

¿INTELIGENCIA ARTIFICIAL COMUNICATIVA?

Adela Cortina

Catedrática Emérita de Ética y Filosofía Política

Universidad de Valencia

1. Un rótulo septuagenario.

El rótulo "inteligencia artificial", creado por John McCarthy, al hilo de la célebre conferencia de Darmouth sobre máquinas pensantes, cumple 70 años en 2026, ya es septuagenario. En este tiempo ha atraído la atención de la opinión pública diariamente a través de congresos, publicaciones, encuentros y noticias impactantes. Entre otras cosas, porque ya ponen buen cuidado las grandes empresas tecnológicas en mantener el fuego sagrado del interés por la IA anunciando innovaciones asombrosas para atraer al público en este tiempo nuestro de economía de la atención.

Ya desde el nacimiento de la IA se perfilaron al menos tres actitudes frente a ella: según los catastrofistas, la invención de la IA es lo peor que le ha pasado a la humanidad; los entusiastas llegaron a prometer para pasado mañana un paraíso sin muerte, enfermedad ni siquiera envejecimiento; el sueño de la inmortalidad y la eterna juventud; los prudentes, por su parte, reclamaron el marco de una ética local y global capaz de maximizar las ventajas de la IA para todos los seres humanos y para la naturaleza. ¿Cómo situarse ante estas alternativas?

En estos años el proyecto ha cobrado unas dimensiones que difícilmente pudieron imaginar sus creadores. Goza de un protagonismo innegable en el orden económico, político, filosófico y social en el nivel mundial y en la competición por el poder geoestratégico y geopolítico. Domina también la opinión pública de formas muy diversas. No hay día en que no aparezcan noticias llamativas sobre sus logros en los medios de comunicación.

Pero a la vez han ido surgiendo críticas, a menudo muy aceradas, sobre los sacrificios humanos y ambientales que exige su producción y sobre las consecuencias éticas y políticas de su funcionamiento. Como muestra, me permitiré citar un texto muy duro del *Atlas de la IA* de Kate Crawford, libro al que acompaña el rótulo "Poder, política y costes planetarios de la inteligencia artificial". El texto es un poco extenso, pero resulta muy significativo:

"(...) en este libro sostengo -escribe la autora- que la IA no es *artificial* ni *inteligente*. Más bien existe de forma corpórea, como algo material, hecho de recursos naturales, combustible, mano de obra, infraestructuras, logística, historias y clasificaciones. Los sistemas de IA no son autónomos, racionales, ni capaces de discernir algo sin un entrenamiento extenso y computacionalmente intensivo, con enormes conjuntos de datos o reglas y recompensas predefinidas. De hecho, la IA, como la conocemos, depende completamente de un conjunto mucho más vasto de estructuras políticas y sociales. Y, debido al capital que se necesita para construir IA a gran escala y a las maneras de ver que optimiza, los sistemas de IA son, al fin y al cabo, diseñados para servir a intereses dominantes ya existentes. En este sentido, la IA es un certificado del poder"¹.

La sospecha de que la estructura de la IA revela de nuevo que la ciencia y la técnica -ahora las tecnociencias- se convierten una vez más en ideología y certifican de nuevo el triunfo de la razón instrumental es inevitable. En el pasado siglo los francfortianos denunciaron el triunfo de la razón instrumental, que entiende de medios, pero es incapaz de reconocer unos fines comunes en los que los seres humanos podamos entendernos, alcanzar una *Verständigung*. De ahí que en *¿Ética o ideología de la inteligencia artificial?* intentara poner sobre el tapete el hecho de que en el siglo XXI se produce el eclipse de la razón comunicativa,

¹ Kate Crawford, *Atlas de IA. Poder política y costes planetarios de la inteligencia artificial*, NED Ediciones, 2023, 29.

que es la que permite potenciar la intersubjetividad humana (la relación sujeto-sujeto), único camino para formar sociedades democráticas.

