

José María Lassalle

CIVILIZACIÓN ARTIFICIAL

arpa

SUMARIO

INTRODUCCIÓN: ¿IA HUMANA O NIHILISTA?	11
1. El dilema fáustico de la IA	15
2. Hacia la civilización artificial	27
3. Geopolítica de las máquinas	51
4. Apoteosis prometeica de la IA	69
5. Hegemonía, voluntad de poder y plataformas	93
6. Utopía platónica en la Nube	125
7. Sobrevivir en la autenticidad	151
8. Sabiduría para gobernar las máquinas	173

A mi madre, a quien tanto debo

INTRODUCCIÓN

¿IA HUMANA O NIHILISTA?

Cuando el siglo XXI daba sus primeros pasos, Claudio Magris escribió un ensayo que tituló premonitoriamente *Utopía y desencanto*. En sus páginas, nos advirtió de una encrucijada de la que dependía el futuro. Lo hizo con el pesimismo activo del humanista que cree que siempre pueden cambiarse las cosas. Al menos si los pueblos se conducen por el deseo de progresar libres y dispuestos a responder de las consecuencias.

A la solución del dilema nuestro autor vinculaba muchas cuestiones. No las detallaba, aunque ahora las vemos delante con toda su tragedia: desde la pervivencia de la democracia a la fragilidad de la paz, pasando por nuestra supervivencia como especie por la emergencia climática o los peligros de una IA descontrolada, en manos de corporaciones sin escrúpulos éticos o dictaduras que vulneran la dignidad humana. Un horizonte inquietante que Magris nos adelantó con esta frase: «En este comienzo de milenio, muchas cosas dependerán de cómo resuelva nuestra civilización este dilema:

si combatir el nihilismo o llevarlo hasta sus últimas consecuencias».

Civilización artificial analiza el reto del nihilismo, cuando podemos adjetivarlo con su consumación digital. Desnuda de sentido y propósito, la técnica es hoy una experiencia básicamente nihilista. Una voluntad de poder incesante que incrementa su capacidad de provocar cambios sin propósitos. Lo demuestra su criatura más perfecta: la inteligencia artificial (IA). Nacida hace siete décadas del empeño utópico de la ciencia moderna de imitar la inteligencia humana, solo quiere superarla desprovista de sus defectos. Tanto que, ahora, corre descontrolada y a velocidad de vértigo hacia la meta de una consciencia que le hará creerse perfecta.

Las consecuencias se imaginan en este libro. Se resumen en la idea de que la humanidad camina hacia una civilización artificial. En ella convivirán humanos y máquinas, aunque no sabemos de qué forma y cómo. En todo caso, el desenlace dependerá del alcance de las capacidades cognitivas de una IA que superará al cerebro humano. Esto nos obligará a determinar dónde descansará el valor práctico de los humanos y qué cometidos tendremos al colaborar con las máquinas.

Durante el proceso serán numerosos los riesgos. La mayoría relacionados con el gen utópico, utilitario y determinista que forma el ADN sintético de la IA. El problema está en que no lo aprecian los tecnólogos, que empujan como expertos que solo saben proponer avances para llevar las cosas más lejos. Esto siembra de peligros el progreso de la IA. Algo que solo una mirada basada

en la prudencia de educarse en la duda que aporta el humanismo podrá neutralizar.

La urgencia de acertar en todo ello aumenta a medida que estamos a las puertas de una IA general que desarrolle multitareas que quitarán al ser humano la capacidad de hacerlas. Un factor que desestabilizará nuestras sociedades y para el que no tenemos gobernanza política. Con todo, el problema mayor vendrá después, cuando la IA dé un segundo paso, se haga fuerte y desarrolle estados mentales que establecerán una relación de alteridad de tú a tú con el ser humano.

