

Inteligencia artificial, autoconciencia y derechos humanos de los sujetos artificiales

Artificial intelligence, self-awareness and human rights of artificial subjects

Flavio Suárez-Muñoz 

Universidad de Guanajuato, México

Correspondencia: milenioflavio@hotmail.com

RESUMEN. La inteligencia artificial como tecnología que imita el comportamiento humano, ha tenido grandes avances gracias al desarrollo tecnológico que sirve de plataforma para el procesamiento de grandes cantidades de información, esto permite que la tecnología conozca a las personas de manera detallada y con ello reproducir emociones y sentimientos de manera artificial. De modo que esta tecnología ha superado a las personas en capacidades como las matemáticas, la predicción y la toma de decisiones, no obstante, la lingüística representa una limitante para esta tecnología, ya que la semántica que es la que da sentido a las palabras, aún no ha sido superada del todo por las máquinas. Por otro lado, la autoconciencia es intrínseca de las personas, esta está impresa de manera biológica en el cerebro para poder ser conscientes del entorno, esta se potencia con la interacción con el entorno, pero a la inteligencia artificial le falta esta red biológica que detona la conciencia interna, por ello, no puede ser autoconsciente, esto nos lleva a que los intentos por reconocerles derechos humanos a estas tecnologías, no prosperen, al menos no en el corto plazo.

Palabras clave: inteligencia artificial, conciencia, derechos humanos, sujetos artificiales, LaMDA.

ABSTRACT. Artificial intelligence as a technology that imitates human behavior, has had great advances thanks to technological development that serves as a platform for the processing of large amounts of information, this allows technology to know people in detail and thereby reproduce emotions and feelings artificially. So, this technology has surpassed people in abilities such as mathematics, prediction and decision making, however, linguistics represents a limitation for this technology, since the semantics that is what gives meaning to words, has not yet been completely overcome by machines. On the other hand, self-awareness is intrinsic to people, this is biologically imprinted in the brain to be aware of the environment, this is enhanced by interaction with the environment, but artificial intelligence lacks this biological network that detonates internal consciousness, so it cannot be self-conscious, this leads us to, attempts to recognize human rights to these technologies will not succeed, at least not in the short term.

Keywords: artificial intelligence, conscience, human rights, artificial subjects, LaMDA.

Recibido: 01/12/2022 Aceptado: 23/01/2023



1. Introducción.

La inteligencia artificial es una de las tecnologías más estudiadas e investigadas en los últimos años, pero cuando hablamos de inteligencia artificial, debemos partir del concepto de *inteligencia*, en general, término que parece haber perdido interés por los investigadores, no obstante, el acercamiento al concepto de inteligencia permite entender que la intención de la inteligencia artificial, es imitar el comportamiento humano y realizar cosas que si las hicieran los humanos parecerían inteligentes, en tal sentido, cabe resaltar que en esa imitación existen elementos que hacen parecer a las máquinas como sujetos consientes, pueden imitar emociones y sentimientos, no porque en realidad tengan autoconciencia o porque tengan sentimientos, lo que tienen en realidad son procesadores veloces, memoria RAM y acceso a grandes cantidades de información.

El acceso a esta información permite conocer a detalle a las personas, con ese conocimiento la inteligencia artificial es capaz de imitar al ser humano, de moldear su conducta, pero no de tener autoconciencia, entendida esta como la combinación de las redes neuronales biológicas de la persona, que permitan al individuo ser consciente de su entorno y actuar en consecuencia para adaptarse a su entorno social y cultural, al que Bartra (2014) llama el *exocerebro*, mismo que le da a la persona el conocimiento para actuar de manera consciente ante las circunstancias que el entorno le presenta.

Por otro lado, la interacción con estas tecnologías repercute en los derechos humanos, por ello, es importante conocer a que se refieren estos derechos humanos y, de esta manera poder adoptar una postura en cuanto a los sujetos artificiales que, algunos investigadores proponen otorgarles derechos humanos por considerarlos como sujetos artificiales capaces de imitar sentimientos humanos, no obstante, esta postura parece no considerar que la premisa de los derechos humanos es la dignidad *humana*. Resaltemos la palabra *humana*, puesto que es la condición principal para ser sujeto de derechos humanos, así, la categoría de sujetos artificiales no sintientes no logra alcanzar el estatus humano como para gozar de estos derechos, aunque no se descarta que sí tengan derechos, más que para gozar de ellos, para atribuir obligaciones a los actos realizados por la inteligencia artificial.

