



Palabras clave: consciencia, inteligencia, inteligencia artificial, era axial, neurona artificial, redes neuronales artificiales, inteligencia artificial general.

LA CONSCIENCIA HUMANA ANTE LA IRRUPCIÓN DE LA INTELIGENCIA ARTIFICIAL GENERAL (IAG)

Consciencia

eIAA

La consciencia y la inteligencia artificial han sido siempre dos temas muy distintos. La primera altamente subjetiva y conectada a lo más profundo de la mente humana, y la segunda muy material, física y producto de la acción humana observable y medible. Aunque siempre ha habido preocupación por la posible relación entre ambas, es ahora cuando hay un interés especial desencadenado por la posibilidad de que las máquinas puedan llegar a ser conscientes como los hombres. En lo que sigue, se analizan algunos aspectos relativos a estas cuestiones, especialmente en cuanto a lo que es la consciencia y lo que puede terminar siendo la inteligencia artificial.



ADOLFO CASTILLA.
Doctor Ingeniero del ICAL
y catedrático de Economía Aplicada

Human consciousness in front of the irruption of artificial general intelligence (AGI)
CONSCIOUSNESS AND AI

Consciousness and artificial intelligence have always been two very different subjects. The first is highly subjective and connected to the depths of the human mind, and the second is very material, physical and the product of observable and measurable human action. Although there has always been concern about the possible relationship between the two, it is now that there is a special interest triggered by the possibility that machines may become conscious like humans.

In what follows, some aspects relating to these issues are discussed, especially as regards what consciousness is and what artificial intelligence may end up being.

Keywords: consciousness, intelligence, axial era, artificial neuron, artificial intelligence general, artificial neuron, artificial neural networks, general artificial intelligence.

ILUSTRACIÓN: ISABEL ALBERTOS



M

No hay temas más dispares a priori que los dos tratados en este trabajo. La inteligencia artificial es solo software, programación e instrucciones para un ordenador hechas por el hombre, mientras que la consciencia es pura subjetividad, lo más interno e inmaterial de la mente de los humanos, que además no sabemos todavía cómo surge o se genera.

Las dos se unen en nuestros días, sin embargo, en el debate actual sobre si la inteligencia artificial puede llegar a tener autoconsciencia como la tienen el hombre y la mujer.

Escribir sobre las dos cuestiones y su interrelación es algo arriesgado a la vista de los cientos, por no decir miles de artículos, libros, memorias de investigación y papers de todo tipo publicados en los últimos tiempos. La mayoría de ellos accesibles en Internet. ¿Es posible decir algo nuevo?

Como se sabe, la consciencia es un área de reflexión muy antigua. Surge con la aparición del pensamiento en la Era Axial (800 a. de C. – 200 a. de C.) según Karl Jaspers (1883-1969). De la mano, en gran manera, de Platón, acompañado de otros pensadores de la época como Parménides, Heráclito o Tucídides, y precedida por la labor de los presocráticos.

No sabemos de dónde procede y qué genera la consciencia.

Esa cuestión de “ser conscientes” es la que todavía no sabemos de dónde procede y cómo se produce en la mente humana. A pesar de los milenios transcurridos, el ser humano no ha avanzado demasiado en el terreno de los procesos mentales de este tipo,

a los que hay que incorporar, además, a los subconscientes y a los inconscientes, así como a los sentimientos más profundos, a la intuición, la imaginación, la inspiración, la creatividad, la inteligencia misma y, en parte, el aprendizaje. Sabemos lo que es la consciencia, pero no cómo aparece en el hombre y qué la genera.

No podemos hacer aquí una revisión detallada de todo lo que el hombre ha aportado en relación con la consciencia a lo largo del tiempo, pero sí deberíamos recordar la labor de René Descartes (1596-1650) en cuanto a la distinción de dos sustancias en el hombre: la pensante, o *res cogitans*, que es la más importante, la que nos hace hombres y la que nos distingue de los animales; y la sustancia extensa, o *res extensa*, formada por el cuerpo y su conexión con el mundo material y físico, el tiempo y otros componentes del universo en el que vivimos.

Dualismo

Este destacado autor fue el introductor del dualismo, aunque parece que él creía en la unión íntima en el hombre de las dos sustancias. A pesar de ello, la historia posterior interpretó sus aportaciones como la existencia de dos “naturalezas”, en vez de dos sustancias, y que había diferencias profundas entre ellas, hasta el punto de que la primera era el ámbito de la fe y las creencias, y la segunda el de la objetividad y la ciencia.

Pronto surgieron críticas a las explicaciones de Descartes, especial-

mente por parte del empirismo inglés, que creía en el cerebro humano como una tabula rasa de partida, que aprendía a través de su interrelación con su hábitat. Es lo que se llamó monismo del que ha habido diversas interpretaciones también. Hoy, por ejemplo, hablamos del monismo trascendental, que considera que las dos sustancias son componentes indisolubles de la naturaleza humana, como las dos caras de una misma moneda, pero con capacidades de la componente mental que aún no podemos explicar y que nos conectan a una posible consciencia cósmica.

