The Horse of Mr. von Osten). A Contribution to Experimental Animal and Human Psychology*, trad. de Carl L. Rahn, Nueva York, Henry Holt, 1911, s. d.

<sup>3</sup> “Clever Hans’ Again: Expert Commission Decides that the Horse Actually Reasons”, en *The New York Times*, 2 de octubre de 1904, disponible en línea: <<https://timesmachine.nytimes.com/timesmachine/1904/10/02/120289067.pdf>>.

mulando o cuando estaba muy lejos de Hans, este rara vez acertaba. Esto los llevó a considerar que quizás alguna señal involuntaria le estuviera dando las respuestas a Hans.

Como Pfungst describió en su libro de 1911, no se equivocaban: la postura del interrogador, así como su respiración y expresión facial, cambiaban sutilmente cuando los golpeteos de Hans alcanzaban la respuesta correcta, lo que lo hacía parar justo en ese momento.<sup>4</sup> Pfungst probó esta hipótesis más tarde en sujetos humanos y confirmó sus resultados. Para él, la parte deslumbrante de este descubrimiento era que, por lo general, los interrogadores no sabían que le estaban dando pistas al caballo. La solución al enigma de Clever Hans, escribió Pfungst, eran las indicaciones inconscientes de los interrogadores.<sup>5</sup> El caballo estaba entrenado para producir los resultados que su dueño quería ver, pero el público se sintió estafado porque esta no era la inteligencia extraordinaria que se habían imaginado.

La historia de Clever Hans es cautivante por muchos motivos: la relación entre el deseo, la ilusión y la acción; el negocio del espectáculo; la manera en que antropomorfizamos lo no humano; cómo emergen los sesgos y la política de la inteligencia. Hans inspiró un término en psicología que se usa para un tipo especial de trampa conceptual, el efecto Clever Hans, también llamado sesgo de expectativa, que describe la influencia que tienen las pistas involuntarias que los investigadores dan a los sujetos de investigación. La relación entre Hans y Von Osten tiene que ver con los complejos mecanismos a través de los cuales los sesgos se abren camino en los sistemas y cómo la gente termina enredada con los fenómenos que estudia. En la actualidad, la historia de Hans se usa en aprendizaje automático como un recordatorio

<sup>4</sup> Oskar Pfungst, *op. cit.*

<sup>5</sup> *Ibid.*

admonitorio de que no siempre se puede estar seguro de lo que ha aprendido un modelo a partir de los datos que se le han dado.<sup>6</sup> Incluso un sistema que parece funcionar de forma espectacular durante las prácticas puede hacer predicciones terribles cuando se le presentan datos nuevos en el mundo real.

Esto nos lleva a la pregunta central de este libro: ¿cómo se “hace” la inteligencia y con qué trampas nos podemos encontrar a partir de este proceso? A primera vista, la historia de Clever Hans es la historia de un hombre que construyó inteligencia entrenando a un caballo para que siguiera pistas y emulara la cognición humana. Pero, en otro nivel, podemos observar que la práctica de crear inteligencia se extendió considerablemente más allá de esta situación. El empeño requirió la validación de múltiples instituciones, incluida la academia, la educación, la ciencia, el público y las fuerzas armadas. Luego, está el mercado en el que se insertaron Von Osten y su notable caballo: inversiones tanto emocionales como económicas que estaban detrás de las giras, los artículos periodísticos, las charlas. Autoridades burocráticas congregadas para sopesar y analizar las habilidades del caballo. Una constelación de intereses financieros, culturales y científicos que cumplieron un rol en la construcción de la inteligencia de Hans y que tenían un interés tangible en que esta fuera realmente notable.

Aquí podemos ver dos mitologías distintas. El primer mito es que los sistemas no humanos (sean computadoras o caballos) son análogos a la mente humana. Esta perspectiva presupone que, con el entrenamiento adecuado o los recursos suficientes, una inteligencia parecida a la de un ser humano se puede crear de cero sin tener en consideración las maneras fundamentales

<sup>6</sup> Sebastian Lapuschkin *et al.*, “Unmasking Clever Hans Predictors and Assessing What Machines Really Learn”, en *Nature Communications*, vol. 10, núm. 1, 2019, pp. 1-8, disponible en línea: <<https://doi.org/10.1038/s41467-019-08987-4>>.

