

LA SINGULARIDAD TECNOLÓGICA: ¿Mito o nueva frontera de lo humano?

José Manuel Elena Ortega
Universidad de Sevilla

Resumen: En el presente documento se realiza una aproximación al concepto de Singularidad Tecnológica y su relación con la Inteligencia Artificial en sus vertientes débil y fuerte, analizando las repercusiones que su desarrollo, tanto el actual como el previsible, tendrán en los aspectos sociales, políticos, económicos y filosóficos de la existencia del ser humano, e incluso sobre la definición de lo que nos caracteriza como tales.

Palabras clave: Singularidad tecnológica, inteligencia artificial, concepto de humano, filosofía.

The Technological Singularity: ¿Myth or New Frontier of the Human?

Abstract: In this document an approximation is made to the concept of Technological Singularity and its relation with the Artificial Intelligence in its weak and strong aspects, analysing the repercussions that its development, both the current one and the foreseeable, will have on the social, political, economic and philosophical aspects of the existence of the human being, and even on the definition of what characterizes us as such.

Keywords: Technological Singularity, Artificial Intelligence, Concept of Human Being, Philosophy.

Recibido: 27/09/2018 **Aprobado:** 20/02/2019

1. Introducción.

En junio de 2008 el magazine *Spectrum*, única publicación que el IEEE¹ distribuye de manera necesaria a todos sus miembros, estuvo dedicado en

¹ El IEEE, *Institute of Electrical and Electronics Engineers*, es la institución más prestigiosa del mundo tecnológico. Tiene en su haber más del 25% de las publicaciones que se realizan en el campo de la Ingeniería Eléctrica, la Electrónica y la Informática a nivel mundial.

su práctica totalidad a la Singularidad Tecnológica, definida como el momento clave en la evolución de la humanidad en el que podríamos contemplar cómo las máquinas² obtendrían niveles de inteligencia equivalentes a la inteligencia humana, haciendo entonces posible la migración de nuestra mente y consciencia a una de ellas.

El hecho de que el IEEE dedicara un *Special Report* a la Singularidad Tecnológica implicaba que no era un tema que, tal y como pudiera parecer, perteneciera al género de la ciencia ficción.

Pero, tras la sorpresa inicial, lo que descubrimos es un campo de investigación científica que parecía abrirnos camino a un nuevo concepto de realidad en la que se manejaba la posibilidad de que, en un futuro no muy lejano, las barreras entre las máquinas y los seres humanos desaparecieran. Las máquinas llegarían a ser, aunque brevemente³, tan inteligentes como los seres humanos que, a su vez, podrían migrar sus contenidos mentales, su consciencia y habitar en el interior de ellas, tal vez en busca de una inmortalidad ¿tecnológica?

En realidad, lo más impactante fue, quizás, el hecho de que prestigiosas figuras del pensamiento científico, tecnológico y filosófico participaran en este debate, algunos negando toda posibilidad a la Singularidad Tecnológica y otros tan convencidos de su proximidad en el tiempo que manifiestan estar modificando sus hábitos de vida para, al vivir durante más años, conseguir llegar a ella.

Es necesario aclarar que, si bien los conceptos que manejamos son cercanos entre sí, no son totalmente equivalentes, pues por una parte está implicada la Inteligencia Artificial en su aspecto más fuerte, concepto que será aclarado más adelante, y por otro el objetivo a conseguir es el desplazamiento de una mente humana desde su entorno natural al interior de una

² En este documento hablaremos de máquinas en su acepción más amplia, como dispositivo capaz de realizar funciones de cálculo o cómputo y como sinónimo de computador (*computer* en la literatura anglosajona) y de ordenador (*ordinateur* en el entorno francófono).

³ Dado que, una vez alcanzado los niveles de inteligencia y autonomía del ser humano, podrían mejorarlos de manera continua y, probablemente, exponencial.

máquina, de un ordenador, donde pueda residir permanentemente una vez realizado el proceso de migración.