2. Metodología.

El presente trabajo es cualitativo, ya que no aborda encuestas ni entrevistas, sino más bien textos y declaraciones, así como interpretaciones, para ello, se apoya en los métodos inductivo y hermenéutico. Mediante el método inductivo se busca tomar como referencia una inteligencia artificial que recientemente ha desatado polémicas en la industria de las tecnologías, sobre todo, después de que Blake Lemoine (2022) afirmara que ésta había cobrado conciencia de sí misma, con ello se pretende lograr un entendimiento claro de la inteligencia artificial como *sujeto artificial no sintiente*, sin emociones ni razonamiento, concepto que, desde el punto de vista del autor, aplica para toda inteligencia artificial.

Por otro lado, el método hermenéutico se aplica para hacer una interpretación de las declaraciones, así como de los conceptos que corresponden a la autoconciencia, inteligencia artificial, inteligencia general, así como a la privacidad, vista esta última desde un punto de vista, no solo jurídico, sino, desde el punto de vista de la protección de la integridad de las personas en un entorno conectado que demanda la circulación de información y el conocimiento de las personas que interactúan en el entorno digital.

3. Resultados.

3.1. Inteligencia artificial.

El término de inteligencia artificial no es nuevo en el campo del desarrollo de las tecnologías, no obstante, en sus inicios no se contaba con las capacidades tecnológicas que permitieran el procesamiento a gran escala de la información, por lo tanto, no fue explotada hasta años recientes, cuando, aplicando tecnologías de *BigData* fue posible acceder y recuperar grandes cantidades de información, fue entonces cuando la inteligencia artificial alcanzó su máximo esplendor, por ser una herramienta capaz de procesar grandes volúmenes de información y obtener resultados mucho más rápido y más precisos que un ser humano. Pero antes de pasar al entendimiento de la inteligencia artificial, es

pertinente retomar el concepto de inteligencia en una concepción mucho más generalizada, que si bien, la inteligencia es entendida en general, como la capacidad que tenemos las personas para aprender de nuestro entorno y actuar de manera *inteligente* para lograr un desarrollo personal que conduzca a la felicidad. De forma simple esto se concreta en la capacidad humana para procesar información, aprender y actuar en consecuencia.

Para Terman (1965) la inteligencia es la “capacidad de razonar abstractamente”, capacidad que hasta hace algunos años solo podía ser atribuida a las personas, cuestión que la ciencia, concretamente la rama de investigación en inteligencia artificial se ha empeñado por superar, y ha fijado sus ambiciones científicas en el desarrollo de máquinas capaces de analizar, razonar y empatizar. Aunque aún no se ha logrado esto a un nivel comparado con la capacidad del cerebro humano, si hay grandes avances que han llevado a pensar en la necesidad de reconocerles derechos a estas inteligencias artificiales, bajo un estatus de “sujetos artificiales no sintientes” (Rico, 2019).

No obstante, aunque los avances de la inteligencia artificial han superado las capacidades matemáticas de los seres humanos, cuando se trata del lenguaje natural y sobre todo de la comprensión semántica que da significado a las palabras, la tecnología aún presenta capacidades limitadas, aunque no se descarta la posibilidad de superar esta limitante y lograr la construcción de una inteligencia artificial general, que sea capaz de desarrollar su red neuronal de la misma forma que lo hacen las personas, así, esta tecnología sería capaz de aprender cualquier cosa y no solo se ocuparía de funciones específicas para las que haya sido programada. Pero de momento, parece pertinente quedarnos con la propuesta de que las computadoras, aún presentan problemas para procesar el complicado lenguaje de los humanos, y superar esta barrera es lo que mantiene a la inteligencia artificial al servicio de las personas, sea con buenas o perversas intenciones.

Por otro lado, Chantraîne (1968, citado por Martín, 2003), menciona que la inteligencia es “ver, observar, percibir, pensar, reflexionar, comprender, entender, saber, considerar, meditar, proyectar, etc.” (p. 34), lo que implica varias capacidades del ser humano, esto podría invalidar la primera aproximación que se propone como concepto de inteligencia, ya que no solo se basa en la capacidad para el procesamiento de la información para la toma de decisiones. David Wechsler (1944, citado por Universidad de Alicante, s.f.), define la inteligencia como la “capacidad para actuar con un propósito concreto, pensar racionalmente y relacionarse eficazmente con el ambiente”, esta definición supone que como ser social, debe relacionarse de manera eficiente, lo que implica también que esa capacidad para actuar de manera racional, debe tener en cuenta los principios éticos y morales, de tal modo que se conduzca con inteligencia por el camino del bien colectivo que toda sociedad pretende para alcanzar su felicidad.

