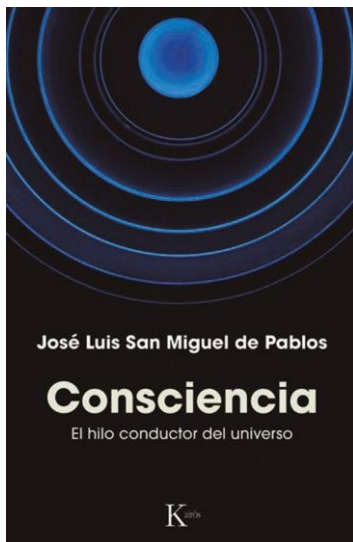


¿Podrían construirse inteligencias artificiales conscientes?

[Leandro Sequeiros] Muchos sociólogos y filósofos debaten sobre la emergencia de una nueva época en el desarrollo cultural de la humanidad: la época de la revolución digital. La que hasta ahora se ha denominado revolución científica parece empezar a ser barrida por la revolución digital. La mal llamada [Inteligencia artificial](#) parece querer desbancar con sus avances lo que se considera el [procesamiento de la información](#). Y muchos – con temor – empiezan a considerar si la [Singularidad](#) (el momento en el que los ordenadores superen las capacidades del cerebro) está llegando. Pero, [¿es realmente imposible que llegue a haber Inteligencias Artificiales conscientes?](#) Esta es la pregunta que se hace José Luis San Miguel de Pablo en un ensayo recién publicado: [Consciencia. El hilo conductor del universo](#).



(2023) CONSCIENCIA: EL HILO CONDUCTOR DEL UNIVERSO

JOSE LUIS SAN MIGUEL DE PABLOS
KAIROS- 9788411211802

Sinopsis de CONSCIENCIA

En tiempos de inteligencia artificial y postverdad se cuestiona sin cesar qué es real y cómo podemos distinguirlo con certeza. Las respuestas basadas en creencias, incluidas algunas que propone el cientifismo hegemónico, son engañosas e incapaces de disipar nuestro creciente desconcierto.

Al final del denso ensayo, su autor concluye: “Yo diría que no lo es *metafísicamente*, pero que sigue habiendo razones poderosas para el escepticismo, tanto a corto como a medio plazo. Ciertamente tanto un ser vivo como un ordenador están hechos de materia y energía, y por lo tanto sus “componentes elementales” podrían contar con un *lado reverso* de interioridad”.

Hay una diferencia: los animales superiores, y entre ellos el ser humano, han surgido en la naturaleza a través de un larguísimo proceso evolutivo que implica una cantidad ingente de **tiempo**, mientras que las IAs (Inteligencias Artificiales) son fabricadas contando tan solo con la acelerada evolución tecnológica que ha tenido lugar en las últimas décadas. No pocos tecnólogos responsables del deslumbramiento actual por la IA creen que esa misma aceleración será capaz de superar con creces a una evolución biológica que ha necesitado cientos de millones de años, y que culminará con el alumbramiento de una nueva entidad consciente y superinteligente

que nos suplantarán, pero mucho me temo que esa creencia, junto con la ambigua (y en el fondo agradable) emoción de película de terror que les produce, no sea otra cosa que un delirio de la *hybris* que dichos tecnólogos cultivan con fruición.

De acuerdo con el autor de este ensayo, [Consciencia. El hilo conductor del universo](#) (Editorial Kairós, 2023), - José Luis San Miguel de Pablos – en este ensayo se toca bastante el tema de la Inteligencia Artificial, y el autor manifiesta una actitud escéptica hacia la posibilidad de fabricar IAs conscientes. Ahora bien, está claro que la mención de esa posibilidad abre preguntas de enorme calado. Parte de la asunción de que la consciencia es idéntica a *la luz de ser*, es decir, a la pura interioridad o lado subjetivo de un ser humano o, como asumen ya la inmensa mayoría de los neurocientíficos, de cualquier ser viviente poseedor de un sistema nervioso complejo.

El nudo del presente comentario puede quedar definido, por tanto, planteando de entrada dos preguntas fundamentales: ¿Cuál puede ser una definición científica válida de “consciencia”? ¿Por qué la poseen los organismos biológicos, o al menos los que cuentan con un sistema nervioso centralizado por un cerebro?

¿Existe una definición científica válida de la “consciencia”?

A la primera pregunta en realidad ya ha respondido al proponer que la *luz de ser* es otra manera válida de llamar a la consciencia. Pero esta no es una definición y mucho menos científica, sino una descripción metafórica de una experiencia directa. Decir esto supone abrir una caja de Pandora. ¿Acaso es posible dar una definición científica de la consciencia? Su respuesta negativa, y hoy comparten ya no pocos científicos y filósofos. Y se justifica si consideramos qué clase de cosas son las que permiten establecer esas definiciones. En principio, cualquier objeto, en sentido amplio, puede llegar a ser definido científicamente. ¿Pero por qué fracasan todos los intentos de aportar una definición objetiva –y por lo tanto, científicamente rigurosa- de la consciencia? Por una razón muy sencilla, y es que **la consciencia no es objeto**.