A lo largo de este documento nos acercaremos a la Singularidad Tecnológica a través de la Inteligencia Artificial para que, una vez la hayamos definido desde un punto de vista general, podamos aproximarnos al concepto de Inteligencia Artificial General, y posteriormente entrar en profundidad en la idea de Inteligencia Artificial Fuerte, como condición previa para la posibilidad de llegar a la Singularidad Tecnológica. A continuación, abordaremos todo este desarrollo bajo el punto de vista filosófico, dadas las importantes cuestiones éticas, morales y de relación con el concepto de lo humano, qué lo caracteriza y, finalmente, determinar la proximidad que pueda llegar a mantener con posibles entidades inteligentes que puedan ser creadas mediante procesos computacionales.

Es nuestra intención acercar al lector a este campo de conocimiento y reflexión sobre algo tan presente y a la vez tan oculto en nuestra realidad cotidiana, siempre dentro de las limitaciones que la naturaleza de este documento nos impone

2. La definición de Inteligencia Artificial

Aunque no existe una definición concreta de Inteligencia Artificial, la mayoría de los autores coinciden en que, en esencia, consiste en lograr que una máquina tenga inteligencia propia, definición que admite muchas aproximaciones conceptuales diferentes, pues mientras para algunos autores la Inteligencia Artificial es el estudio de cómo hacer que los ordenadores hagan cosas que, en estos momentos, hace mejor el hombre, para otros, es una ciencia que intenta la creación de programas para máquinas que imiten el comportamiento y la comprensión humana, que sean capaces de aprender, reconocer su entorno y pensar. En síntesis, sea cual sea el concepto, la Inteligencia Artificial está orientada a conseguir que las máquinas realicen trabajos donde se aplique la inteligencia, el razonamiento y el conocimiento de un ser humano.

Si pretendiéramos hacer un esbozo de la historia de la Inteligencia Artificial podríamos remontarnos al *Mito del Madero Parlante*, dispuesto por Atenea en la nave Argos (Carrasquilla, 2016), a las *Estatuas Parlantes* del medievo, al *Turco*, el supuesto androide de Von Kempel, del que se afirma que jugó al ajedrez con Napoleón, o al *Motor Analítico* de Charles Babbage que calculaba logaritmos. Sin embargo, todos estos intentos, ya fueran fraudulentos o reales, resultaron infructuosos, hasta que en 1943 Culloch y Pitts propusieron un modelo de neurona del cerebro humano y animal. Estas neuronas del sistema nervioso, simuladas informáticamente, proporcionaron una representación simbólica de la actividad cerebral. Nbert Wiener tomó estas y otras ideas y las elaboró dentro de un mismo campo que se llamó *Cibernética*, a partir de cual nacería, la Inteligencia Artificial.

Fue en el Congreso de Dartmouth⁴ en 1956 en el cual se llegó a la definición de los presupuestos básicos del núcleo teórico de la Inteligencia Artificial, estableciendo como estructuras conceptuales básicas de la naturaleza del pensamiento las siguientes suposiciones sobre su naturaleza y propiedades:

Puede ocurrir fuera del cerebro, es decir, en máquinas.

Puede ser comprendido de manera formal y científica.

La mejor forma de entenderlo es mediante simulación en computadoras digitales.

Sin embargo, el auténtico comienzo de la carrera por la IA debemos situarlo en el artículo seminal de Alan Turing, *Maquinaria, computadora e inteligencia* (Turing, 1950), publicado en 1950 y recogido en 1963 por Alan Ross, doctor en Filosofía, en su obra *Controversia sobre mentes y máquinas* (Ross, 1984) donde define la forma de determinar si una máquina piensa, a través de una prueba de imitación mediante interacción con examinadores y examinandos humanos, conocida desde entonces como *Prueba de Turing*,

⁴ Minsky, McCarthy y Shanon (Universidad Dartmouth College, Hanover, Nuevo Hampshire, Estados Unidos de Norteamérica). Es posible consultar información al respecto en el enlace siguiente: <https://home.dartmouth.edu>

utilizando lo que se daría en llamar *Maquina Universal de Turing*⁵, considerada actualmente como base de estandarización para la mayor parte de los ordenadores,

Controversia sobre mentes y máquinas adelanta en casi medio siglo la discusión previamente citada sobre la Singularidad Tecnológica del anteriormente mencionado magazine *Spectrum* del IEEE, pues nueve autores, incluyendo a Alan Turing y al propio Alan Ross, se posicionan a favor o en contra de la posibilidad de la existencia de la Inteligencia Artificial.

