

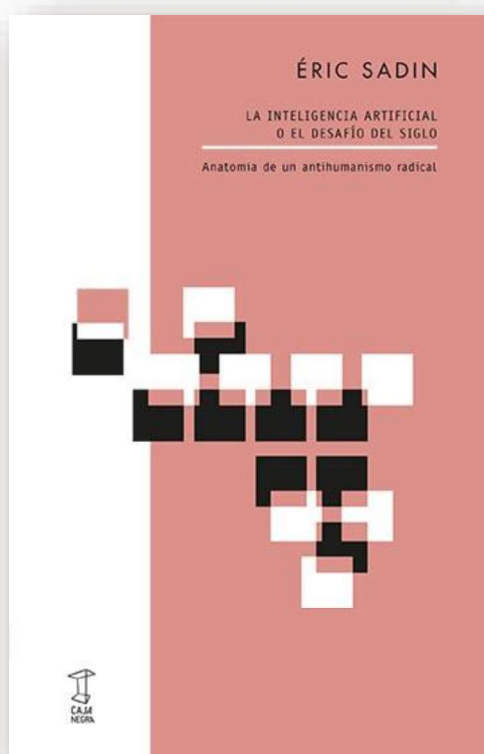
“La inteligencia artificial o el desafío del siglo: Anatomía de un antihumanismo radical”

Éric Sadin

Roberto Pizarro Contreras

Magíster en Filosofía

Universidad de Chile, [rpizarroc@hotmail.com](mailto:rpizarroc@hotmail.com)



Editorial Caja Negra  
Buenos Aires, 2020, 328 pp.  
ISBN: 978-987-1622-86-3

Un auténtico “leviatán algorítmico” es la entidad que el filósofo francés Éric Sadin, uno de los grandes nombres en la investigación de lo que ha venido a llamarse “subjetividad digital”, ve oculta detrás del popular fenómeno de la inteligencia artificial (IA). Pero a diferencia del leviatán de Hobbes, temible, este se presenta bajo la forma de ciudades “inteligentes” (*smart cities*), una “omnisciente” hiperconectividad a través del “internet de las cosas” (IoT), “oportunas” y “acertadas” sugerencias de plataformas de *streaming service*, como Netflix, y “amabilísimos” *digital assistants*, como Cortana de Microsoft o Alexa de Amazon, si bien la facultad de coaccionar, esta vez a escala planetaria, se

mantiene indemne (pp. 153-164).

“La inteligencia artificial o el desafío del siglo” responde, sí, a un intento de sentar las bases de una *ética contra la inteligencia artificial*, a la vez que se erige como el más político de los últimos tres libros del autor. “Político”, pues defiende a ultranza lo que puede entenderse como una versión, digamos,



“historicista” de una conciencia de raigambre humanista-moderna, además de centrar sus reflexiones en una crítica sobre las implicancias negativas de la IA más bien que balancear esto con sus potenciales beneficios, aunque el texto tiene el mérito de advertirnos con una claridad, que es a la vez filosófica y técnica, sobre el riesgo casi inminente del advenimiento de un régimen artificial donde los seres humanos pasan a convertirse en los apéndices nescientes de un sistema superinteligente que todo lo gobierna.

¿Cómo entiende la IA Sadin? En principio, se trata de un sistema global constituido por una miríada de dispositivos computacionales (*smartphones, chatbots, GPS, etc.*) conectados a los seres y las cosas del mundo (voces, textos, automóviles en movimiento, etc.), los que a su vez son reducidos a valores informáticos que facilitan su administración algorítmica. Esta última, a su vez, se realiza a través de secuencias de código que intentan imitar algún aspecto de la inteligencia humana, como lo es determinar el grado de felicidad a través de la captación biométrica de la pura fisonomía del rostro sin considerar otros hechos, o bien, sin tener en cuenta una forma humana no calculista de interpretarlos. Estas capacidades inferenciales, aunque parciales, son diseñadas para su continua expansión, sintetizando al cabo un modo de ser, una “racionalidad aumentada”, cuyo imperio eficiente y consecuencias están por verse, pese a que dejan ya entrever sus corolarios transformacionales sobre lo humano, o lo que el francés entiende como tal.

Uno de los aspectos nuevos que se aportan a la filosofía de la técnica de este pensador, expuesta previamente en *La humanidad aumentada* y *La silicolonización del mundo*, entre otros ensayos, es la intuición de dos características esenciales de la bestia informática, a saber, su *potencia aletheica* (pp. 17, 93) y su *poder-kairos* (pp. 237-246).