A cualquiera de estos lo podemos llegar a conocer como una realidad fenoménica, es decir, como algo que se presenta a la consciencia por medio de los sentidos, pero la consciencia misma es el espacio de subjetividad en y para el cual hay fenómenos, y por lo tanto objetos que son cognoscibles como tales. *La luz de ser*, la consciencia, es lo que ilumina los objetos convirtiéndolos en fenómenos. Y la consciencia se vive directamente, sin necesidad de recorrer camino epistémico alguno, a diferencia (radical) de los objetos, que se conocen siempre indirectamente, a través de algún método de aproximación... a la consciencia, por más que su en-sí, como hizo notar Kant, resulta inaccesible.

Vivir experiencias conscientes es, pues, el único “método” de conocer la consciencia, y eso lo estamos haciendo todo el tiempo. Mientras que todo lo demás es entrar en un laberinto de circularidades viciosas muy poco aconsejables desde el punto de vista de la salud mental. Como vio con absoluta claridad Bertrand Russell (a quien nadie podrá acusar de irracionalismo), si hay conocimiento(s) científico(s) es porque hay consciencia, ya que ella es la precondition ineludible de todo conocimiento.

Nos enfrentamos, sin embargo, - nos dice el autor - al enigma científico y filosófico del correlato de la actividad neural con la consciencia y sus distintos estados. Pero sin asumir lo anterior es más que dudoso que se pueda avanzar en orden a resolverlo, por mucho que se identifiquen áreas cerebrales y dinámicas neuronales asociadas a experiencias conscientes; y ello porque el hiato ontológico entre lo objetual (como las neuronas y los electrones que se mueven entre ellas) y lo subjetivo es abismal.

¿Por qué los seres vivos poseen interioridad?

Estas consideraciones nos ponen ya frente al segundo interrogante: ¿por qué los seres vivos dotados de sistemas nerviosos evolucionados poseen interioridad, mientras que la presencia de esta en dispositivos informáticos, por sofisticados que sean, a muchos nos resulta inverosímil?

John Searle, en una célebre controversia que mantuvo hace ya tiempo con David Chalmers, emitió una opinión tajante. “La consciencia –dijo- es ante todo un fenómeno biológico y está tan confinada dentro de lo que constituye su propia biología como puede estarlo la secreción de la bilis o la digestión de los hidratos de carbono”.

Y, sin embargo, a lo largo de los años transcurridos, cada vez más estudiosos del tema le han ido dando la razón a Chalmers y quitándosela a Searle: un panpsiquismo raíz como explicación última del “problema fuerte” de la consciencia no solo no puede ser excluido sino que aparece como la más verosímil, la que menos contradicciones lógicas plantea.

Contradicciones, por supuesto, de orden filosófico, como las que originan un auténtico *shock* en muchos seguidores de la tradición científicista materialista al oír la palabra “panpsiquismo”, pues para ellos la sola mención de dicho término equivale a invocar lo más oscuro y depreciable del pensamiento mágico. Y sin embargo la “mágica” consciencia está aquí (no *ahí fuera*) y hay que notar que muchos de los que reaccionan de ese modo son los mismos que no saben, literalmente, qué hacer con ella.

¿Es posible el panpsiquismo?

El autor reconoce ser una de las personas que asumen un panpsiquismo no ingenuo (es decir que no implica creer que los electrones sean una suerte de duendecillos danzantes y juguetones) presente en los fundamentos de la *physis*. Que nos resulte inconcebible que el electrón tenga una *protointerioridad*, e incluso que el espacio-tiempo y la energía puedan ser solo el lado objetivo o fenoménico (es decir, el percibido por nosotros, focos de consciencia) de lo que sería, por su “lado reverso”, un campo de consciencia, no me parece una objeción de verdadero peso, pues ya deberíamos estar acostumbrados a las dificultades de concebir unas paradojas cuánticas que parecen ilustraciones perfectas de los *koan* del budismo zen.

Pero es que, además, la totalidad de los dispositivos informáticos, incluidas las IAs, son fabricados para satisfacer los fines de sus hacedores. Carecen, por tanto, de la autofinalidad que caracteriza a los seres vivos, hasta los más elementales, que son

agentes autónomos como los denomina Stuart Kauffman. Y todas las especulaciones sobre la robótica, desde Asimov hasta el miedo actual a un desbordamiento de la humanidad por unas IAs super-avanzadas, tiene que ver con la posibilidad o imposibilidad de que la tecnología sea capaz de crear agentes autónomos verdaderos, pues solo eso supondría realmente crear vida, ya que únicamente ella implica esa autonomía radical.

En el ensayo se plantea cómo podemos salir del laberinto de las apariencias engañosas que nos impiden reconocer *lo que es*. Surge de entrada la pregunta de si podemos reconocerlo con absoluta certeza en algún caso, y la respuesta es la misma que encontró Descartes, aún más simplificada: no “cogito ergo sum” sino solo “sum”, soy como pura experiencia directa de ser. Puede que sea imposible acceder al en-sí de las cosas, como pensaba Kant, pero todos accedemos al *en-mí*, puesto que todos vivimos la propia *luz de ser* en primera persona. Esto es, por tanto, una certeza absoluta que a algunas personas, como al autor, les lleva a mirar con conmiseración a los negacionistas de la consciencia.

Leandro Sequeiros. Presidente de ASINJA (Asociación Interdisciplinar José de Acosta) y colaborador de la Cátedra Hana y Francisco J. Ayala de Ciencia, Tecnología y Religión.