en que las personas se encarnan, se relacionan y se ubican dentro de contextos más amplios. El segundo mito es que la inteligencia es algo que existe de forma independiente, como algo natural y separado de las fuerzas sociales, culturales, históricas y políticas. Este concepto de inteligencia ha causado un daño enorme durante siglos y se ha usado para justificar relaciones de dominación desde la esclavitud hasta la eugenesia.<sup>7</sup>

Estas mitologías prevalecen particularmente en el campo de la inteligencia artificial (IA), donde la creencia de que las máquinas pueden formalizar y reproducir la inteligencia humana ha sido axiomática desde mediados del siglo xx. Así como se consideraba la inteligencia de Hans equiparable a la de un ser humano, cuidadosamente fomentada como la de un niño durante su educación básica, en muchas ocasiones se han descrito los sistemas de IA como formas simples de inteligencia pero parecidas a la humana. En 1950, Alan Turing predijo que “para el fin de siglo el uso de palabras y opiniones generales razonadas habrá cambiado tanto que uno podrá hablar de máquinas pensantes sin esperar a que lo contradigan”.<sup>8</sup> En 1958, el matemático John von Neumann aseguró que el sistema nervioso de los seres humanos es, “a primera vista, digital”.<sup>9</sup> En cierta ocasión, ante la

<sup>7</sup> Véase el trabajo de la filósofa Val Plumwood sobre los dualismos inteligencia-estupidez, emocional-racional y maestro-esclavo. Val Plumwood, “The Politics of Reason: Towards a Feminist Logic”, en *Australasian Journal of Philosophy*, vol. 71, núm. 4, 1993, pp. 436-462, disponible en línea: <<https://doi.org/10.1080/00048409312345432>>.

<sup>8</sup> Alan M. Turing, “Computing Machinery and Intelligence”, en *Mind*, 1° de octubre de 1950, pp. 433-460, disponible en línea: <<https://doi.org/10.1093/mind/LIX.236.433>> [trad. esp.: “Maquinaria computacional e inteligencia”, trad. de Cristóbal Fuentes Barassi para la Universidad de Chile, Santiago de Chile, 2010, disponible en línea: <<http://xamanek.izt.uam.mx/map/cursos/Turing-Pensar.pdf>>].

<sup>9</sup> John von Neumann, *The Computer and the Brain*, New Haven, Yale University Press, 1958, p. 44 [trad. esp.: *El ordenador y el cerebro*, Barcelona, Bosch,



pregunta de si las máquinas podían pensar, el profesor del Massachusetts Institute of Technology (MIT) Marvin Minsky respondió: “Por supuesto que las máquinas pueden pensar; nosotros podemos pensar y somos ‘máquinas hechas de carne’”.<sup>10</sup> Pero no todo el mundo estaba tan convencido. Uno de los precursores de la invención de la IA, Joseph Weizenbaum, creador de ELIZA, el primer programa conversacional, creía que la idea de que los seres humanos son meros sistemas procesadores de información involucraba una noción demasiado simplista de la inteligencia, y que esa noción estaba detrás de la “perversa fantasía de proporciones” que suponía pensar que los científicos de la IA llegarían a crear una máquina que aprendiera “al igual que un niño”.<sup>11</sup>

Esta ha sido una de las discusiones centrales en la historia de la IA. En 1961, el MIT fue sede de una serie de importantes conferencias con el título de “Administración de empresas y el ordenador del futuro”. El programa incluía la participación de una serie estelar de científicos informáticos, entre los que se encontraban Grace Hopper, Joseph Carl Robnett Licklider,

---

1980]. Este método fue muy criticado por Hubert L. Dreyfus en *What Computers Can't Do. A Critique of Artificial Reason*, Nueva York, Harper and Row, 1972.

<sup>10</sup> Véase Joseph Weizenbaum, “On the Impact of the Computer on Society: How Does One Insult a Machine?”, en *Science*, vol. 176, núm. 4035, 1972, p. 612. Tras su muerte, Minsky se vio implicado en serias acusaciones relacionadas con el pedófilo y violador convicto Jeffrey Epstein. Minsky era uno de los muchos científicos que se reunieron con Epstein y lo visitaron en su isla de retiro, donde se obligaba a niñas a tener sexo con los miembros de su camarilla. Como ha observado la académica Meredith Broussard, esto formaba parte de una cultura de exclusión más amplia, que se volvió endémica en la IA: “Por muy creativos que fueran Minsky y su séquito, también solidificaron la cultura de la tecnología como un club de amigos multimillonarios. Las matemáticas, la física y las otras ciencias ‘duras’ nunca han sido hospitalarias con las mujeres y la gente de color; la tecnología siguió este ejemplo”. Véase Meredith Broussard, *Artificial Unintelligence. How Computers Misunderstand the World*, Cambridge, MIT Press, 2018, p. 174.