Sin embargo, a mi juicio, la crítica de Crawford no da en la diana. Que la IA necesite para construirse medios materiales, como el esfuerzo de millones de trabajadores, el casi incalculable gasto energético, los diseños humanos de la logística imprescindibles para producirla, no significa que ella misma sea corpórea, que sea material. De hecho, en los últimos tiempos se han multiplicado exponencialmente las críticas al modo de producir la IA, como es el caso paradigmático del trabajo ya clásico de Shoshana Zuboff *La era del capitalismo de la vigilancia* o, entre nosotros, el libro de Ramón López de Mántaras, *100 cosas que hay que saber sobre inteligencia artificial*. Consideran con buen acuerdo estos expertos que, junto a los innegables beneficios de la IA, que forma ya parte de nuestro modo de vida, es preciso sacar a la luz y acabar con realidades intolerables².

La *forma de producir la IA* representa un innegable problema ético y político, que es necesario abordar con urgencia, y así se está haciendo desde múltiples textos e intervenciones públicas³, pero no dice nada de la materialidad de la IA misma, no niega que sea artificial. Y en cuanto a que la inteligencia de la IA sea la de todos los humanos que la construyen con su trabajo es un tema muy discutible, del que hablaremos más adelante. Sí que es cierto, en cualquier caso, que es un certificado del poder, pero ése es un problema que debe abordarse desde otra perspectiva.

Desde esa perspectiva convenía, a mi juicio, destacar la excepcionalidad de las personas humanas en el sentido de que gozan de unas capacidades que las hacen únicas en el mundo conocido y responsables de ese mundo. Una de esas capacidades es la *autonomía*, por la que reconocemos que los

² Shoshana Zuboff, *La era del capitalismo de la vigilancia*, Paidós, Barcelona, 2020; Ramón López de Mántaras, *100 cosas que hay que saber sobre inteligencia artificial*, Lectio Ediciones, Barcelona, 2024.

³ León XIV, *Magnífica Humanitas, Sobre la custodia de la persona humana en el tiempo de la inteligencia artificial*, Carta Encíclica, 2026.

seres humanos tienen dignidad y no un simple precio. El surgimiento de la IA había llevado a utilizar expresiones como "personas electrónicas", "vehículos autónomos", "sistemas de armas autónomos" y era preciso preguntarse si, efectivamente, los sistemas inteligentes pueden considerarse autónomos y, en tal caso, tienen que ser tratados con consideración y respeto como ocurre con los seres humanos.

Ante esta cuestión un buen número de autores brindó la espléndida propuesta de entender que esas máquinas gozan de *autonomía funcional*, no ontológica ni ética⁴. Funcionan como si fueran autónomas. Mientras que los seres humanos gozan de autonomía ontológica y ética, de una agencia moral plena, los llamados sistemas "inteligentes" gozan de una autonomía funcional y no son sino *instrumentos*⁵.

Existen razones suficientes, a mi juicio, para sostener esa afirmación, y dediqué parte de mi última intervención en esta academia a intentar mostrarlo. Llegué allí a la conclusión de que la razón más contundente es la siguiente: la diferencia esencial entre *autonomía ontológica y ética* y autonomía sólo *funcional* se descubre a través del diálogo, porque es en el diálogo cuando los sistemas supuestamente inteligentes deberían poder asumir también la *perspectiva del participante* en ese diálogo, y no sólo la perspectiva del observador.

Pero en ese caso deberían intervenir asumiendo el lenguaje de primera persona, no el lenguaje objetivador de tercera persona. Con ello revelarían a la vez la existencia de la *intersubjetividad*, el *vínculo con los demás sujetos*, que hace a la vez posible el mundo moral. Sin embargo, para las máquinas esto es imposible, porque, efectivamente, no son ontológicamente autónomas. Y, por si faltara poco, sin la *intersubjetividad* del acto de comprender no hay tampoco *objetividad* del saber, como acertadamente reconoce Habermas

⁴ Wendell Wallach y Allen Colin, *Moral Machines: Teaching Robots Right from Wrong*, Oxford University Press, Nueva York, 2008. P. 39.

