

INTELIGENCIA ARTIFICIAL

Reflexiones desde la filosofía y la ética

Por el Académico de Número
Excmo. Sr. D. Diego Gracia Guillén*

SUMARIO

El tema de la «inteligencia artificial» ha cobrado en estos últimos años una actualidad no solo sorprendente sino también preocupante. De hecho, ha puesto en guardia a todo tipo de instituciones, que han comenzado a plantearse cómo controlar este sorprendente fenómeno, preocupadas por no llegar tarde a su necesaria regulación y control.

Hasta hace relativamente poco tiempo, el debate sobre la inteligencia humana se hacía siempre por contraste con la llamada «inteligencia animal», de tal modo que los protagonistas eran por lo general etólogos o psicólogos. Hoy el escenario es algo distinto, porque el contraste se establece con la llamada «inteligencia artificial», debido a la importancia que en las últimas décadas ha cobrado la computación y esa nueva ciencia llamada informática. Ni que decir tiene que esto exige aclarar qué significa inteligencia o de qué estamos hablando cuando usamos esa palabra.

¿Qué es inteligencia? La definición más elemental la identifica con la capacidad de procesar información, algo que hacen tanto los seres vivos como ahora también los ordenadores. Lo que no podían hacer éstos era aprender,

* Sesión del día 31 de octubre de 2023.

algo que parecía privativo de los seres vivos. Pero a día de hoy las máquinas también aprenden, en sistemas como el Chat GPT.

¿Hay algo que diferencie la inteligencia humana de las máquinas y los animales? La filosofía ha respondido por lo general afirmativamente. Y a la pregunta por el origen de esa característica propia y peculiar de la especie humana ha dado dos respuestas, la clásica, que afirma la existencia de una sustancia espiritual en el interior del cuerpo humano, y la emergentista, para la que la novedad del psiquismo humano emerge por pura complejidad de los elementos estructurales del cuerpo. En este caso, se plantea la posibilidad de que la complejidad cada vez mayor de los sistemas informáticos acabe haciendo posible el que en algún momento acabe surgiendo esa cualidad nueva e irreductible a todo lo anterior que llamamos inteligencia específicamente humana. Algo que, en cualquier caso, parece muy alejado de las capacidades científicas y técnicas actuales.

INTRODUCCIÓN

El tema de la «inteligencia artificial» ha cobrado en estos últimos años una actualidad no solo sorprendente sino también preocupante. De hecho, ha puesto en guardia a todo tipo de instituciones, que han comenzado a plantearse cómo controlar este nuevo fenómeno, preocupados por no llegar tarde a su necesaria regulación y control¹.

La sorpresa ha dado lugar a todo tipo de especulaciones sobre lo que puede llegar a suceder en el próximo futuro. La verdad es que no lo sabemos. Pero sí parece llegado el momento de iniciar la reflexión sobre este sorprendente fenómeno, que a no dudar tendrá presencia recurrente en esta Academia a partir de ahora. De hecho, Adela Cortina se ocupaba ya de este problema en su reciente artículo *¿Personas electrónicas?*²

DE DÓNDE VENIMOS: LA INTELIGENCIA NATURAL

El término «inteligencia artificial» se ha acuñado sobre un patrón de vieja progenie en nuestra cultura, el de «inteligencia natural». Los filósofos griegos, y más en concreto Aristóteles, definieron al ser humano, precisamente, por su inteligencia. *Zôon lógon ékbon*, animal dotado de *lógos* lo llama Aristóteles

¹ Cf. European Commission, *White Paper on Artificial Intelligence*, COM(2020) 65 final.

² CORTINA, A., «¿Personas electrónicas?», *El País* el 7 de septiembre de 2023.