<sup>11</sup> Joseph Weizenbaum, *Computer Power and Human Reason. From Judgment to Calculation*, San Francisco, W. H. Freeman, 1976, pp. 202 y 203.

Marvin Minsky, Allen Newell, Herbert Simon y Norbert Wiener, para discutir los rápidos avances que se estaban produciendo en la computación digital. En su osada conclusión, John McCarthy argumentó que las diferencias entre las tareas humanas y las de las máquinas eran ilusorias. Se trataba simplemente de que las máquinas requerirían más tiempo para formalizar y resolver algunas de las tareas humanas más complicadas.<sup>12</sup>

Pero el profesor de filosofía Hubert Dreyfus contraargumentó, preocupado por el hecho de que los ingenieros en su conjunto “ni siquiera consideraran la posibilidad de que el cerebro pudiera procesar información de una manera completamente diferente a una computadora”.<sup>13</sup> En una obra posterior, *What Computers Can't Do*, Dreyfus señala que la inteligencia y la pericia humanas dependen en gran medida de muchos procesos, inconscientes y subconscientes, mientras que las computadoras requieren que todos los procesos y los datos sean explícitos y estén formalizados.<sup>14</sup> Como resultado, los aspectos menos formales de la inteligencia tienen que ser abstraídos, eliminados y aproximados, lo que vuelve a las computadoras incapaces de procesar información sobre situaciones como lo harían los seres humanos.

Mucho ha cambiado en la IA desde la década de 1960, incluido un giro de los sistemas simbólicos a la más reciente ola hiperbólica sobre las técnicas de aprendizaje automático. En varios sentidos, las primeras discusiones sobre lo que puede hacer la IA han sido olvidadas y el escepticismo se ha desvanecido. Desde mediados de la década de 2000, la IA se ha expandido rápidamente, como campo académico y como industria. Hoy en día, un pequeño nú-

<sup>12</sup> Martin Greenberger (ed.), *Management and the Computer of the Future*, Nueva York, John Wiley and Sons, 1962, p. 315.

<sup>13</sup> Hubert L. Dreyfus, *Alchemy and Artificial Intelligence*, Santa Monica, RAND, 1965, s. d.

<sup>14</sup> Hubert L. Dreyfus, *What Computers Can't Do*, op. cit.

mero de compañías tecnológicas poderosas hace uso de sistemas de IA a escala planetaria, y sus sistemas son aclamados, una vez más, como similares o incluso superiores a la inteligencia humana.

Sin embargo, la historia de Clever Hans también nos recuerda cuán estrechamente consideramos o reconocemos la inteligencia. A Hans se le enseñó a imitar tareas dentro de un rango muy acotado: sumar, restar y deletrear. Esto refleja una perspectiva limitada de lo que los caballos o los seres humanos pueden hacer. En términos de comunicación entre especies, Hans llevaba a cabo hazañas notables, además de actuar frente al público y tener una considerable paciencia; sin embargo, ninguna de estas acciones se reconocía como inteligencia. En palabras de la autora e ingeniera Ellen Ullman, esta creencia de que la mente es como un ordenador, y viceversa, ha “infectado décadas de pensamiento en las ciencias de la computación y cognitivas”, y se volvió una especie de pecado original en el campo.<sup>15</sup> Se trata de la ideología del dualismo cartesiano en la IA, en la que esta se limita a ser entendida como una inteligencia incorpórea, liberada de cualquier relación con el mundo material.

## ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL? NI ARTIFICIAL NI INTELIGENTE

Hagamos la pregunta, en apariencia simple, de qué es la IA. Si le preguntas a alguien en la calle, puede que mencione a Siri de Apple, la nube de Amazon, los coches de Tesla o el algoritmo de búsqueda de Google. Si les preguntas a los expertos en aprendizaje profundo, puede que te den una respuesta técnica acerca

<sup>15</sup> Ellen Ullman, *Life in Code. A Personal History of Technology*, Nueva York, MCD, 2017, pp. 136 y 137.