Turing nos muestra un recorrido sobre las afirmaciones realizadas en ese momento para evidenciar la imposibilidad de la Inteligencia Artificial que, sorprendentemente, podrían verse reflejadas en cualquier artículo publicado en la actualidad (Ross, 1984: 25-41):

La objeción teológica: El pensamiento es una función del alma inmortal del hombre. Dios ha dado un alma a cada hombre y mujer, pero no a ningún animal o máquina. Por tanto, ningún animal o máquina pueden pensar.

La objeción del avestruz: Las consecuencias de que las máquinas puedan pensar serían horribles. Creamos y esperemos que eso no será nunca posible.

La objeción matemática: Por ejemplo, los Teoremas de Incompletitud de Gödel, que afirman que, en cualquier sistema matemático coherente, es decir, sin contradicciones lógicas, y lo suficientemente potente, pueden formularse afirmaciones que no pueden demostrarse ni refutarse dentro del propio sistema.

El argumento de la consciencia: No es suficiente que una máquina sea capaz de reflexionar o de sentir, es necesaria además que pueda saber que lo hace. Ningún mecanismo será capaz de tener alegría por sus logros ni tristeza por sus fracasos.

El argumento de las incapacidades diversas: Si bien es posible que se construyan máquinas que haga esto y lo otro, seguro que nunca se construirán máquinas que logren hacer otras cosas, tales como apreciar un buen helado, enamorarse, etc.

⁵ Máquina abstracta, síncrona y determinista, con un número finito de estados internos, con un conjunto finito de caracteres que pueden generar un conjunto infinito de palabras y operando símbolo a símbolo, es decir, de manera secuencial.

La objeción de Lady Lovelace: Las máquinas sólo pueden hacer aquello que se les ordene, nunca podrán crear nada, sólo realizaran lo que les ordenemos.

El argumento de la continuidad del sistema nervioso: Al ser el sistema nervioso una estructura de carácter continuo nunca podrá ser simulada por dispositivos de carácter discreto, como computadores digitales.

El argumento de la informalidad de comportamiento: No se puede elaborar un conjunto de reglas que permitan describir lo que una persona hace en todas las circunstancias posibles.

Podemos añadir algunos argumentos más, incluido el que las máquinas nunca podrán tener *percepción extrasensorial* (?), aquello de que las máquinas son feas, unos trastos que nunca podrán hacer nada por iniciativa propia, etc.

La gran variedad de objeciones, y réplicas a las mismas, ante la posible existencia de inteligencias de génesis artificial es muestra de la gran atención con que se sigue este proceso, tanto por parte de la Ciencia como de la Filosofía, estando siempre presente, además, en el imaginario popular, en multitud de obras literarias, cinematográficas, juegos de ordenador, etc.

A todas estas objeciones, y a algunas más, da respuesta Alan Turing en el mencionado artículo *Maquinaria, computadora e inteligencia* al que, dadas las limitaciones de este documento, dirigimos al lector para un más amplio conocimiento, si bien dada la amplia difusión de trabajos y publicaciones actualmente existente sobre Inteligencia Artificial, es posible seguir argumentaciones a favor y en contra en casi cualquier publicación tanto de tipo específico como general, incluidas revistas y programas audiovisuales.

Aunque podríamos extendernos sobre la historia más reciente de la Inteligencia Artificial vamos a analizar los distintos tipos existentes.

3. Las variantes de la Singularidad Tecnológica

Tal y como hemos indicado con anterioridad, la Singularidad Tecnológica presenta dos aspectos muy relacionados, pero de diferente naturaleza

y con diferentes repercusiones, en su sentido ético, político, económico y filosófico.

En su primer aspecto tratamos todo lo referente al proceso de dotar de inteligencia humana, o incluso superior a la humana, a una máquina, y en el segundo nos acercaremos a la posibilidad de migración de nuestra mente, y consciencia, al interior de una de ellas.