al comienzo de su *Política*.³ En esto ha visto toda la tradición occidental la definición de ser humano, a diferencia de los animales, carentes de *lógos*, dado que no están dotados de lo que Aristóteles llama en su tratado *De anima*, *noûs*, inteligencia o intelecto, una cualidad por demás peculiar, ya que, como afirma, «es separable, sin mezcla e impasible, siendo como es acto por su propia entidad.» Y añade que el *noûs*, «una vez separado, es solo aquello que en realidad es y únicamente esto es inmortal y eterno»⁴. Tales afirmaciones sobre el intelecto agente las hace Aristóteles no sin confesar sus dudas al respecto. «Por lo que hace al intelecto y a la potencia especulativa no está nada claro el asunto, si bien parece tratarse de un género distinto de alma y que solamente él puede darse separado como lo eterno de lo corruptible»⁵. Y más adelante aclara: «La facultad sensible no se da sin el cuerpo, mientras que el intelecto es separable»⁶. Esta entidad inmortal y eterna, que diferencia a los seres humanos de los animales, tiene la capacidad de abstraer las formas inteligibles de la materia. Los textos citados de Aristóteles son complejos y han dado lugar a diversas interpretaciones. En cualquier caso, la teoría que se impuso fue la de que, a diferencia de las almas de las plantas y los animales, que son materiales, el alma de los seres humanos es espiritual y capaz de penetrar en las formas de las cosas, que son universales, y por tanto captar su esencia.

EL DEBATE SOBRE LA SEDE DE LA INTELIGENCIA: CARDIOCENTRISMO Y CEREBROCENTRISMO

Afirmado lo anterior, inmediatamente surgió la pregunta de dónde asentaba ese *noûs* que es tan esencial y definitorio del ser humano. Y aquí las respuestas se dividieron. Hubo un amplio conjunto de pensadores, entre ellos Aristóteles, que fueron «cardiocéntricos», porque defendieron que la sede del psiquismo superior estaba en el corazón, en tanto que otros fueron «cerebrocéntricos», la tesis que con el tiempo acabó desplazando a la anterior. Conviene recordar, con todo, que fue un convencido cardiocéntrico, William Harvey, quien en pleno siglo XVII descubrió y describió la circulación mayor de la sangre en un libro que ya desde el título, *De motu cordis*, recuerda su progenie cardiocéntrica⁷.

Pero el cerebrocentrismo fue imponiéndose poco a poco, como consecuencia de los datos aportados por la Medicina. El cerebro parecía ser la sede de la conciencia. La anatomía y la fisiología fueron aportando datos cada vez

³ *Pol* I 2: 1253 a 9-10.

⁴ *De an* III 5: 429 b 18-23.

⁵ *De an* II 2: 413 b 25-28.

⁶ *De an* III 4: 429 b 5-6.

⁷ Cf. GRACIA, D., «Las razones del corazón». *Naturaleza y Gracia* 2004; LI (2/3): 345-392.

más convincentes sobre ello. La sede de la conciencia y de la propia vida, por tanto del alma, parecían asentar en el cerebro. Todavía Descartes, en pleno siglo XVII, especuló con que el punto de unión entre el cuerpo y la mente asentaba en la glándula pineal o epífisis.

Hoy no se duda de que el SNC es la sede de la conciencia, y que dentro de él hay una parte, la de aparición más reciente en el proceso evolutivo, el llamado neocortex o corteza cerebral, en que asientan las cualidades psíquicas específicamente humanas. Los neurofisiólogos suelen englobarlas en un término filosóficamente tan impreciso como «conciencia», *consciousness*. De etapas previas de la evolución hemos heredado otras estructuras que constituyen el llamado paleocortex, que es condición necesaria pero no suficiente para la conciencia humana. Lo cual explica que en el ser humano haya una gran cantidad de actos, la mayoría, que no son conscientes en el sentido preciso del término. Son los actos inconscientes.

DEL DEBATE FILOSÓFICO AL DEBATE CIENTÍFICO SOBRE LA INTELIGENCIA

Hasta hace relativamente poco tiempo, el debate sobre la inteligencia humana se hacía siempre por contraste con la llamada «inteligencia animal», de tal modo que los protagonistas eran por lo general etólogos o psicólogos. Hoy el escenario es algo distinto, porque el contraste se establece con la llamada «inteligencia artificial», debido a la importancia que en las últimas décadas ha cobrado la computación y esa nueva ciencia llamada informática.

Ahora la pregunta es si las máquinas son inteligentes, o, volviendo al vocablo tan traído y llevado en el debate anglosajón, si los ordenadores serán capaces de tener «conciencia» o llevar a cabo procesos «conscientes». La literatura sobre este tema se ha disparado en estos últimos años y parece que el debate no ha hecho más que comenzar⁸. Ni que decir tiene que esto exige aclarar qué significa inteligencia o de qué estamos hablando cuando usamos esa palabra.