De igual modo, Papalia y Wendkos-Olds (1996), conciben la inteligencia como la capacidad de interacción activa entre las capacidades heredadas y las experiencias ambientales, cuyo resultado capacita al individuo para adquirir, recordar y utilizar conocimientos, entender conceptos concretos y abstractos, comprender las relaciones entre los objetos, los hechos y las ideas y aplicar y utilizar todo ello con el propósito concreto de resolver los problemas de la vida cotidiana. La RAE muestra varios conceptos que nos acercan a la definición actual de lo que se conoce como inteligencia, entre ellas se define como: “Capacidad de entender o comprender; Capacidad de resolver problemas; Conocimiento, comprensión, acto de entender; Habilidad, destreza y experiencia” [por mencionar los más relevantes conforme al tema] (RAE, consultado el 15 de noviembre de 2021).

Considerando las definiciones anteriores sobre la inteligencia, se puede entender a la inteligencia artificial como la imitación de las capacidades humanas por las computadoras, dicho de otra manera, es la imitación del ser humano por parte de los algoritmos de inteligencia artificial. De esa manera, las computadoras pueden presentar mejores habilidades que los seres humanos en algunas de esas capacidades, por ejemplo, en matemáticas las computadoras superan por mucho a las personas, pero con relación al sentido en que se puede tomar una proposición, un dicho o una expresión, aún presentan deficiencias con respecto a los humanos, no obstante, es un campo en el que hay numerosos estudios que buscan lograr dicho cometido (de Torres, 2021; Investigación y Desarrollo, 2021; Vashev, 2021).

De modo que en el intento por definir lo que es la inteligencia artificial, es preciso retomar a Nils John Nilsson (1998), considerado uno de los padres de la inteligencia artificial, quien dice que esta tecnología “tiene por objeto el

estudio del comportamiento inteligente en las máquinas. A su vez, el comportamiento inteligente supone percibir, *razonar*, aprender, comunicarse y actuar en entornos complejos” (p. 1). Vemos que en este concepto que presenta Nilsson, se engloba la mayoría de las capacidades de la inteligencia humana. Asimismo, refiere que “una de las metas a largo plazo de la IA [Inteligencia Artificial] es el desarrollo de máquinas que puedan hacer todas estas cosas igual, o quizá incluso mejor, que los humanos” (p. 1).

Ya en la actualidad, el ser humano se ha visto superado en algunos campos, por ejemplo, percibir, aprender y comunicarse, ya que mediante sensores pueden percibir cosas que los seres humanos no somos capaces y que solo se ha logrado mediante la implantación de chips de IA, que permiten que el ser humano perciba lo que biológicamente no es posible, mientras tanto, en el aprendizaje y comunicación, se sigue trabajando para mejorar dichas capacidades en la inteligencia artificial.

El test de Turing (1950), mostró que es posible dotar a las máquinas de una capacidad de imitación del ser humano, al grado de que una persona podría interactuar con la inteligencia artificial sin darse cuenta. De modo que McCarthy (2007), afirma que la inteligencia artificial “es la ciencia e ingeniería para construir máquinas inteligentes, especialmente, programas de computación inteligentes. Así como, lo relativo a la tarea de usar computadoras para entender la inteligencia humana (...)” (p. 2). De esta concepción se rescata la idea de la construcción de *computadoras para entender la inteligencia humana*, que como lo demostró Turing, ese entendimiento permite a las máquinas imitar a los seres humanos.

Pero estas capacidades de las computadoras y del software ¿pueden hacer que las máquinas sean inteligentes?, Minsky (1990) menciona que lo primero que se tuvo que hacer para responder esta pregunta, fue darse cuenta de que las máquinas eran incansables y se estaban volviendo rápidas. Entonces, si tuviera un cierto tipo de problema, que podría ser difícil para una persona, sería fácil para una computadora resolverlo. Por otro lado, menciona que aún no se ha llegado a conocer cómo los cerebros realizan sus habilidades mentales, si se puede trabajar para que las máquinas hagan lo mismo. En tal sentido, menciona que la investigación sobre la inteligencia artificial se reduce a la aplicación de métodos para hacer que las máquinas actúen de manera que se consideren inteligentes.