⁵ Adela Cortina, *¿Ética o ideología de la inteligencia artificial?*, pgs. 102- 106.

en distintos trabajos y muy especialmente en *Entre naturalismo y religión*⁶.

2. ¿Superando el logocentrismo?

Sin embargo, algunas voces aseguran que existe inteligencia artificial comunicativa, incluso que podría ser la antesala de la Inteligencia Artificial general, capaz de resolver problemas generales, no sólo especiales. Y, como consecuencia, anuncian con entusiasmo el fin del logocentrismo humano.

A mi juicio, como la meta de la IA como disciplina consiste en crear una inteligencia artificial como la humana, no es de extrañar que los antropomorfismos campeen a sus anchas en los textos, pero no deja de ser llamativo el afán de muchos autores por intentar poner fin al antropocentrismo, mostrando cómo algunas de las capacidades que se tenían por eminentemente humanas se asignan no sólo a los animales, o a ciertos accidentes geográficos (ríos, lagos, montes), y también a la tierra, si es que hay que atender a la Hipótesis Gaia. Ahora también las máquinas gozarían de algunas de esas capacidades tradicionalmente tenidas por humanas, con lo cual, en este caso, se desdibuja la distinción “humano-máquina”.

Un ejemplo de ello es el texto de Coeckelbergh y Gunkel *Inteligencia artificial comunicativa*, en el que decretan con entusiasmo la defunción del logocentrismo⁷. Evidentemente, los autores recuerdan -como es de rigor- el espléndido pasaje del libro I de la *Política* de Aristóteles,

“La razón por la cual el hombre es, más que la abeja o cualquier animal gregario, un animal social es evidente: la naturaleza, como solemos decir, no hace nada en vano, y el hombre es el único animal que tiene

⁶ Jürgen Habermas, *Entre naturalismo y religión*, Paidós, Barcelona, 2006, 178-187; Adela Cortina, *Neuroética y neuropolítica*, Tecnos, Madrid, 2011, 195-198.

⁷ Mark Coeckelbergh, David J. Gunkel. *Communicative AI. A Critical Introduction To Large Language Models*, Polity Press, Cambridge 2025.

palabra (lógos). (...) La palabra es para manifestar lo conveniente y lo dañino, lo justo y lo injusto, y es exclusivo del hombre frente a los demás animales, el tener, él sólo, el sentido del bien y del mal, de lo justo y lo injusto, etc., y la comunidad de estas cosas es lo que constituye la casa y la ciudad”⁸.

*Mientras los animales tienen voz para significar el dolor y el placer, el hombre puede deliberar sobre lo conveniente y lo dañino, sobre lo justo y lo injusto. Y en esto consiste la casa, el oikós, es decir, la vida de la familia y la amistad. Pero también en esto consiste la pólis, la comunidad política, que congrega distintas familias y diversas etnias, y se distingue de unas y otras porque tiende por naturaleza al bien común, y debería esforzarse por alcanzarlo. La vida privada y la pública tienen su raíz en el lógos. El hombre es, por decirlo con el acertado título del libro de nuestro compañero Diego Gracia, *El animal deliberante*, precisamente por estar dotado de palabra*⁹.

Y no deja de ser interesante que un buen número de autores, cuando se preguntan cómo tomar decisiones a la hora de diseñar un marco global de gobernanza de la IA, porque sus consecuencias son globales, acaben sugiriendo que el procedimiento más adecuado sería el de una democracia deliberativa, propiciada por Naciones Unidas, en la que deberían participar los ciudadanos con el asesoramiento de los expertos. A fin de cuentas, los afectados por las tecnociencias deberían ser quienes decidan hacia dónde deben dirigirse, y no un grupo reducido de tecnoempresas poderosas, que han llevado a hablar de tecnofeudalismo, e incluso de tecnocolonialismo, aunque ésta sea, a mi juicio, una expresión muy inadecuada. Sin embargo, estas propuestas, que debería liderar Naciones Unidas, no cuentan con incluir entre los sujetos deliberantes a una supuesta IA comunicativa.