⁸ Cf. MONSERRAT, J., «El transhumanismo de Ray Kurzweil. ¿Es la ontología biológica reductible a computación?» *Pensamiento*, 2015;71:1417-41.

¿QUÉ ES INTELIGENCIA? PRIMERA RESPUESTA: LA INTELIGENCIA COMO CAPACIDAD DE PROCESAR INFORMACIÓN

Hubo que esperar al nacimiento de la Psicología experimental y a la aparición de los llamados tests de inteligencia, para que la idea de inteligencia cambiara de modo drástico. El criterio de inteligencia de la psicología no es filosófico sino operativo, práctico, y viene a identificarse con la «capacidad de procesar información». Hay personas que tienen una enorme capacidad para procesar información numérica, o el lenguaje, etc. Esto llevó a distinguir varios factores de inteligencia, y luego uno general, el llamado factor g. Estos factores se miden mediante unos tests o pruebas en los que puede comprobarse la capacidad de procesar información de cada persona, que es distinta en los diversos factores y que converge en el llamado factor general de inteligencia.

Con la aparición de los ordenadores, que también procesan información, y a veces en cuantía y velocidad muy superiores a las humanas, se ha comenzado a hablar también de «inteligencia artificial». Los ordenadores, se dice ahora, son más o menos inteligentes, porque procesan información en medida mayor o menor.

¿QUÉ ES INTELIGENCIA? SEGUNDA RESPUESTA: LA INTELIGENCIA COMO CAPACIDAD DE APRENDIZAJE: LA INTELIGENCIA ANIMAL

Hasta muy recientemente se ha venido pensando que los ordenadores tenían capacidad de procesar información, pero no de aprender, algo que sí hacen los seres vivos, por ejemplo, los animales. El aprendizaje de conductas es una cualidad que suele atribuirse al psiquismo, pero que es profundamente biológica. Los cuerpos vivos aprenden. Es lo que el fisiólogo Walter Cannon llamó *The Wisdom of the Body*⁹. El organismo aprende a dormir en un cierto momento, a despertarse en otro, a tener hambre a una cierta hora, etc., etc. Los reflejos condicionados de Pavlov son un sistema de aprendizaje. En eso consiste la domesticación de perros, animales, etc. Y en cierto modo en eso consiste también la educación de las personas. Pero como veremos inmediatamente, quizá la educación no es mera domesticación. Alguno de los teóricos de la psicología del aprendizaje, como Skinner, creyó que con puros reflejos podría conseguirse el mundo feliz de Aldoux Huxley. Recuérdese su novela *Walden dos*¹⁰.

⁹ CANNON, W., *The Wisdom of the Body*, New York, Norton, 1939.

¹⁰ SKINNER, B. F., *Walden Dos: Hacia una ciudad científicamente construida*, Barcelona, Martinez Roca, 1987.

Pero esto, que pasaba por ser específico de los seres vivos, el aprendizaje, resulta que ahora pueden hacerlo también las máquinas. Esto es lo que hoy se denomina «inteligencia artificial». Su expresión más clara la encontramos en sistemas como el Chat GPT, hoy de rabiosa actualidad y que está causando tanta admiración como espanto.

¿HAY ALGO ESPECÍFICO EN LA INTELIGENCIA HUMANA?

Tras este rápido recorrido que hemos llevado a cabo, parece claro que por inteligencia se han entendido y se entienden cosas sumamente diversas. Desde el tiempo de los antiguos filósofos griegos, por inteligencia se ha venido entendiendo el *intus-legere* o *inter-legere*, la capacidad de abstraer las formas de la materia y captar la esencia. Se definía como la capacidad de *intus-legere*, donde *legere* remite a *lógos*, razón. Por eso al ser humano se le definía como *zôon lógon ékbon*, *animal rationale*, animal racional. El *intus legere* se entendía como la capacidad de «abstraer» la «forma» respecto de la «materia» de las cosas, y hacerse, de ese modo con las «ideas» de las cosas. Esta era la definición clásica, de progenia plotónica y aristotélica. En el interior del ser humano hay una realidad o sustancia de características completamente distintas a las materiales, que tiene la capacidad de trascender la materia y situarse en el universo de las ideas puras. En la naturaleza es posible que nunca haya habido un triángulo perfectamente equilátero, pero la mente humana puede pensar y razonar sobre el triángulo equilátero, con independencia de que tenga existencia empírica o mundanal. Por tanto, las ideas puras habitan en otro mundo distinto del material. Es el *hyperouránion tópos* de Platón, o lo que en toda la tradición occidental se ha entendido por la mente divina. De esto se deduce que la inteligencia es una capacidad o prerrogativa específicamente humana. Y más que humana, divina, de tal modo que en el ser humano no se da de modo completo, sino participado, como decía Platón. Dios es intelecto puro, *nóesis nóeseos nóesis*, dice Aristóteles¹¹, y por ello Tomás de Aquino lo definirá como *actus purus*, dado que en él no puede haber nada en potencia.