3.1 La Inteligencia Artificial débil o estrecha y la Inteligencia Artificial fuerte o general

Fue John R. Searle quien propuso los conceptos de Inteligencia Artificial débil y fuerte. La Inteligencia Artificial débil es aquella inteligencia artificial racional que se centra en una tarea concreta y determinada, y abarca todos los automatismos que ejecutan tareas de un ámbito concreto y con una percepción y actuación sobre un entorno estrecho y limitado, tales como los robots industriales que fabrican total o parcialmente nuestros bienes de consumo, bienes industriales, automóviles, etc. Por ello recibe frecuentemente la denominación de Inteligencia Artificial estrecha, pues las máquinas podrían actuar como si fuesen inteligentes y simulando serlo, podrían superar a sus programadores, en campos restringidos debido a su gran velocidad de cómputo y capacidad de memoria, pero siempre en campos concretos de actuación y sin llegar a poseer capacidad alguna de conocimiento general de su entorno, sino sólo el necesario para ejecutar las tareas encomendadas.

En cambio, los defensores de la Inteligencia Artificial fuerte afirman que todo lo que caracteriza a la mente es poseer un programa. Esto nos llevaría a la conclusión de que las máquinas podrían llegar realmente a pensar, exactamente igual que el humano. Bastaría con implementar el tipo adecuado de programas para que puedan pensar en forma real y tener conciencia de lo que están haciendo. Para ello se programarían rutinas y mecanismos autorreflexivos, de autoanálisis, autocontrol y autocrítica de lo que están ejecutando, conducentes a la autocorrección y a la optimización de su tarea,

para lo cual deben poseer una visión del mundo como los humanos, lo que habitualmente denominamos sentido común.

Es evidente que nuestro campo de atención se debe centrar en la Inteligencia Artificial fuerte, en particular en la Inteligencia Artificial general, en la que nos van a surgir las cuestiones de tipo filosófico que intentaremos abordar a lo largo de este documento.

En efecto, la posibilidad de crear mentes de naturaleza artificial que puedan evolucionar por sí mismas hasta alcanzar lo que se ha venido en denominar *superinteligencia*, tiene una proyección de suficiente alcance como para provocar desenlaces tan felices como una humanidad prácticamente liberada del trabajo necesario hasta ese momento para su supervivencia, al igual que una entidad superinteligente podría, por sí misma o guiada por terceros, ejercer un dominio total y absoluto sobre esa humanidad, responsable en última instancia de su creación y pudiendo ser causa, finalmente, de su destrucción.

3.2 La Singularidad Tecnológica como migración de la mente a una máquina

Es dentro del campo de la Inteligencia Artificial fuerte donde podemos hablar de la Inteligencia Artificial General y de la otra variante de la Singularidad Tecnológica, aquella que defiende que cuando las máquinas alcancen el nivel de inteligencia del ser humano, o superior a él, podremos migrar nuestros contenidos mentales a una de ellas, participando por tanto de las ventajas, actuales o futuras, que se suponen poseen.

Tendríamos aquí un problema añadido, pues para copiar algo necesitamos conocer perfectamente aquello que copiamos, no pudiéndose decir que en la actualidad poseamos ese nivel de conocimiento sobre nuestro cerebro, sobre la naturaleza de nuestra mente.

Es por ello que la mayor parte del esfuerzo realizado en este campo está situada sobre el estudio y análisis del cerebro, tanto del humano como el de otros seres vivos, su estructura y funciones y, aunque se producen constantemente nuevos avances y se obtienen grandes conocimientos, no parece

que podamos tener, en un futuro muy próximo, el nivel de conocimientos necesarios para realizar la integración de una mente humana en una máquina.

Si ese proceso de migración pudiera realizarse, esto supondría la posibilidad de hacer copias o respaldos⁶ de nuestra mente, realizar mejoras de nuestra velocidad de razonamiento, parejas a las posibles mejoras del soporte físico informático en el que residiríamos, alcanzar enormes capacidades de almacenamiento de información, en definitiva de conocimiento y, naturalmente la inmortalidad, porque bastaría con ir migrando de una máquina a otra de superiores capacidades para tener una *vida eterna*. Incluso en el caso de un accidente o acto mal intencionado que interrumpiera el funcionamiento de la máquina en la que residiríamos, bastaría con acudir a la copia de respaldo para volver a la vida.