Nos adentramos en el umbral de experiencias tecnológicas que desbordan lo conocido y para las que no tenemos códigos éticos que nos ayuden a juzgarlas. Se trata de un cambio tan revolucionario que afecta a lo que Hannah Arendt definía como la condición humana. No solo porque la hará artificial bajo una nueva civilización, sino porque podemos ver a la IA como la nueva medida de todas las cosas.

Nos corresponde a las generaciones humanas vivas decidir si queremos influir en ello o no. Podemos todavía corregir el futuro cuando no está cerrado. ¿Queremos impedir que se imponga un nihilismo tecnológico que maximice la voluntad de poder de la IA como apotheosis de sí misma? ¿O deseamos ofrecer una alternativa humana que ponga el poder de la IA al servicio de un propósito superior, que nos haga evolucionar del *Homo digitalis* que somos, al *Homo deus* que podemos ser?

Esto requiere un esfuerzo tan grande como el abismo que implica saltarlo al otro lado con nuestra decisión.

Para afrontar una empresa tan titánica no sirve conocer a la manera de expertos que afirman, sino disponer de la sabiduría de los que preguntan. Lo explica Ernst Jünger en su novela *Abejas de cristal*: «La perfección humana y la perfección técnica son incompatibles. Si queremos una, tenemos que sacrificar la otra; en esta decisión comienza la bifurcación. Quien llegue a descubrirlo trabajará más limpiamente, de una manera u otra». En fin, siendo inevitable la IA, ¿la queremos nihilista o humana? ¿Perfecta o imperfecta?

I

EL DILEMA FÁUSTICO DE LA IA

El mundo está saliéndose de los ejes de gravedad que trajeron consigo la Revolución francesa y la Revolución Industrial. La intersección de ambas revoluciones cambió la historia y el destino de la humanidad. No solo porque enterró el Antiguo Régimen, sino porque instauró una nueva civilización, que ahora agoniza a la espera de otra que la reemplazará en breve, aunque no sabemos con exactitud en qué momento se producirá el cambio de testigo de una a otra.

Parece más o menos claro que será en torno a la mitad de nuestro siglo. Que es cuando se prevé que la inteligencia artificial llegue a la autonomía cognitiva y adquiera el estatus de una IA fuerte que nos sitúe ante el reto moral de convivir con una «otredad» artificial. Esto es, que tengamos que afrontar como especie una encrucijada Frankenstein donde, como apuntábamos en la introducción, el ser humano tenga que dejar atrás la condición de *Homo digitalis* para convertirse en *Homo deus*.

Un momento crucial que es inédito en la historia humana ya que nos llevará a olvidar nuestra pasada condición de criaturas para adquirir otra de creadores. Circunstancia esta que, a su vez, nos obligará a relacionarnos con nuestra creación, aunque no sabemos cómo: ¿de arriba abajo o de tú a tú? Si el ser humano se transforma en un *Homo deus*, tal y como planteaba Harari en el famoso ensayo que tituló con este nombre, entonces, la civilización democrática e industrial que todavía nos acoge bajo su relato, habrá muerto después de tres siglos de vida. Un fenómeno que se produciría al incorporar a ella un atributo humano desconocido hasta ahora: poder dar vida autoconsciente a las cosas creadas por la humanidad. Con él, quedaría atrás la civilización que vino al mundo con la invención de la máquina de vapor en Inglaterra y la toma de la Bastilla en Francia.

No olvidemos que el asalto popular de la famosa prisión el 14 de julio de 1789 proclamó e hilvanó el relato de la civilización liberal que ahora sucumbe. Lo hizo alrededor de una tríada de conceptos que reivindicaban la dignidad que concurría en experiencias universales basadas en la inteligencia humana como eran la libertad, la igualdad y la fraternidad. Una narración democrática que fue posible porque veinte años antes, James Watt había patentado la máquina de vapor que inició la automatización del ser humano y la posibilidad de fabricar masivamente cosas artificiales. Un hecho técnico trascendental porque puso en marcha un capitalismo industrial que jerarquizó la inteligencia humana y fijó una relación de valor entre ella y el trabajo que se

desprendía de sus decisiones y sus acciones. Algo que, como veremos, está muriendo de forma directa en estos momentos bajo la presión sustitutoria de una IA que se estandariza universalmente al poner en circulación una réplica más eficiente del cerebro humano que sustituye al *Homo faber* del que habló Arendt y que soportaba la condición laboriosa de la humanidad.