Es curioso, pero a mediados de la década de los años 50 se produjo la interrelación del pensamiento humano con el mundo de los ordenadores o, dicho de otra forma, las ciencias cognitivas incorporaron a sus intereses la programación de la lógica humana en las computadoras. Como fecha de comienzo de la inteligencia artificial se suele indicar el verano de 1956 en el que se organizó una primera reunión de personas destacadas, que se distinguían por estar trabajando en ordenadores, y más específicamente en su programación. El trascendental encuentro en el Dartmouth College, en New Hampshire (Estados Unidos) ocurrió dos años después de la trágica muerte de Alan Turing¹. John McCarthy (1927-2011), profesor de dicho centro, convocó una conferencia a la que asistieron: Claude Shannon (1916-2001), Marvin Minsky (1927-2016), Frank Rosenblatt (1928-1971), Herbert Simon (1916-2001) y Allen Newell (1927-1992). En ella, y

especialmente en la documentación elaborada para pedir financiación, se utilizó la expresión inteligencia artificial. Hoy sabemos, y está todo documentado, que fue una idea feliz de McCarthy (1927-2011), quien había trabajado bastante en el tema antes de la reunión y la había definido como “el proceso de hacer que una máquina se comporte de formas que serían llamadas inteligentes si un ser humano las hiciera”.

Es verdad que muchos nos preguntamos por la verdadera inteligencia, que conectamos de forma muy clara a la consciencia. Aquí está el *hard problem* de la inteligencia artificial como lo denomina el filósofo, australiano de nacimiento, pero estadounidense de residencia, David Chalmers (nacido en 1966). ■■■

¹ Allan Turing (1912-1954) fue un notable investigador británico, matemático, lógico, informático teórico, criptógrafo, filósofo y biólogo teórico, que después de haber prestado grandes servicios a su país, entre ellos su contribución a descifrar los códigos empleados por los nazis en su mensajería encriptada, especialmente los de la máquina Enigma, construir dos de los primeros ordenadores, crear la idea de programación continua, trabajar teóricamente en la inteligencia artificial y muchas otras aportaciones, fue perseguido por homosexual, encarcelado y castrado químicamente, suicidándose a la temprana edad de 41 años.

Sabemos lo que es la consciencia, pero no cómo aparece en el hombre y qué la genera

Hay un punto al que, de momento, no llegan los ordenadores: el de la consciencia, la introspección y la actividad mental superior, entre otras cosas porque no sabemos lo que es ya que los procesos conscientes (y los inconscientes) no son algorítmicos y porque no existe (en realidad, es una ilusión).

Y lo cierto es que ha habido inviernos y veranos en esta disciplina, es decir, épocas con grandes expectativas y épocas de grandes decepciones. La actual parece claramente del primer tipo, sobre todo por la aparición del machine learning, el deep learning y el learnig reinforcement. Así como por el gran impacto social que ha tenido la llamada, inteligencia artificial general (IAG)², la que hemos llamado hasta ahora inteligencia artificial fuerte.

Basándose en esas formas de hacer que las máquinas aprendan y en los avances impresionantes llevados a cabo en la programación, la empresa OpenAI y otras están lanzando al mercado productos deslumbrantes como ChatGPT y DALL-E 2. Esta y otras empresas, menos conocidas y más especializadas, están creando una verdadera revolución en las sociedades actuales de todo el mundo. Algunos hablan ya de la V Revolución Industrial y anuncian cambios superiores a la introducción de la electricidad y el motor eléctrico.

A pesar de esos progresos, hoy se considera mayoritariamente que las máquinas no son conscientes y que no podrán serlo a menos que se avance mucho, y no solo en el terreno de la tecnología digital. Habrá que complementarlas con lo que hoy empezamos a llamar tecnología biodigital, que tiene que ver, muy ampliamente, con lo

que creemos que es el hombre: interconexión profunda de ser humano-naturaleza-tecnología.

Podemos concluir recopilando las posiciones de los hombres hoy en cuanto a la consciencia y las máquinas con las siguientes posiciones: 1) la de los que creen que la consciencia no existe y que en realidad es una ilusión; 2) la de los que creen que es un epifenómeno del cerebro, algo así como el vapor de agua que se produce al calentar el líquido elemento; 3) la de los que creen que solo hay cerebro, es decir, materia, y antes o después descubriremos cómo se produce la consciencia; 4) la de los que creen que la consciencia no es algorítmica y que hay que utilizar la mecánica cuántica para su explicación; 5) la de los que creen en la existencia de una consciencia cósmica o universal, lo que presupone la existencia de un ser superior, de la que procede la consciencia humana.

2 Algunos la llaman, erróneamente, inteligencia artificial generativa, mezclando en esta denominación lo que en realidad es ChatGPT: un *Generative Pre-trained Transformer*.

Bibliografía

- Arias, A.: (2021): *Introducción a la ciencia de la conciencia. El estudio de la experiencia subjetiva en filosofía, psicología y neurociencias*. Madrid, Los Libros de la Catarata.
- Campillo, J. E. (2021): *La consciencia humana. Las bases biológicas, fisiológicas y culturales de la consciencia*. Barcelona, Arpa.
- Koch, C. (2004): *La consciencia. Una aproximación neurobiológica*. Barcelona, Ariel.
- Monserrat, J. (2015): *El gran enigma: ateos y creyentes ante la incertidumbre del más allá*. Madrid, San Pablo.
- Neumann, E. (1970): *The Origins and History of Consciousness*, Princeton, N.J., Princeton University Press.
- Penrose, R. (1996): *Las sombras de la mente*. Barcelona, Crítica.
- Russell, S. y Norvig, P. (2014): *Inteligencia artificial. Un Enfoque Moderno*. Madrid, Pearson.
- Tibayrenc, M. y Ayala, F. (2021): *Lo que nos hace humanos. Biología, medicina, lenguaje, mente, ética y religión*. SalTerrae-Comillas