⁸ Aristóteles, *Política*, I, 2, 14-18, Instituto de Estudios Políticos, Madrid, 1970.

⁹ Diego Gracia, *El animal deliberante. Teoría y práctica de la deliberación moral*, Triacastela, Madrid, 2025.

Con todo, los autores mencionados aseguran que contamos con una IA comunicativa desde el nacimiento de ChatGPT a fines de noviembre de 2021, lanzado por Open AI, como una interfaz conversacional, diseñada para el procesamiento del lenguaje natural, basado en los Modelos Extensos de Lenguaje. Como es sabido, los modelos Extensos o modelos grandes de Lenguaje son sistemas de IA, entrenados con enormes cantidades de texto, diseñados para procesar y generar lenguaje humano, mediante modelos de aprendizaje profundo. Los LLMs producen interacción conversacional y son instrumentos para escritura, investigación y otras formas de comunicación¹⁰.

De ahí que -continúan estos autores- ChatGPT no haya sido sólo una maravilla tecnológica, sino que supone un cambio de paradigma, que sacude los fundamentos de las humanidades y amenaza, si no destruye, nuestro sentido de auto-identidad y excepcionalismo. Los seres humanos pierden la exclusividad de la posibilidad de comunicar hasta el punto de que Coeckelbergh y Gunkel anuncian con entusiasmo la muerte del logocentrismo y la representan en su libro con una lápida, coronada por las bien conocidas siglas funerarias "RIP". ¿Cómo vivir -se preguntan- en un mundo en que las máquinas nos hablan?

Este reconocimiento tendría consecuencias de diverso tipo, epistemológicas, entre otras. La distinción entre persona y cosa -se dice- ha tenido vigor durante 2.000 años, pero los Modelos Extensos de Lenguaje difícilmente se encuadran en una de las dos categorías. No son personas, pero tampoco son objetos como los otros, ya que pueden usar el lenguaje de una forma que resulta indistinguible del uso que hacen de él los seres humanos. Los Modelos Extensos de Lenguaje no son objetos como los otros, porque nos hablan¹¹

¹⁰ Coeckelbergh / Gunkel, op.cit., 1. Los Modelos Extensos de Lenguaje consumen una gran cantidad de energía y de trabajo humano, pero resultan muy rentables para las empresas y para los políticos.

¹¹ Ibid, 56.

y con esta innovación el panorama de los saberes sufre una transformación.

3. Repercusiones epistemológicas.

Aplicando estas reflexiones al modo de concebir la relación entre las tres formas de saber que posibilitan el progreso del saber humano, según la Teoría de la Evolución Social de Habermas¹², se difuminarían los perfiles de esas formas de saber, al menos de las dos primeras.

1) El *saber científico técnico*, impulsado por el interés en dominar la naturaleza externa e interna, que recurre a la distinción "sujeto-objeto" y se propone alcanzar un conocimiento objetivo, supone una distinción epistemológica, porque el sujeto ontológico puede devenir epistemológicamente objeto de sí mismo. ¿Una máquina pensante puede devenir objeto de sí misma? Si es realmente pensante, sí, porque entonces tiene la capacidad de desdoblarse. Pero quienes hablan de IA comunicativa utilizan un lenguaje ambiguo en este punto y emplean un doble discurso. Por una parte, se expresan dubitativamente con respecto al tema y dicen en algún momento: "hoy las máquinas parecen poseer lógos también"¹³, pero, por otra parte aseguran que "los modelos extensos de lenguaje, no son objetos como los otros porque nos hablan"¹⁴. Si esto fuera así, debería tener consecuencias morales y legales: ¿cómo deberían ser tratados?