Pero ¿cómo determinar si nuestra máquina, además de inteligente, es consciente? ¿Necesariamente inteligencia implica consciencia?

Podría suceder que, si traspasamos nuestros contenidos mentales a una máquina supuestamente capaz de recibirlos, en realidad estemos cometiendo suicidio, porque nuestra máquina no fuera capaz de sustentar consciencia alguna.

Y si, por precaución, migráramos una copia de nuestros contenidos de forma satisfactoria, sin destruir el original ¿dónde residiría nuestro yo, nuestro ser consciente?

Podríamos además considerar que nuestra mente, una vez migrada al interior de una máquina, pasaría a ser pura información de nuestros contenidos mentales y nuestra consciencia, por lo cual no tendría por qué estar localizada en un lugar concreto y podría, al igual que sucedería con una posible *superinteligencia*, habitar simultáneamente en muchas máquinas e incluso habitar en una red de información distribuida como, por ejemplo, Internet, lo cual nos llevaría prácticamente a conceptos como la omnipresencia y la omnisciencia para una mente en esa situación, pues tendría

⁶ Denominados habitualmente *Backups* en lenguaje informático.

acceso a prácticamente todo el conocimiento de lo que ocurriera en cualquier lugar y su capacidad de almacenar información podría ser prácticamente ilimitada, al depender directamente de las posibilidades de ampliación de su soporte físico o *hardware*.

Esta estructura hipotética ha dado origen a la idea de lo que se ha dado en denominar *God Equivalence Hypothesis*⁷, estructura hipotética de carácter recursivo que sostiene la posibilidad de que vivamos en una simulación informática generada por una mente cibernética que, residiendo inicialmente en una máquina, haya sido capaz de desarrollar totalmente su potencialidad y conocimiento hasta el punto de poder controlar el espacio y el tiempo, lo cual le permitiría, entre otras muchas cosas, cambiar nuestra propia historia, situándose en el origen de los tiempos, es decir, en el origen de nuestra realidad.

Dejamos a la curiosidad del lector la búsqueda de información complementaria al respecto.

4. La función de la Filosofía

Ante lo que, de manera resumida, hemos visto sobre la Singularidad Tecnológica y la Inteligencia Artificial, las preguntas que podríamos hacernos son:

¿Tiene algo que aportar en este proceso la Filosofía?

¿Cómo debe realizarse la aproximación a estructuras del pensamiento que aparentemente sólo incumben al ámbito de las ciencias?

La respuesta a la primera pregunta no ofrece dudas, pues la Filosofía, desde los más antiguos pensadores, siempre se ha ocupado de la mente, espíritu o alma del ser humano y de sus relaciones con su sustrato corporal o físico.

⁷ Puede consultarse información al respecto en el siguiente enlace: <https://www.simulation-argument.com/simulation.html>

Pero la segunda cuestión no tiene una respuesta tan inmediata y directa, pues parecen predominar las interpretaciones sustentadas en la metafísica, a veces con mucha más *doxa* que *episteme*.

Si siguiendo, por ejemplo, a Wilfrid Sellars, consideramos que (Sellars, 1963:1-3)

El propósito de la Filosofía, formulado en abstracto, es comprender cómo las cosas, en el sentido más ampliamente posible del término, se relacionan entre sí, en el sentido más amplio que sea posible del término.

Y si asumimos además que:

Para conseguir los objetivos que se pretenden mediante el análisis filosófico de algo, es necesario conocer tanto su entorno como lo que se conoce previamente de ello.

Vemos entonces que, lo realmente importante de este proceso de análisis filosófico, es el cómo nos aproximamos al objeto en cuestión, y cuál es nuestra propia vía de acceso.

Y esta vía de acceso no implica que tengamos que conocer todo lo que se conoce de una disciplina científica determinada, lo importante es ver, analizar, conocer las estructuras que la unen a los otros conocimientos con los que se relaciona.