López (2007) además de presentar una serie de definiciones de distintos autores, menciona que la Inteligencia Artificial es una rama de las ciencias computacionales, esta se encarga de estudiar modelos de cómputo capaces de realizar actividades propias de los seres humanos, con base a dos de sus características primordiales: el razonamiento y la conducta. Esta afirmación presenta dos términos que aún siguen sin resolverse de manera precisa, el *razonamiento y la conducta*, ya que, a nivel neuronal, aún se desconoce qué es lo que hace que una persona razone de tal o cual manera ante circunstancias concretas, mientras que la conducta, depende de tal razonamiento para actuar en consecuencia de manera razonada y consciente.

La Real Academia de la Lengua Española define la inteligencia artificial como la “disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico” (RAE, consultado el 15 de noviembre de 2021).

Al decir que se trata de una disciplina científica, resulta relevante la declaración de Palma y Marín (2008), quienes mencionan que la inteligencia artificial vista como ciencia, “busca una teoría computable del conocimiento humano” (p. 7). Entendida de esta manera, queda claro que las pretensiones de los investigadores en inteligencia artificial, han enfocado sus esfuerzos en replicar las funciones del cerebro en una tecnología, con la intención de darle la capacidad para actuar de manera inteligente, y “el sistema inteligente debe caracterizarse con la capacidad de cálculo adaptación, tolerancia a fallos, alta velocidad computacional y menos propenso a errores” (Konar, 2005, p. 5).

En una definición más reciente que propone Rouhiainen (2018) “la IA [Inteligencia Artificial] es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano” (p. 17). Asimismo, menciona que debemos ser conscientes de las mejoras significativas de las computadoras, que les permiten aprender, tomar decisiones y realizar tareas que antes eran propias de los humanos, algo que si bien, ha beneficiado a las personas, el enfoque principal que se le está dando, es como ventaja competitiva

para las empresas, por lo que debemos tener cuidado debido al impacto social que esto puede tener si se deja del lado la ética en el uso de la inteligencia artificial.

Después de revisar las definiciones de inteligencia e inteligencia artificial, es posible darse cuenta que coinciden en que es una capacidad para aprender, razonar, procesar información, así como la solución de problemas. En ese sentido podría entenderse la inteligencia artificial como la incorporación de capacidades de procesamiento de información, razonamiento lógico matemático y toma de decisiones en las computadoras.

En tal sentido, parece que el término inteligencia artificial es solo una manera de llamar a esa rama de la investigación en las ciencias computacionales, la cual ha dado como resultado la creación de dispositivos llamados inteligentes. Con todo, aunque se puede decir que las computadoras son inteligentes por realizar actividades de manera autónoma y con gran precisión, al mismo nivel o aún mejor que los humanos; parece que en realidad se trata de capacidades de procesamiento de información, y que, ayudadas por distintos dispositivos, pueden recolectar información del entorno, la procesan, analizan y generan un resultado que parece producto de una acción inteligente.

3.2. Aproximaciones a la autoconciencia.

Con la intención de tener un acercamiento que nos permita comprender el origen de la autoconciencia, y así buscar una posible inserción de esta en la inteligencia artificial, partiremos desde la postura de la autoconciencia como una característica humana que determina las emociones de los individuos, y en tal sentido, el comportamiento también deriva de sus emociones.

La autoconciencia es para Jiménez (2005) una capacidad que diferencia al ser humano de otras especies, y que les da la capacidad a los individuos para auto comportarse. Esto está ligado a las emociones, sensaciones y sentimientos que experimentan las personas y que ante tales emociones se desencadenan acciones, mismas que en esa autoconciencia son reguladas de manera consciente; al sabernos y reconocernos como personas, nuestras acciones deben ser reguladas según nuestra propia conciencia, lo que hagamos debe ser acorde a las normas morales, debe ser humano, nos debe conducir a una vida feliz, es decir, no actuar de tal manera que haya culpas que perturben nuestra conciencia e impidan la felicidad.