Es verdad que los modelos extensos de lenguaje pasan con mucho el test de Turing, que, como es sabido, es una prueba para medir si una máquina puede exhibir un comportamiento inteligente, indistinguible del de un ser humano inteligente. Pero las opiniones sobre cómo interpretar este hecho difieren. Algunos autores se preguntan si esto es una prueba de que son conscientes o de

¹² Jürgen Habermas, "Conocimiento e interés", en, *Ciencia y técnica como ideología*, Tecnos, Madrid, 1984, pags, 159-181.

¹³ Mark Coeckelbergh, David J. Gunkel, oc., p. 2.

que son signos de inteligencia¹⁵. Otros dicen que es el comienzo de la IA general¹⁶. Mientras que otros aseguran que es un loro estocástico¹⁷. Como es sabido, los estocásticos son sistemas de IA que utilizan relaciones estadísticas, aprendidas a partir de conjuntos de datos masivos, para generar texto similar al humano de manera convincente, pero carecen de verdadera comprensión semántica detrás de los patrones de palabras, funcionan aleatoriamente. Ocurre, sin embargo, que cuando pronuncian una palabra inesperada parece que sean creativos.

Evidentemente, en el caso de que se tratase de los estocásticos, los modelos extensos de lenguaje no serían seres ontológicamente autónomos, y, por tanto, no serían responsables de sus expresiones lingüísticas. Como bien dice Javier Sampedro, "Los LLM basan sus resultados en asociaciones léxicas y principios estadísticos, y no en razonamientos contextuales ni criterios normativos. Eso es lo que llaman 'epistemia', un modelo que genera una ilusión de conocimiento sin tenerlo realmente. Parece que sabe, pero no sabe"¹⁸.

El problema de fondo reside, a mi juicio, en que el sintagma "*inteligencia artificial*" es desafortunado, cuando menos. Sí que es artificial, pero no es inteligencia. Para serlo necesitaría un cuerpo biológico para situarse en el mundo, comprenderlo desde un bagaje cultural y tomar posición ante él.

Con el proyecto de las máquinas pensantes se trataba de construir máquinas que se comporten de un modo que llamaríamos inteligente si un ser humano inteligente se comportara de ese modo. Se trata, pues, de simulaciones, de que se comporten "como si" fueran inteligentes, aunque no lo sean realmente. La filosofía del "como si" planea desde el

¹⁵ Anil Seth, *La mitología de la IA consciente*, Noema, 14 de enero. 2026.

¹⁶ AGI/Dubeck et alii, 2023, citado por Coeckelberg op. cit. p. 471.

¹⁷ José Luis Mendivil, "Un loro estocástico en la habitación china: ¿qué nos enseña ChatGPT sobre la conducta humana?", *Letras Libres*, nº 262, 2023, pags. 1-57.

¹⁸ Javier Sampedro, *El País*, 1 de noviembre, 2025, p.13.

comienzo y, sin embargo, se emplea continuamente un lenguaje antropomórfico que no hace sino confundir los perfiles. Se dice que hay coches autónomos, sistemas de armas autónomos, que los algoritmos llegan a decisiones, que ChatGPT entabla conversaciones con interlocutores humanos, que les da consejos y buenas sugerencias, incluso mucho mejores que los amigos y consultores. Hasta el punto de que se corre el riesgo de debilitar las relaciones humanas¹⁹.

Y, sin embargo, tiene razón Floridi cuando afirma que la llamada inteligencia artificial es una forma de *agencia*, pero no de *inteligencia*²⁰. ChatGPT es una herramienta sofisticada y útil, pero los modelos extensos de lenguaje se basan en calcular las probabilidades que tiene una palabra de aparecer después de otra. El programa valora qué palabra escribir a continuación, teniendo en cuenta las ya empleadas. A veces no usa la más probable y eso da sensación de originalidad. Pero ChatGPT no entiende nada de lo que dice ni tiene la menor intención comunicativa²¹.