En la filosofía moderna, y de modo especial en el idealismo alemán, eso es lo que se denominará *Vernunft*, razón, a diferencia de la *Verstand*, entendimiento. Y de nuevo la *Vernunft* es condición divina, de la que el ser humano participa en alguna medida. En Dios no hay *Verstand* sino solo *Vernunft*.

Si acudimos a un representante paradigmático de la filosofía alemana del siglo xx, Heidegger, nos dirá que lo propio y característico del ser humano es la «comprensión» (*Verstehen*, *Verständniss*) del Ser, algo que no se da en los

¹¹ *Met* XII 9: 1076 b 34.

animales ni en las máquinas. Y con él toda la filosofía hermenéutica, al menos desde los tiempos de Dilthey.

En el pensamiento español, Zubiri ha definido la inteligencia como la capacidad de actualizar las cosas en tanto que «realidades». Y «cosa-real» es para él lo opuesto a «cosa-estímulo», la que se agota en desencadenar una respuesta, que es el modo como parece que procede eso que se llama inteligencia animal, y por supuesto la inteligencia artificial. La inteligencia es una «formalidad», la propia de la realidad humana, que permite «formalizar» o «actualizar» las cosas como «de suyo» o «en propio» reales, y por tanto como realidades y no como meros estímulos que suscitan respuestas. Si este es el sentido propio del término inteligencia, debería concluirse que, en el rigor de los términos, no es correcto hablar de inteligencia animal, y menos de inteligencia artificial.

Los estudios de los investigadores de la Universidad de Leipzig, en especial Michael Tomasello, parecen indicar que lo específico de la inteligencia humana respecto a la de los otros primates es la capacidad de tener en cuenta en las decisiones a los otros, en vez de tener en cuenta solo a uno mismo, como parece que sucede en los primates. De tal modo que el paso del modo «yo» al modo «nosotros» sería lo más específico y propio de la inteligencia humana¹².

Pero sucede que esta enorme ampliación del concepto de inteligencia respecto de su sentido usual a lo largo de los siglos, se ha hecho, como no podía ser de otro modo, a costa de perder precisión. Como diría un lógico, ha ganado en extensión pero ha perdido en comprensión. Quien mucho abarca poco aprieta, dice el refrán castellano. Si todo es inteligente, resulta que no vamos a poder definir con ello lo que es propio y específico de la especie humana, la inteligencia humana, que es por donde comenzó esta historia.

Michael Tomasello piensa que es el «modo nosotros». La tesis de Zubiri es distinta, aunque no incompatible: la aprehensión de las cosas como «realidades». Habría que ver si estos dos enfoques no están relacionados, y si el paso al modo «nosotros» no se debe, precisamente, a la capacidad de aprehender las cosas, no ya como estímulos que desencadenan respuestas, sino como cosas «de suyo» o «realidades». De ser esto así, habría que concluir que lo específico de la inteligencia humana es, precisamente, esa capacidad de actualizar las cosas como cosas en sí, y no solo como estímulos que suscitan respuestas.