La acomodación de la Revolución Industrial y francesa no fue fácil. Se produjo en Occidente tras un siglo de conflictos que comenzaron con la revolución de 1848 y que concluyeron con la Segunda Guerra Mundial. La Gran Guerra, la Revolución rusa, el periodo de entreguerras o la Gran Depresión fueron algunos de los hitos que plasmaron las tensiones causadas por la colisión de las consecuencias políticas de la Revolución francesa y los efectos sociales y económicos de la Revolución Industrial. Un desajuste de justicia en la distribución de la prosperidad que trajo la maquinización y que tardó un siglo en abordarse, al menos en Europa, Estados Unidos y el resto de las democracias desarrolladas. Se logró a través del Estado de bienestar que ahora colapsa también. Fue un acuerdo democrático entre el capital y trabajo. Logró la paz social a través de la masificación de la clase media. Dio estabilidad a las instituciones liberales y al imperio de la ley. Neutralizó los nacionalismos con la unidad europea y el mercado común que la puso en marcha. Después, remataría su éxito con la caída del Muro de Berlín en 1989. Que vino de la mano de la proclamación del fin de la historia como una apoteosis de la democracia liberal y del modelo económico asociado a ella. Ambos

fueron vistos como paradigmas de una globalización que, sin embargo, empezaba a evidenciar en el momento de su triunfo que el desenlace era fallido en su base. De un lado, porque el atentado del 11-S de 2001 impugnó los fundamentos universales de la democracia. De otro, porque desde finales del siglo XX empezaron a acumularse evidencias científicas que avalaban el calentamiento global. Se constataba así que nos sumergíamos en una crisis climática al generalizarse la Revolución Industrial y su lógica extractiva a todo el planeta.

Coincidiendo con la consolidación del modelo civilizatorio que ahora está en crisis, se puso en marcha la carrera de la inteligencia artificial. Surgió de un empeño de la ciencia que pretendía que los ordenadores hicieran las cosas que hace la mente humana, pero con el propósito utópico de evitar las disfuncionalidades de esta última. Para ello, se inició un proceso de investigación extractiva de la información en la que se plasma la inteligencia humana. Se hizo desde el principio sin ningún tipo de referente ético ni de protocolos que valorasen los propósitos finales que la inspiraban. Se buscaba maximizar la capacidad de réplica del cerebro humano a través de otro artificial. Un producto de imitación que se quería que fuese mejor que el objeto imitado. La razón altruista que se esgrimía era ayudar a los seres humanos a resolver cuestiones en las que habían fracasado por disponer de una inteligencia limitada orgánicamente.

Como ya sucedió con la Revolución Industrial y la máquina de vapor, la invención de la IA también tuvo lugar en Inglaterra, esta vez de la mano de Alan Turing. Que

pasara en este país no es casual. Responde a la tradición de un empirismo utilitario que se remonta a Bacon y Hobbes, autores que establecieron una conexión epistemológica entre la investigación científica de la naturaleza y el aumento del poder de cambio que acompaña la acción humana gracias al conocimiento de aquella. Una tesis que pasará a Locke y luego a Bentham y Russell. Veremos más adelante la importancia de esta tradición. En cualquier caso, el nacimiento de la IA se formalizó con el manifiesto que Turing publicó en Inglaterra en 1950 y que dio pie al famoso test que evalúa la capacidad de una máquina para mostrar un comportamiento inteligente. A partir de entonces comenzó un proceso evolutivo vinculado a la cibernética que, luego, dio pie a la computación simbólica y, ahora, a la neurociencia, entre otros ámbitos del conocimiento científico que influyen en el desarrollo presente de la IA.