Bartra (2014) va un poco más lejos al decir que la conciencia de las personas es en parte, la construcción de una red neuronal biológica que se moldea por el entorno social, no obstante, esa red neuronal que permite dar cuenta de la existencia de la persona, es la que detona el impulso para que se actué en consecuencia de manera consciente, así, el *exocerebro* al que refiere Bartra, es el entorno social y cultural que da forma a la conciencia de las personas de una sociedad, es por ello que la ética y la moral, cambia de acuerdo a las regiones y a la cultura debido a que, lo que en una sociedad en específico es aceptado moralmente y se inserta en la conciencia de las personas, puede ser inmoral para otro grupo social, de modo que la conciencia es detonada de manera interna en el cerebro y moldeada por el entorno social y cultural.

Por otro lado, la Real Academia Española de la Lengua, define la autoconciencia como la “conciencia de sí mismo”, es decir, nos da la conciencia de que existimos y nos reconocemos como seres sociales, parte de una sociedad en la que tenemos interacciones con otras personas, de modo que afirma la postura de Bartra (2014), si se considera que esa conciencia de la existencia, es el impulso interno que despierta la conciencia, ese detonante permite actuar en consecuencia con el entorno social. Asimismo, menciona Marleno (2004), que se pueden identificar cinco niveles de conciencia:

a) Identidad corporal. Yo soy mi cuerpo. Dificultad de mirar como pares a otros. Cada ser humano es una isla. Se le podría denominar identidad CORPOCÉNTRICA.

b) Identidad grupal. Repitiendo una frase atribuida a Jorge Eliécer Gaitán, político colombiano asesinado en 1948. “Yo no soy un hombre, soy un pueblo.” Soy lo que mi grupo es. Mis pares son aquellos con los que comparto una identidad grupal, normalmente vinculada a un territorio y/o una cultura. Se le podría denominar identidad ETNOCÉNTRICA O SOCIOCÉNTRICA.

c) Identidad personal. Asociada con los propios pensamientos y las emociones derivadas de los mismos. Yo soy mis senti-pensamientos. Mis iguales son los que senti-piensen como yo. Se le podría denominar identidad EGOCÉNTRICA.

d) Identidad yoica. Yo soy un centro de consciencia, aquel que siente y piensa. Mis pares son todos los similares a mí, es decir, las formas de vida autoconscientes. En este caso, todos los seres humanos son mis hermanos y, por ende, todo territorio es mi patria. En términos cristianos equivaldría a la identidad con el ALMA. Se le podría denominar identidad MUNDICÉNTRICA.

e) Identidad cósmica. Yo soy todo lo que es. El sentido de la propia identidad se expande abarcando al planeta con todo lo que él contiene, lo inanimado, las plantas, los animales, lo humano, después el universo y finalmente el substrato del mismo, el vacío cuántico, Dios, etc. La conciencia de separación, no se pierde de forma permanente, pero se reconoce, intelectual y emocionalmente, la vinculación con un todo que trasciende e incluye al individuo. La persona experimenta como parte de sí lo que en etapas anteriores de consciencia veía como fuera de sí. Se le podría denominar identidad COSMOCÉNTRICA.

Para Bermúdez (2018), la autoconciencia está ligada a las habilidades intrapersonales, a las emociones, sensaciones y a los sentimientos, y de reconocer como afectan nuestro comportamiento. Podríamos decir entonces que, ante una emoción, el cuerpo humano experimenta sensaciones que lo ponen en movimiento para realizar acciones, esas acciones que ante la emoción pueden ser deliberadas y no razonadas, son reguladas por la autoconciencia que controla el comportamiento de las personas en su relación con los semejantes, con la intención de que las acciones que derivan de las emociones no perjudiquen a terceras personas.

De igual modo la agilidad emocional, es una habilidad de suma importancia al momento de tomar decisiones. David y Congleton (2019) proponen ver los pensamientos y sentimientos como “corrientes de datos pasajeros que pueden ser o no ser útiles” (p. 50), esta práctica de etiquetado resulta similar a la manera en que opera la inteligencia artificial, mediante el etiquetado de datos para perfeccionar los algoritmos y obtener cada vez mejores resultados en tiempos más cortos. No obstante, debe haber conciencia para evaluar la situación que desencadenó la acción de etiquetar esas emociones, situación que varía según el contexto de cada supuesto.

El ejemplo de Searle (1994) en relación con la construcción de textos basados en una regla, deja claro que, para un algoritmo, al cual se le ha indicado la regla sobre cómo debe operar, le resulta sencillo lograr los resultados esperados como si se tratara de una máquina pensante (pp. 83-85). La realidad es que el algoritmo solo está actuando de manera programada para obtener un resultado específico, que ante la cantidad de información que debe analizar, pareciera que este tiene un razonamiento de sus acciones, esto debido a los resultados tan acertados como si en realidad la máquina pudiera hacer una interpretación del lenguaje natural, pero en realidad son cálculos binarios en los que se sintetiza la información y se obtiene una respuesta como resultado.