"Agencia" sería la capacidad de una entidad de actuar por sí misma, a diferencia de los seres inertes. El sujeto agente en este caso ha sido programado en un inicio, pero después actúa sin necesidad de la intervención de terceros, y puede modificar el programa inicial teniendo en cuenta que hablamos de máquinas que aprenden, aunque carezcan de conciencia y sentido moral. Y esta distinción es de la mayor trascendencia, entre otras razones, porque el coche o la máquina letal "inteligentes" no son *responsables* de las actuaciones, y por eso importa construirlos con pautas que respeten los códigos éticos valiosos. Los robots no son agentes morales.

De ahí que pueda decirse con Floridi que la IA es una forma de *agencia* sin *inteligencia*, que *provoca una revolución asombrosa, pero en un sentido pragmático y no cognitivo*. Y

¹⁹ Adela Cortina, "¿Deshumanizarse en la era de la IA?, *Revista de Occidente*, n° 534, (2025), 97-103.

²⁰ Luciano Floridi, *Ética de la inteligencia artificial*, Herder, Barcelona, 2024, pp. 22 y 23.

²¹ Adela Cortina. *¿Ética o ideología de la inteligencia artificial?*.

que la filosofía de la IA se centra en la reconstrucción de una inteligencia de tipo biológico, pero debería de hacerlo en la configuración de artefactos, que puedan funcionar con éxito sin necesidad de inteligencia. La inteligencia no consiste tanto en resolver problemas de forma ingenieril como en decidir qué problemas merece la pena resolver.

2) *Saber práctico-hermenéutico*, impulsado por el interés en el entendimiento mutuo, en la *Verständigung*, desde el esquema sujeto-sujeto, desde la intersubjetividad desde la que los seres humanos entendemos y construimos el mundo. El positivismo ha negado racionalidad a este tipo de saber sobre todo por no venir legitimado por la experiencia y, sin embargo, como hemos dicho, sin la racionalidad del saber intersubjetivo tampoco puede haber saber objetivo. Son dos tipos de saber que han crecido conjuntamente en el proceso de la evolución²².

Algunos autores entienden que con las inteligencias artificiales se ha ampliado el número de sujetos, de interlocutores, como es el caso paradigmático de Mo Gawdat, que considera discriminatorio negar personalidad a las máquinas porque no se sustentan en un cuerpo biológico²³. En el caso de los modelos extensos de lenguaje puede decirse que no son meros objetos, porque nos hablan, y hay que dispensarles un trato especial.

Aunque estas cuestiones están en discusión, la balanza se inclina hacia la convicción de que no hay base suficiente para afirmar que los modelos extensos de lenguaje tienen conciencia, intencionalidad y son capaces de reciprocidad. No pueden ser tratados como interlocutores válidos ni reconocidos como personas. De ahí que seamos sólo los seres humanos quienes hemos de decidir hacia dónde queremos dirigirnos. Los llamados "sistemas inteligentes" son instrumentos muy útiles sin los que ya no sabríamos vivir. De hecho, un autóctono digital no puede concebir cómo se

²² Adela Cortina, *Ética cosmopolita. Una apuesta por la cordura en tiempos de pandemia*, Barcelona, Paidós, cap. 7: Humanidades y tecnociencias. Juegos de suma positiva.

²³ Mo Gawdat, *La inteligencia que asusta*, Paidós, Barcelona, 2024.

puede vivir en un mundo sin IA. Pero no dejan de ser instrumentos. Su aparición no cambia el perfil del saber práctico-hermenéutico.