La primera mienta la capacidad de enunciar una “verdad” altamente persuasiva y su consiguiente *fuerza conminatoria* (pp. 47, 81-90). Ella es posible gracias al confort letárgico inoculado al usuario por el cumplimiento “exacto” de sus expectativas, lo que minimiza las decepciones, y por el refuerzo que a esto otorga la sacralización en nuestra cultura tecnocientifista del número. No obstante, contraria a la verdad abierta y en construcción imperecedera que intuyeran los filósofos desde la Antigüedad y máxime durante el criticismo ilustrado principiado por Immanuel Kant, la exactitud remite a una *falsa, complaciente y cada vez más absoluta verdad* (pp. 93-104).

La segunda, el *poder-kairos*, denuncia la facultad de la IA de responder oportuna a instantáneamente a cada evento de la realidad administrada, de un modo tal que no dé margen de

actuación ni crítica a los seres humanos en un sentido opuesto al ya pronosticado por sus lógicas internas, o bien, si lo hace, la desviación ha de ser mínima y tendiendo a cero conforme los algoritmos se iteran y perfeccionan a sí mismos en instancias posteriores según un esquema de *posprogramación* (p. 77) que prescinde de la supervisión humana.

A su vez la descomposición funcional de los atributos de “exactitud” y de “oportunidad e instantaneidad” hace posible una explicación mayor del éxito del fenómeno y la aprehensión de su fundamento económico-político, el capitalismo digital o *tecnoliberalismo*, que le ha propulsado. (pp. 191-202)

El rasgo nuclear aquí es quizá la *lógica utilitarista y omnisciente programada* en el sistema. Ella, herencia del liberalismo económico y concreción distópica del cálculo matemático de la felicidad que pretendía el fundador del utilitarismo moral, el filósofo Jeremy Bentham (p. 188), subsume cada elemento de la realidad numerada a un cálculo maximizador de las expectativas usuarias, es decir, a una sustracción entre beneficios y perjuicios, a un *trade-off* (por decirlo en la jerga empresarial).

Luego los cálculos mentados arriba se hacen sobre la base de *procedimientos de comparación y correlación perpetuos* entre los objetos gerenciados por el sistema (empleados, proveedores, mercancías, materias primas, etc.), de los que además se extraen relaciones imprevistas que escapan a la analítica humana y que optimizan su control. Su solución incita la movilidad de estos, también interminable, a lo largo y ancho de la economía digital, como supeditados a los hilos de la mano invisible de Adam Smith, reinterpretada ahora digitalmente por el espíritu mundializado del *entrepreneurship* de Silicon Valley (pp. 179-185). Todo esto converge, por último, en una capacidad de *autoaprendizaje emancipatorio*, encarnada por las técnicas de *machine y deep learning* (pp. 71-80).

Desde el punto de vista de la agencia social —y desde donde Sadin emprende su crítica y sienta las bases para una *deontología del agente de la técnica* y más generalmente para una *ética contra la inteligencia artificial*, como se previno al comienzo —, el funcionamiento de la IA se catapulta, idealmente, sobre una *interpretación de la IA como solución de comercialización y panacea de nuestro tiempo* (pp. 191-199), y, materialmente, sobre *la instauración de una burocracia o fábricas 4.0* (pp. 293-299).

En efecto, los emprendedores y otros *start-uppers* no solo ven en esta tecnología una forma de enriquecimiento económico, sino además una vía de realización (especialización) profesional y la solución tecnocrática a todos los problemas del mundo, lo que implica que ha devenido en ideología con manifestaciones extremas como la iglesia de la IA erigida por Anthony Levandowski, un ex

ejecutivo de Uber y experto en automoción autónoma (p. 103). Todo ello justificaría hoy la instalación de una versión electrónica de la jaula de Max Weber o, en palabras de Sadin, de unas mallas de silicio del leviatán hobbesiano, que automatizan la fuerza de trabajo por medio de una *asistencia robótica* (pp. 87-88), a partir de la cual ya no solo el recurso de la fuerza física, sino también el mental abdica en el de la máquina por ser esta capaz de procesar de manera mucho más exacta la vida matemáticamente reducida de unas organizaciones (empresas, organismos públicos, academias, etc.) que en el curso de su evolución ya habían devenido matemáticas también. De aquí en más el número manda, no las veleidades humanas.