Aquí, en la visión amplia, interdisciplinar, en la aproximación comprensiva e inteligente, es donde la Filosofía puede actuar, aportando lo que el científico, centrado habitualmente en su campo específico, en principio, y salvo excepciones, no puede.

Podríamos preguntarnos si todo este proceso de incorporación de la IA a nuestra cotidianidad, como demuestran tantos y tantos anuncios de productos y tecnologías “inteligentes”, tiene realmente la capacidad de influir en el desarrollo de nuestras vidas, en nuestra forma de abordar la realidad.

Pues parece ser que sí, pues si la primera década de este siglo estuvo dedicada al cerebro, ahora la década actual, y posiblemente las siguientes,

van a estar dedicadas a la Inteligencia Artificial, a sus avances y las repercusiones que tendrá sobre nuestra sociedad y, sobre todo, sobre nuestra forma de pensar.

5. Los interrogantes filosóficos

Una vez que una máquina supera el test o prueba de Turing, es decir, sea indiscernible cualquiera que sea su naturaleza, pues bajo cualquier punto de vista o análisis físico, conductual o intelectual no podamos distinguirla entre nosotros, ¿cómo podremos negarle la condición de humana? ¿Bajo qué premisas podremos evitar ser arbitrarios en tal decisión?

Si a lo largo de nuestra historia a determinadas razas o etnias le ha sido negada la condición de seres humanos ¿qué sucederá con aquellos androides de última generación de los que no podamos afirmar de manera indubitable su condición de hombre o máquina?

Si todo esto llegara a suceder es evidente que en realidad es la propia naturaleza de lo humano la que se pone en juego, pues no será suficiente barrera de separación el hecho de que un intelecto similar al nuestro tenga un sustrato orgánico, mecánico o, en definitiva, electrónico.

Si las máquinas logaran alcanzar una inteligencia y autonomía intelectual igual a la de los humanos, esto sucedería durante un periodo muy breve, pues el intelecto natural o de base orgánica tiene tiempos de evolución y posible mejora muy largos, dependiente en general del proceso evolutivo de la especie, mientras que la inteligencia de sustrato inorgánico tiene muy fácil incrementar sus potencialidades, sustituyendo procesadores, cambiando tecnologías de implementación de los sistemas informáticos, creando nuevas formas de programación, incrementando la memoria de una forma masiva, etc.

En efecto, si una máquina llegara a poseer el mismo nivel de inteligencia de los seres humanos, entre sus habilidades estaría la capacidad de diseñar otras máquinas con mejores características, en definitiva, con mayor inteligencia que ella misma que, en un proceso recursivo, serían capaces de

mejorar de nuevo los niveles de inteligencia de las sucesivas generaciones de máquinas, estableciéndose así un proceso de explosión literal en la inteligencia disponible.

Pero ¿seguirían siendo *nuestras máquinas*, o llegaría a invertirse la situación, pasando nosotros, los humanos, a ser *sus humanos*?

Quienes afirman que eso nunca podría llegar a ser posible, con afirmaciones tales como que las máquinas no podrían desarrollar objetivos propios pues nunca tendrían percepción subjetiva de sí mismas, es decir, consciencia, están siendo fuertemente contestados por los avances actuales en el estudio conductual de los sistemas automatizados, que afirman que no es necesaria la presencia de consciencia para que una máquina pueda fijar sus propios objetivos. Para un más amplio conocimiento de estas cuestiones remitimos al lector a la bibliografía incorporada al final del presente documento.

¿Son inteligencia y pensamiento términos equivalentes? J. Searle diseñó un experimento mental, habitualmente denominado la *habitación china de Searle*, en el que desde una habitación cerrada se recibían cuestiones y mensajes en lengua china, un operario especializado utilizaba un conjunto de instrucciones y algoritmos que le permitía entregar la respuesta correspondiente con precisión y también en lengua china, a pesar de desconocer totalmente dicha lengua. Desde el exterior parecería que el sistema entendía el chino y era capaz de actuar como eficaz operador para responder cuestiones en dicha lengua, pero no era más que pura apariencia. Igualmente, una máquina podría *simular* inteligencia, y comportarse *como si* tuviera mente, pero eso no implica necesariamente que pueda *pensar* realmente.