¹² Cf. HERRMANN, E.; CALL, J.; HERNÁNDEZ-LLOREDA, M. V.; HARE, B.; y TOMASELLO, M. «Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis». *Science*, 2007; 317 (5843): 1360–1366. Tomasello, M.; Carpenter, M.; CALL, J.; BEHNE, T.; y MOLL, H. «Understanding and sharing intentions: The origins of cultural cognition». *Behavioral and Brain Sciences*, 2005;28: 675-735. WARNEKEN, F.; HARE, B.; MELIS, A. P.; HANUS, D.; TOMASELLO, M.; «Spontaneous Altruism by Chimpanzees and Young Children». *PLoS Biology*, 2007;5: 1414–1420. I. Pablo, Sahagún Kunhard y Gonzalo Alonso-Bastarache, «Lo que nos hace humanos es la intencionalidad compartida: Michael Tomasello», en: Juan Arana (coord.), *Concepciones antropológicas de los protagonistas de la revolución neurocientífica*, Valencia, Tirant Humanidades, 2023, 361-389

DE LA CUESTIÓN FORMAL A LA PREGUNTA POR LA GÉNESIS: EL EMERGENTISMO

Llegados a este punto, es lógico preguntarse por el origen de esta peculiaridad de la mente humana. De dónde procede esta peculiar *psyché* del ser humano, ha recibido distintas interpretaciones a lo largo de la historia. Hay tres posibilidades interpretativas: la *animista*, que afirma la existencia de una sustancia espiritual, irreductible a la materia, en el ser humano, la *reduccionista* que reduce la mente a la materia orgánica, y la *emergentista*, que la interpreta como un salto cualitativo, y por tanto como una novedad estructural no reducible a sus componentes¹³. Al menos desde el tiempo de Hegel se habla del «salto de la cantidad a la cualidad». Por tanto, en la naturaleza hay saltos, y saltos que hacen aparecer cualidades nuevas, irreductibles a los elementos que las originaron. Esto hoy es tópico en química, y está en la base de teorías hoy muy actuales, como la del caos. Pequeñas modificaciones elementales acaban produciendo saltos estructurales en que aparecen novedades inesperadas. Ya decía Aristóteles que el todo es más que la suma de sus partes. Y Zubiri habla, por ello, de unas «propiedades sistemáticas», distintas e irreductibles a las «propiedades elementales»¹⁴.

Las tres hipótesis interpretativas citadas han dominado en diferentes épocas de la historia. La tradicional ha sido el «animismo». El materialismo, especialmente en el siglo XIX, defendió muy beligerantemente el «reduccionismo». Y en el siglo XX fue cobrando cada vez mayor fuerza la hipótesis emergentista.

La cuestión se reduce a saber si la inteligencia específicamente humana es una propiedad elemental o sistemática. La primera respuesta es la propia del creacionismo y la segunda la propia del emergentismo. Hoy por hoy no parece que haya razones que diriman completamente este dilema. Sí hay que decir que, aunque solo sea por permanecer fiel al principio de economía del pensamiento, parece más razonable la respuesta emergentista que la creacionista.

Según esta tesis, la emergentista, la inteligencia específicamente humana es una propiedad sistemática nueva, emergida de la complejidad estructural

¹³ Cf. BEORLEGUI, C., «El emergentismo humanista de Javier Monserrat», *Pensamiento* 2015; 71 (269): 1165-1200.

¹⁴ Cf. GRACIA, D., «Zubiri en los retos actuales de la antropología», en Pintor-Ramos (coord.), *Zubiri desde el siglo XXI*, Salamanca, Publ. Univ. Pontificia de Salamanca 2009, pp. 111-158. GRACIA, D. «Problemas filosóficos de la génesis humana», *Franciscanum*, 27:79 1985; 27(79): 17-32; Ibid. «Problemas filosóficos en genética y en embriología», en F. Abel; C. Cañón (eds.), *La mediación de la filosofía en la construcción de la bioética*, Madrid, Federac. Internac. Universidades Católicas / Univ. Pont. Comillas 1993., pp. 215-154; D. Gracia, «La antropología de Zubiri», en J. A. Nicolás; O. Barroso (eds.), *Balance y perspectivas de la filosofía de Zubiri*, Granada, Comares 2004, pp. 87-116. J. Sáez, «Los retos actuales de la filosofía zubiriana y su 'giro trascendental' desde la noología a la trascendencia metafísica», *Cuadernos salmantinos de filosofía* 2010; 37: 397-438.

de los elementos constitutivos del sistema, que abre al ser humano a un mundo nuevo, el de lo que podemos llamar la vida del espíritu, o a lo que Zubiri denominaría el orden de la realidad en cuanto realidad. Ese mundo nuevo es el mundo del espíritu.