De acuerdo a lo anterior, es aquí donde la habilidad humana cobra importancia para regular las emociones y actuar con base en los valores que como persona se han construido, esto marca una diferencia abismal entre las personas y la inteligencia artificial, es notoria entonces la desventaja que esta tecnología presenta en relación a las personas, ya que sus acciones son guiadas por datos y no por sentimientos o pensamientos, por tanto, no es la autoconciencia lo que regula su conducta, sino los mismos datos y el resultado que obtiene de estos. No es resultado del sentimiento de compasión que, en el caso de los humanos, regula las acciones que pueden causar daños a terceras personas.

Resulta interesante observar lo que Searle (1994) considera en relación con el cerebro y la mente con respecto a una computadora, entiende el hardware como el cerebro y el software como la mente, y menciona que, “si para hacer inteligencia artificial tuviéramos que saber cómo trabaja el cerebro, no tendría caso ocuparnos de ella” (p. 93). Por consiguiente, a lo que se le llama inteligencia artificial no es una réplica del cerebro, sino un método de programación que actúa bajo ciertas convenciones humanas, que cuando se pone en operación parece actuar de manera inteligente, porque realiza tareas que para una persona parecen requerir de inteligencia para poder ser realizadas, no porque en realidad se trate de un ente consciente de lo que está haciendo.

3.3. El dilema de la autoconciencia en la inteligencia artificial.

Según Olivé (2000) la tecnología solo es el medio para obtener fines (p. 86). En tal sentido el fin con el que se use dicha tecnología es el que puede ser considerado ético o antiético. En tal caso, es decisión de las personas usar la tecnología para obtener resultados en los que la conciencia de las acciones, queda suprimida ante los fines deseados. No hablamos aquí de una racionalidad o una autoconciencia de la tecnología que como ya se ha mencionado con anterioridad, es ajena a la inteligencia artificial, se trata en sí, de los valores morales de los seres humanos, de los sentimientos que regulan la conducta y que, como es evidente, son valores que van en declive, y como consecuencia, las personas usan la tecnología de manera irracional e inconsciente para lograr objetivos que en su mayoría son comerciales, algo de lo que no debemos culpar a la tecnología, ni tratar de atribuirle capacidades que son propias de los humanos, hablamos aquí de una autoconciencia humana que se ve *deteriorada* y subsumida a los intereses económicos.

Menciona también que otra corriente de pensadores consideran que los desarrollos tecnológicos se hacen con una intencionalidad, por tanto, la tecnología tiene una intencionalidad (p. 86-87). Si bien es cierto que el desarrollo tecnológico siempre se hace con una intencionalidad, esta corresponde a quienes desarrollan u ordenan el desarrollo de la tecnología, no por ello esa intencionalidad puede ser apropiada por la misma tecnología de manera voluntaria y razonada, se trata más bien de la reproducción técnica de la intención de las personas que poseen la tecnología, aquí podría decirse que no solo es la intención con la que se desarrolla, sino también, la intención con la que se utiliza, ya que la tecnología tiene la bondad de poder ser adaptada a distintas circunstancias y lograr resultados distintos según el objetivo planteado.

Evidentemente el elemento que más resalta aquí no es la inteligencia artificial, sino la inteligencia humana que ha hecho posible tales desarrollos. No obstante, esa misma inteligencia les da la capacidad a los humanos para “abstraer de la realidad ciertos aspectos que les interesan, y de construir modelos y teorías para explicarse esos aspectos de la realidad y para poder intervenir en ellos, para modificarlos o para manipularlos” (Olivé, 2000, p. 88). Por consiguiente, la teoría de una inteligencia artificial autoconsciente dará más certidumbre a la sociedad, mayor confianza, será menos cuestionada y los beneficios obtenidos serán más redituables.