¿Cuál será el estatus legal, jurídico, laboral y social de nuestros robots andróides, aquellos humanoides que, parece ser, pronto pasarán a compartir con nosotros el planeta?

En una época en la que incluso los animales están protegidos legalmente contra los abusos y daños que le puedan ser infligidos, nuestras futuras máquinas pensantes, quizás incluso autoconscientes ¿estarán protegidas de

igual manera contra el abuso laboral, la esclavitud y la conculcación de sus posibles derechos?

Nuestros robots actuales, que ya fabrican buena parte de nuestros productos de uso cotidiano, lavadoras, neveras, vehículos, etc. no presentan, todavía, problemática alguna y, por sus limitadas capacidades y su apariencia no humana, los vemos bajo el punto de vista de la Inteligencia Artificial débil o estrecha.

Pero, a nuestros futuros asistentes personales, de configuración y comportamiento antropomórficos, que estarán mucho más cercanos a la Inteligencia Artificial fuerte, a la Inteligencia Artificial general ¿Qué trato les daremos?

¿Hablabremos de explotación sexual e incluso prostitución cuando se pongan a la venta androides como asistentes para actividades de tipo sexual? ¿O consideraremos que, por fin, aquellos que, por discapacidad de algún tipo, no puedan llevar una vida sexual satisfactoria, puedan estar atendidos en este aspecto? ¿Cuáles serían las objeciones religiosas, morales y éticas que habrían de ser tenidas en cuenta?

Y respecto al uso militar, drones, máquinas de uso militar y autónomas para el combate, cuyo uso está ya generalizándose ¿Tendrán reglas de corte ético en su funcionamiento autónomo?

Dado que la condición de humano se la otorgamos a quienes aparentemente son como nosotros y parecen pensar y actuar como nosotros, cuando nuestros androides pasen sin dificultad alguna la prueba de Turing y sean indistinguibles de cualquier humano ¿le daremos acceso a esa condición humana?

¿Consideraremos que existe explotación laboral cuando les hagamos trabajar durante el día y la noche, dado que pueden ser programados para no sentir cansancio? ¿Llegarán a tener derechos civiles?

Las repercusiones legales, económicas, laborales, éticas, incluso de defensa en el plano militar, que van implícitas en este proceso, no pasan inadvertidas para los poderes políticos en los aspectos domésticos y transnacionales, lo cual se puede advertir en las distintas publicaciones emitidas

en los EE.UU. de Norteamérica por el Pentágono⁸ y la Casa Blanca⁹ y, más cercanos a nosotros, por la Comunidad Europea¹⁰.

Como muestra valga la publicación nueve años después, en junio de 2017, de un nuevo número monográfico del magazine Spectrum del IEEE sobre este tema, cuyas líneas temáticas principales son, entre otras: “Aprendiendo a amar las máquinas que puedan pensar”, “Cómo implementar un cerebro artificial”, “¿Puede una máquina ser consciente?”, etc.

Pero ahora ya no se discute si se puede llegar a la Singularidad Tecnológica, se trazan las líneas del cómo se va a hacer, en realidad de cómo se está haciendo.

En los últimos cincuenta años hemos ido mucho más allá de la denominada *Ley de Moore*¹¹, cuya predicción de doblar cada 18 o 24 meses los niveles de integración de los dispositivos semiconductores, con los que se construyen las máquinas de cómputo, se ha cumplido escrupulosamente hasta ahora. Pero ahora estamos haciendo ingeniería inversa del cerebro, trabajamos con microchips integrados en tres dimensiones para ir más allá de la simulación informática de sistemas neuro-difusos¹² tipos I y II y, por fin, comenzamos el trabajo de construir sinapsis neuronales reales, no meras simulaciones de las mismas, gracias al *Memristor*¹³, elemento electrónico largamente esperado y que, por sus especiales características, nos permite

⁸ Información disponible en el enlace: <https://www.nytimes.com/es/2018/08/29/inteligencia-artificial-pentagono/>