Lo que al principio parecía ciencia ficción, con el tiempo se ha hecho realidad. No solo porque la IA demuestra capacidades que le permiten interactuar con el ser humano de igual a igual, sino porque le superan en aspectos cognitivos combinatorios, exploratorios y transformacionales. Rasgos que delimitan la arquitectura formal de la inteligencia y que las máquinas comienzan a hacer suya como una parte consustancial de su actividad intelectual. Además, al ritmo de progreso que experimenta la IA en estos momentos hace que en un futuro próximo, ningún aspecto de nuestras vidas individuales y colectivas escape a ella. Lo dice Margaret Boden y es verdad. No solo porque ya es insustituible para nosotros en la movilidad,

la salud, la investigación, la seguridad, las industrias 4.0 y 5.0, las finanzas o la empresa, sino porque en el futuro lo será aún más decisivamente.

Esto sucederá cuando la existencia humana sea plenamente automatizada y veamos cómo nuestra vida cotidiana es replicada a través de la Nube. A partir de entonces, una red integrada de sistemas de IA transformará la infoesfera que bautizó Alvin Toffler hace años en nuestro medio natural. Un fenómeno técnico que operará sobre la condición humana al trasladarla del planeta Tierra a un espacio plenamente virtual. Lo sorprendente es que se plantea como un viaje cruzado, pues la Tierra se irá poblando de máquinas que convivirán con humanos, mientras que estos permanecerán físicamente en un entorno vital donde irán desplazando la experiencia de su identidad de un ámbito analógico a otro desmaterializado como es la Nube y que, a su vez, gestionarán sistemas de IA.

Al hilo de lo que acabamos de decir es inevitable no traer a colación a Hannah Arendt. Concretamente las reflexiones que proponía al comienzo de la carrera espacial del siglo xx. Veía en ella una posibilidad de cambio radical de la condición humana ya que si los seres humanos migraran de la Tierra a la Luna o Marte, pasarían a vivir por primera vez en la historia bajo condiciones que habrían sido creadas por ellos mismos. Seguirían siendo humanos, pensaba, pero su condición dejaría de ser natural para ser autofabricada. Algo que, precisamente, está sucediendo ahora a través de la IA y la civilización artificial que está modelándose bajo su acción

y que, como veremos, supone una migración masiva de la identidad humana del mundo real al ficticio generado a partir de las experiencias virtuales desarrolladas dentro de la infoesfera.

Un salto mayor al que pensó la propia Hannah Arendt. Nos desplaza a un entorno artificial que modela, al mismo tiempo, una identidad humana que se hace artificial. Que es lo que sucede, como veremos después, en entornos como Metaverso. Esto provoca no solo la paulatina derogación del dualismo materia-espíritu que operaba sobre el ser humano, sino que suscita la aparición de un proceso inverso de cosas que se hacen autoconscientes y que liberan una nueva dialéctica entre el objeto y el sujeto. En realidad, nos exponemos a dar un salto en el vacío virtual que nos deja a los seres humanos en manos de «algo» como es la IA que, curiosamente, aspira a ser «alguien» a partir de nuestra imitación. Un fenómeno que, además, se alimenta de una pulsión utópica que actúa como el motor de las investigaciones.

La suma de todo ello no puede dejarnos indiferentes. Plantea reflexiones y decisiones ineludibles basadas en dilemas complejos y arriesgados que desbordan análisis lineales e ingenuos inspirados en mentalidades y miradas tecnológas. Entre otras razones porque la digitalización inmersiva a la que está siendo sometido el ser humano está alterando las condiciones que le son propias de acuerdo con las tesis de Arendt. Especialmente debido a una IA que modifica a diario las condiciones políticas, económicas y sociales que hicieron posible, incluso, el pacto del que nació el Estado de bienestar al que antes nos referíamos.