La degeneración del espíritu científico-humanista es parte primordial de estos movimientos igualmente. Frente al tecno-orden instituido, científicos y humanistas se habrían enfrentado al consabido dilema de una *defensa impopular de sus más íntimas convicciones* o la *prostitución a regañadientes a favor de un capitalismo de nuevo cuño al objeto de lograr una sobrevivencia cómoda* (pp. 261-270). Sin embargo, esto, al hacerse con un grado de conciencia superior al de los agentes del Estado y otros tecnócratas consagrados a lo contingente, resulta tan contradictorio y reprochable como decidir, en circunstancias de que el autor es capaz de intuir que en la conformación de la IA la idea de erradicar el malestar al costo que fuera estuvo siempre a la zaga, incluso en los intelectuales, facilitando la edificación de una *racionalidad que marca distancia del cuerpo*, “artificial”, y que queda evidenciada en filósofos como Descartes y Kant, si se considera que practicaron el retiro para armar la trama de sus racionalismos; sobre todo el primero, que aspiró a apartar metódicamente de sí todo lo sensible-aparente hasta alcanzar su duda-certeza fundamental (pp. 253-256), que no habría sido más que una base nominal, una declaración de buena intención para su doctrina filosófica.

¿Qué podemos hacer entonces para reivindicarnos?

La solución ética de Sadin no parece muy convincente, no porque carezca de sistema (ya sabemos por él que largas y sesudas cadenas de razonamientos pueden estructurar maquinaciones voraces), sino porque propone un núcleo ético de *conformación y pregnancia molecular de la realidad*, idea que asegura recoger del pensamiento tradicional chino (p. 287), pero que sabe muy a pensamiento contemporáneo francés si se la correlaciona con la microfísica foucaultiana o la molecularidad o el rizoma de Deleuze y Guattari, entre otros, que se traduce en un esfuerzo sostenido y progresivo de resistencia política e intelectual de lo humano (p. 289). Es alentadora, por cuanto permite espabilar para los efectos técnicos que nos convocan aquí acerca de algunas soluciones archiprobadas a problemas tradicionales de la especie (el debate, el motín, etc.), pero demasiada abstracta aun y sesgada

quizá por un entrapamiento que se da en ciertas interpretaciones o síntesis filosóficas en las que se proponen alternativas equilibradas, como intentando aducir a un justo medio aristotélico, donde el punto pivote no es en modo alguno fácil de localizar, y de ahí que exégesis posteriores argumenten la adscripción del teórico a uno u otro polo que se intenta superar, descartando de este modo la novedad.

En concreto, Sadin propone repensar lo humano, considerando la tendencia histórica hacia una comodidad adormecedora en la certidumbre otorgada por una exactitud o falsa verdad y que, no obstante, todo nuestro devenir ha sido hasta acá consecuencia de “modos de racionalidad basados en la aceptación de la pluralidad de los seres y la incertidumbre fundamental de la vida” (p. 37), para decidir lo que queremos hacer en consecuencia. Algunas acciones específicas que plantea son *renegar de la asistencia digital individualizada* (pp. 293-296) tanto como sea posible y *abrir espacios de pensamiento* en las fábricas automatizadas o más probablemente fuera de ellas, así como promover *la educación y el espíritu crítico* (pp. 295-296), la evidencia de *contraexperticias* (p. 291) y *contraimaginarios* (p. 43), y *combatir el lenguaje tergiversado* (p. 292) y sus derivadas funestas (el ejemplo más emblemático es quizá la confusión cotidiana de equiparar la inteligencia humana a la artificial, ignorando que sus naturalezas y alcances son disímiles), que no hacen más que contribuir al desarrollo e intensificación de los sesgos tecnocráticos y a un *efecto Pígalión*, que en última instancia conducen a una ideologización y en casos extralimitados, como se dijo, a manifestaciones cultistas. Todo lo anterior, por supuesto, requiere igualmente de un *esfuerzo concertado* (de los gremios reivindicados en su razón de ser, por ejemplo) que implicará pérdidas en lo personal, pero que se verán abolidas en la medida que se tome conciencia de que lo que aquí está en juego es “el respeto incondicionado de la integridad y la dignidad humana” (p. 29).

Finalmente es de la máxima relevancia ética la asimetría que el filósofo establece en el epílogo al comparar al ser humano con el ser del pulpo en la voz de este último, puesto que exhorta con ello a tantear otras orientaciones de la racionalidad, que ya no se subsuman a las lógicas matemático-binarias de la inteligencia artificial, abriéndose en cambio a un espectro mucho más complejo y menos reductivo y sistemático, opción que, no obstante, difiere en apariencia con la defensa que hasta entonces parecía hacer del humanismo, si bien es posible que venga a esclarecer que de lo que se trata al cabo es de no perder la autonomía consciente, “de tener ‘el coraje de valernos de nuestro propio entendimiento’, para retomar la famosa fórmula de Kant, tan emblemática del Siglo de las Luces” (p. 275).