LAS HIPÓTESIS ACTUALES

La neurofisiología lleva mucho tiempo investigando el modo como surge la conciencia específicamente humana. Hoy por hoy, no ha conseguido dar una explicación plausible que haya concitado la aprobación general y pueda darse por satisfactoria. Está más claro el modo o los modos como se pierde la conciencia, que el mecanismo por el que se gana. En lo que existe un acuerdo muy amplio, aunque desde luego no total, es en que ese milagro de la conciencia se consigue a través de la casi inextricable maraña de conexiones que se producen entre las neuronas del cerebro, y sobre todo de su corteza.

El año 1990, Francis Crick, el descriptor junto con James Watson de la estructura terciaria del ADN, publicó un libro titulado *The Astonishing Hypothesis: The Scientific Search For The Soul*, «La hipótesis sorprendente o revolucionaria: la búsqueda científica del alma». Su hipótesis no es más que eso, una hipótesis, pero hoy aceptada por la mayoría de la comunidad científica. Lo que dice es que «las actividades mentales de una persona se deben por entero al comportamiento de sus células nerviosas, células gliales y los átomos, iones y moléculas que las componen y las influyen»¹⁵. Lo de *astonishing* tiene que ver más con el marketing que con la realidad, y demuestra la ignorancia sorprendente por parte del laureado Crick del estado de la filosofía de su tiempo, ya que solo parece conocer, y poco, la de tiempos muy pasados, esa que somete a dura crítica. En cualquier caso, la idea hoy más aceptada es que el psiquismo específicamente humano es una propiedad sistemática dependiente de la estructura cerebral.

¿PODRÁ LA INTELIGENCIA ARTIFICIAL ALCANZAR ESTE GRADO DE DESARROLLO?

El salto de la cantidad a la cualidad es un misterio que no sabemos muy bien cómo se produce. Lo que sí está claro es que el resultado es algo cualitativamente distinto a todo lo anterior. Eso es lo que llamamos un salto

¹⁵ CRICK, F., *La búsqueda científica del alma: Una revolucionaria hipótesis para el siglo XXI*, Barcelona, Debate, 1994, p. 345.

cualitativo. Se produce en las combinaciones químicas, en las que de elementos con determinadas propiedades surgen compuestos con propiedades nuevas, irreducibles a las anteriores. En la evolución humana, parece que ese salto también se ha producido, y probablemente no una vez, sino varias. Zubiri se ocupó de esto, hace ahora más de medio siglo, en su artículo «El origen del hombre»¹⁶. La inteligencia humana es cualitativamente distinta de la animal e irreducible a ella, pero a la vez dependiente de ella. Surge, como decía Zubiri, por una especie de «desgajamiento exigitivo», que a fin de cuentas es un tipo de emergencia.

¿Podría suceder algo similar en los procesadores de información que utiliza la inteligencia artificial? ¿Podría darse en ellos el salto de la cantidad a la cualidad, y por tanto el salto a la inteligencia específicamente humana? No parece que haya argumentos para negar esa posibilidad, que ya se ha dado en el mundo animal. ¿Podrían llegar a los ordenadores a ser inteligentes, en el sentido preciso y específico o propio de la inteligencia humana? Podría ser. No hay razones para negar esa posibilidad.

Pero eso aún no ha sucedido y no sabemos si sucederá. Lo que sí podemos decir es que ese salto a día de hoy no se ha producido, y que esto que se llama en la actualidad inteligencia artificial, no lo es en el rigor de los términos, es decir, cuando la palabra inteligencia se entiende en sentido fuerte, y no como mera capacidad de procesar información o simple capacidad de aprendizaje.

CONCLUSIÓN

De ser un puro asunto de especialistas, el tema de la inteligencia artificial ha pasado a ocupar y preocupar a la opinión pública, hasta convertirse en uno de los temas y problemas más sorprendentes, a la vez que preocupantes de nuestra actualidad. La sorpresa del fenómeno ha dado lugar a todo tipo de especulaciones sobre lo que puede llegar a suceder en el próximo futuro. De hecho, ha puesto en guardia a las principales instituciones internacionales, que han comenzado a plantearse cómo controlar este nuevo y revolucionario fenómeno, preocupadas por no llegar tarde a su necesaria regulación y control. De lo que no parece haber duda razonable es de que promete ser una de las mayores revoluciones técnicas acaecidas en la historia entera de la humanidad, llamada a transformar muy profundamente la vida de nuestras sociedades. De ser esto así, la reflexión sobre ella es muy razonable pensar que habrá de ocuparnos de modo cada vez más intensivo a lo largo de los próximos años, quizá décadas.