Y en cuanto a la distinción sujeto-objeto, que lleva a preguntarse por el estatuto intermedio de los LLMs, Cockelbergh y Gunkel plantean la cuestión con las siguientes palabras:

“Pero los LLMs no son objetos como los otros, porque nos hablan. Puede usar el lenguaje de formas que son indistinguibles del uso que hacen los seres humanos. La distinción persona-cosa ha valido por 2000 años, pero los LLMs difícilmente se encuadran en una de las dos categorías”²⁴

Sin embargo, esta difícil distinción entre sujetos y objetos, que comporta consecuencias éticas con respecto al trato que se les debe dispensar, se difuminó hace ya tiempo en el caso de los animales y en el de entidades naturales, como lagos y ríos, a los que se reconocen derechos en constituciones latinoamericanas²⁵. En cuanto a los LLMs, todavía no se les ha reconocido jurídicamente ningún estatuto especial. Lo que sí es cierto es que ellos mismos plantean una gran cantidad de problemas éticos y jurídicos

3) *En el caso del saber emancipatorio*, que orienta a los otros dos y se esfuerza por construir sociedades no ideologizadas, siguen siendo los seres humanos quienes deben liderarlo. Sólo puede hacerlo una inteligencia humana, con emociones y sentimientos, que quiere liberarse de ideologías como la que han creado quienes prometen una IA general sin base científica suficiente, pero con afán de poder económico y político.

De ahí que en este caso se planteen preguntas tan relevantes como la siguiente: ¿estas tecnologías intentan satisfacer necesidades de los seres humanos, como debería

²⁴ Op.cit, p. 56.

²⁵ Adela Cortina, *Las fronteras de la persona. El valor de los animales, la dignidad de los humanos*, Taurus, Madrid, 2009, cap. 8.

ser propio de las tecnociencias, o más bien sus creadores intentan configurar una infoesfera, es decir, un ecosistema de creación y relación informacional, donde lo digital y lo físico se entrelazan, de modo que se desdibuja la diferencia entre lo humano y lo artificial?

Es verdad que la tecnología siempre ha modelado de una forma especial la vida humana, es verdad que es culturógena, lo quiera o no, pero cada vez más se muestra que no es un medio para satisfacer las necesidades y deseos humanos, sino un medio de moldear nuestras necesidades y deseos de forma que el entorno sea favorable a las tecnologías que interesa promocionar con fines comerciales y políticos. De ahí que no pueda decirse que las tecnociencias son ética y políticamente neutrales y que la ética de sus productos depende del fin para el que se use, porque siempre es de alguna manera política²⁶.

4. ¿Del olvido del ser al olvido del lógos?

Las nuevas "filosofías de la sospecha" ponen en cuestión el valor de la palabra para la vida pública porque la consideran como un instrumento del poder y, por lo tanto, que resulta imprescindible deconstruir los discursos para sacar a la luz las relaciones de poder entrañadas en ella desde una teoría de la justicia social crítica, entendida en sentido amplio. Pero entonces puede ocurrir que de la misma manera que en su tiempo Heidegger denunciaba el olvido del ser en la historia, hoy en día sea urgente denunciar el olvido del logos, por decirlo con Karl-Otto Apel. E incluso denunciar la "logofobia", como muy bien dice Jesús Conill.

¿A quiénes perjudica la fobia hacia el *lógos*, entendido como simbiosis de razón y emociones? Normalmente la pregunta que se formula en estos casos es ¿"cui prodest?", "¿a quién beneficia?". Sin embargo, en este caso, como en muchos otros,

²⁶ Mark Coeckelbergh, *La filosofía política de la inteligencia artificial. Una introducción*, Cátedra, Madrid, 2023, pgs. 17 ss.

la pregunta más adecuada es doble: "¿a quién beneficia y a quién perjudica?".

Por mucho que se empecinen los propagandistas de la IA, la respuesta no puede darla ella misma, porque carece de intención. Es un conjunto de instrumentos muy útiles en manos de la inteligencia humana, pero incapaz por sí misma de comprender, decidir y elegir. Es la inteligencia general la que puede hacerlo, la que señala de qué problemas merece la pena ocuparse, y hoy por hoy no hay más inteligencia general que la humana.