⁹ Información disponible en el enlace: <https://www.voanoticias.com/a/inteligencia-artificial-en-la-casa-blanca/4385351.html>

¹⁰ Información disponible en el enlace: http://europa.eu/rapid/press-release_IP-18-3362_es.htm

¹¹ Información disponible en el enlace: <https://www.intel.la/content/www/xl/es/silicon-innovations/moores-law-technology.html>

¹² Información complementaria disponible en el enlace: <http://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/viewFile/3270/2733>

¹³ Información detallada en el enlace: <http://www.memristor.org/reference/research/13/what-are-memristors>

conseguir la implementación de estructuras cerebrales y neurológicas similares a las orgánicas o naturales.

Más allá de todo lo hasta aquí expresado, una de las principales cuestiones que tendremos que determinar es si se puede copiar algo si no sabemos a ciencia cierta qué es, pues seguimos sin conocer la naturaleza de lo que denominamos consciencia, y proseguimos la búsqueda y análisis de lo que es la inteligencia, más allá de determinadas definiciones de carácter utilitarista.

Y es, de nuevo, en el campo del pensamiento y la especulación filosófica donde podemos intentar avanzar en el conocimiento relativo a estos temas:

Siguiendo la corriente filosófica *reduccionista* ¿buscaremos el conocimiento mediante un proceso analítico que vaya desde el proceso global hasta cada una de las partes que lo componen?

¿O, por el contrario, nos apoyaremos en el *emergentismo filosófico*, considerando que el todo es algo más que las partes que lo componen?

6. Conclusiones

Basta con escuchar, leer, observar y, en definitiva, analizar lo que sucede a nuestro alrededor, para ser conscientes de que la Inteligencia Artificial en sus distintas vertientes va a ser, y de hecho está siendo, el principal motor de la economía y el crecimiento industrial, en sus vertientes civil y militar.

Es tal la importancia y repercusión que sobre nuestro destino tendrá la Inteligencia Artificial que están surgiendo importantes movimientos éticos, filosóficos, morales e incluso de carácter económico y político para conseguir lo que se ha dado en llamar Inteligencia Artificial *amigable* mediante el control de todo este proceso y, de esta forma, intentar evitar que lo que estamos realizando se convierta en una grave amenaza para el futuro de nuestra especie.

Si a todo esto añadimos la escasa posibilidad de limitar estos avances, pues nunca el ser humano ha dejado de hacer aquello que estaba a su alcance, parece quedar bien claro el posible valor mediador y de configuración

de la Filosofía en un evento que ha sido comparado con la Revolución Industrial en su capacidad de producir un giro fundamental en el entorno de lo humano.

7. Bibliografía empleada

N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Londres, Reino Unido, 2017.

J. Carrasquilla: *Mitología Griega Vol. III*, Cultiva Libros S.L. 2016

A. Cortina y M. Serra, ¿Humanos o posthumanos?: Singularidad tecnológica y mejoramiento humano, Fragmenta Editorial, Barcelona, 2015.

A. Ross Anderson, *Controversia sobre mentes y máquinas*, Cuadernos Ínfimos 124, Tusquets Ed. 1984.

W. Sellars *Philosophy and the Scientific Image of Man*, documento obtenible en el enlace: <http://selfpace.uconn.edu/class/percep/SellarsPhilSciImage.pdf>

M. Shanahan, *The Technological Singularity*, The MIT Press Essential Knowledge Series, Cambridge, Massachusetts, EE.UU., 2015.

Singularity Now ¡The Artificial Intelligence Timeline: Automation, Tech Titans, and Their New Brave World, Ed. Independientes, Amazon EU, febrero 2018.

M. Tegmark, *Vida 3.0: Qué significa ser humano en la era de la inteligencia artificial*, Taurus, Madrid, 2018.

A. Turing, *Minds and machines*, Mind, Vol. LIX, nº 236, 1950

Todos los enlaces telemáticos han sido confirmados y están activos a la hora y fecha de finalizar este documento, enero de 2019.

José Manuel Elena Ortega
jmelena@us.es