¹⁶ Cf. ZUBIRI, X., «El origen del hombre», en *Escritos menores (1953-1983)*, Madrid, Alianza, 2006, 65-102.

En estos inicios del proceso, la primera pregunta que cabe hacerse es la referida a lo que los escolásticos llamaban el *quid nominis*. ¿Es correcta la expresión inteligencia artificial? ¿Qué significa exactamente? ¿Sabemos con precisión de qué estamos hablando?

La palabra inteligencia procede de la raíz latina *lego* que significa leer, por lo que originariamente tiene el sentido de leer dentro o captar la esencia o realidad de algo. En la filosofía antigua, esto acontecía mediante el proceso mental de abstracción de la llamada forma respecto de la materia, de tal modo que entender algo era captar con la mente lo que ella es, eso que los filósofos griegos llamaron su forma, abstrayéndola de la materia. Al decir que algo es una silla, la materia de la silla no penetra en nuestra mente, pero sí su forma, eso que nos hace decir que es una silla y no cualquier otro objeto. Esa captación de la forma, que viene a identificarse con la esencia de las cosas, en este caso de la silla, es la que expresamos atribuyéndola un nombre, silla. El lenguaje es, según esto, la expresión de esa esencial formal de las cosas. Los griegos utilizaron una palabra, *lógos* para designar este sorprendente fenómeno, y de ahí que ese término signifique en griego tanto razón o intelección como palabra. Y Aristóteles definió al ser humano, a diferencia de los demás seres vivos, como el animal dotado de *lógos*, es decir, de razón y de palabra, lo que le distingue de todos los demás animales.

Esta es la definición de inteligencia que ha circulado a lo largo de toda la historia de la cultura occidental, hasta hace relativamente poco tiempo, en que la ciencia que se ocupa del estudio del comportamiento animal, la etología, empezó a hablar de «inteligencia animal», algo incomprensible en toda la tradición anterior. Mucho más reciente es aún la expresión «inteligencia artificial», nacida con el auge de la informática en el último medio siglo. Y surge la pregunta: ¿cabe hablar de inteligencia a propósito de las máquinas? ¿Son las máquinas inteligentes?

De acuerdo con la definición griega de inteligencia, evidentemente, no. Las máquinas, al menos por ahora, no son capaces de llevar a cabo el proceso de abstracción de que hablaron los filósofos antiguos, ni tampoco de aprehender las cosas como realidades. Su gran habilidad está en el proceso y manejo de la información, de tal manera que las máquinas son inteligentes si por inteligencia entendemos, como hoy es usual, la capacidad de procesar información. En este sentido amplio, amplísimo del término, no hay duda de que los ordenadores son inteligentes, y los animales también. Es más, poco a poco vamos asumiendo el hecho de que, así definida la inteligencia, los animales, y sobre todo las máquinas, son en ciertos tipos de funciones más inteligentes que los propios seres humanos.

Pero inmediatamente surge la objeción de que las máquinas no tienen conciencia de lo que hacen, ni por tanto pueden salir responsables de sus actos.

Tanto de los animales como sobre todo de ellas nos separa, entre otras características, la ética. No parece que las máquinas tengan conciencia del deber. ¿Podrán adquirirla en algún momento? No lo sabemos, pero tampoco podemos descartarlo completamente. En la naturaleza son continuos los saltos de la cantidad a la cualidad, de tal manera que de la complejización de estructuras materiales acaban surgiendo propiedades no solo nuevas sino cualitativamente distintas a las de las anteriores e irreductibles a ellas. Es la teoría que se conoce con el nombre de «emergentismo». Sin ella es imposible entender la evolución de los seres vivos, en la que los saltos de la cantidad a la cualidad parecen haber sido muchos. ¿Llegarán algún día las máquinas a ser inteligentes, no en el sentido amplio y actual de ese término, el de capacidad de procesar información, sino en el sentido de tener conciencia de lo que hacen? ¿Podrá darse ese salto? No lo sabemos. Lo que sí sabemos es que por ahora ese salto no se ha dado y que incluso en esta época de euforia y optimismo ante los resultados de la inteligencia artificial, el momento parece lejano. Pero tampoco puede descartarse.