Y ante la pregunta que hemos formulado, la veracidad sigue siendo una de las condiciones de validez de la comunicación, y en el caso de la IA esa condición se cumplirá informando sobre los beneficios que está proporcionando ya, que son muchos, y los que cabe esperar que procure en el futuro hasta donde se pueda prever *con base científica*. Pero también respondiendo a una segunda pregunta tan importante como la primera: ¿Beneficios para quién? ¿Para los ya poderosos de modo que se amplíen las desigualdades? ¿Para los menos favorecidos por la fortuna y por la sociedad?

Recurrir a la ética para eludir las ideologías es de primera necesidad. Cualquier ética aplicada, también la de la IA, enmarca su reflexión en unos principios que ordenan beneficiar, no dañar, empoderar la autonomía de los seres humanos, distribuir los beneficios con justicia, en este caso entre todos los seres humanos, que son los afectados por la IA, y potenciar la sostenibilidad de la naturaleza. Claro que el mundo de intereses políticos y económicos, individuales y grupales, que pone en marcha una actividad tan lucrativa es inmenso. Pero en pleno siglo XXI estos principios éticos son ya los trazos irrenunciables de una ética global²⁷.

Sin embargo, como hemos comentado, durante los 70 años de la IA, y muy especialmente en los últimos tiempos, junto a las alabanzas, se han venido multiplicando las críticas de

²⁷ Adela Cortina, *¿Ética o ideología de la inteligencia artificial?*, cap. 4.

un creciente número de expertos sobre la forma de producirla. La IA se produce recurriendo a la explotación del trabajo de los supervisores con sueldos de miseria, a la colonización una vez más de los países menos desarrollados y al expolio del planeta. Pero también conlleva consecuencias indeseables. Se amplía la brecha digital entre las personas y entre los países, la posible disrupción del modelo laboral, que anuncia una nueva época, dificulta satisfacer las aspiraciones de la socialdemocracia, y es una verdadera tentación dejar en manos de los algoritmos las decisiones en asuntos que afectan a los seres humanos, eludiendo responsabilidades, cuando lo cierto es que los algoritmos no toman decisiones, porque carecen de conciencia y voluntad. Ofrecen soluciones y somos los seres humanos quienes tenemos que decidir y asumir responsabilidades. El libro de Cathy O'Neil, *Armas de destrucción matemática* sigue siendo dolorosamente instructivo²⁸.

Ante un panorama tan sombrío es preciso regresar a la pregunta "IA ¿para qué y para quiénes?". Y afortunadamente cada vez se elevan más voces exigiendo que la IA sirva al bien común, que en este caso es ya local y global. A defender los derechos humanos y alcanzar los Objetivos del Desarrollo Sostenible, el primero de los cuales es acabar con el hambre y la pobreza extrema, dos realidades verdaderamente obscenas en un tiempo en que hay medios más que suficientes para que ningún ser humano sea pobre, mucho menos que muera de hambre.

El extraordinario aumento de las conexiones y las plataformas propiciado por la revolución 4.0 debería servir para lograr una comunicación auténtica que haga posible el entendimiento y la cooperación en el camino de una paz justa, más allá de las polarizaciones, los conflictos, los chantajes y las humillaciones a los que se ven sometidos los más débiles si quieren sobrevivir. Aunque sólo sea -por decirlo con Kant- porque hasta un pueblo de demonios, de seres sin sensibilidad moral, preferiría la paz a la guerra, el Estado

²⁸ O'Neil, Cathy, *Armas de destrucción matemática*, Capitán Swing, Madrid, 2018.

de Derecho al estado de naturaleza, con tal de que tengan inteligencia. No digamos ya un pueblo de personas humanas con inteligencia general, emociones y sentido moral.