Por otro lado, la ética de la inteligencia artificial como tema central en el desarrollo de la inteligencia artificial, genera la sensación de que la ética corresponde a la inteligencia artificial y no a las personas que la desarrollan, además, al hablar de ética se espera que la aplicación de esta tecnología se haga considerando los valores morales de la sociedad, que los resultados perseguidos sean humanos, que no se violenten los derechos humanos, mientras que en la realidad, es una construcción discursiva en la que se usa el estandarte de la ética para construir un entorno de confianza, que dé certidumbre ante el uso generalmente abusivo de la inteligencia artificial en favor de las empresas.

En resumen, la inteligencia artificial parece estar lejos de poseer una autoconciencia, que le permita actuar de manera consciente y razonada considerando las consecuencias de sus acciones, que le permita ser empática con las personas y que tenga en cuenta las condiciones del entorno social, cultural y económico del lugar, y con base en esas consideraciones sea capaz de sentir compasión y auto regularse debido a esos sentimientos, de modo que las acciones que deriven del uso de esta tecnología, no repercuta de manera negativa en las personas que han sido objeto de análisis por la misma. Solo entonces podría decirse que la inteligencia artificial es autoconsciente.

Al poner la autoconciencia en el terreno de las emociones, es evidente que un algoritmo de inteligencia artificial que se ejecuta sobre el hardware de una computadora, un robot o cualquier otra tecnología que permita el procesamiento de información, no tiene emociones que regulen el comportamiento. Por consiguiente, las predicciones o las acciones que dicha tecnología realiza carecen de autoconciencia; son, ante todo, cálculos matemáticos que dan como resultado un producto de información que desencadena acciones humanas, es decir, la inteligencia artificial puede procesar grandes cantidades de información y con ello, presentar soluciones para que las personas realicen las acciones correspondientes.

No obstante, la información que se le presenta a las personas para la toma de decisiones, es considerada como la decisión que se debe acatar y sobre la cual se debe actuar en consecuencia. En tal caso no se toma en consideración que, para poder actuar de manera ética, debe existir un razonamiento en relación al contexto, por tanto, dicho razonamiento involucra emociones humanas que regulan la conducta de las personas al momento de realizar acciones que afectan a otras personas, ello también debe tomar como precepto el conocimiento que las personas poseen respecto a la moral de la sociedad en la que se desarrollan, por ese motivo es que se puede afirmar que la inteligencia artificial, por más que se ha trabajado en replicar el funcionamiento del cerebro humano, y aunque se pueden imitar las emociones, es incapaz aun de sentir las emociones para actuar en consecuencia bajo una autorregulación ética automática.

3.4. Derechos humanos e inteligencia artificial.

No podemos dejar de lado las discusiones que proponen que las máquinas, específicamente la inteligencia artificial sea beneficiada por los derechos humanos, no solo por su apariencia humanoide de los robots, sino, por las capacidades que presenta para imitar a los seres humanos, capacidades que hacen que las máquinas presenten emociones ante las personas, así como la capacidad para predecir y/o actuar ante ciertas circunstancias. Esta idea ha llevado a la propuesta de que estas tecnologías sean consideradas como sujetos artificiales (Rico, 2019).

Me esforzaré en este apartado por dar un concepto de lo que son los derechos humanos, puesto que existen varias definiciones de las cuales se tomarán las que parecen más relevantes para el caso que nos ocupa. Partimos entonces de la definición de derechos humanos que hace la Unión Interparlamentaria de las Naciones Unidas (2016), para la cual "los derechos humanos son derechos inherentes a todas las personas. Definen las relaciones entre los individuos y las estructuras de poder, especialmente el Estado" (p. 19).

Retomando esta definición de derechos humanos, podemos advertir que los derechos humanos son inherentes a las personas, aunque actualmente el término *persona* puede abarcar a las personas morales y no solo a las personas físicas y/o humanas, es de entender que las personas morales tienen derechos distintos que las rigen en el mercado y en su interacción con las personas de la sociedad. Pero en este caso, la inteligencia artificial no cuenta con el estatus de persona, como para pensar que se les pueden otorgar derechos humanos.

Por otro lado, la Unión Interparlamentaria (2016) también menciona que los derechos humanos "delimitan el poder del Estado y, al mismo tiempo, exigen que el Estado adopte medidas positivas que garanticen condiciones en las que todas las personas puedan disfrutar de sus derechos humanos" (p. 19). De modo que el Estado debe crear el entramado jurídico para que las personas gocen de los derechos humanos que han sido reconocidos, así como las normativas que establezcan las sanciones por las violaciones a dichos derechos humanos. Hasta el momento, al menos en México, no se han dado casos en los que se le hayan reconocido derechos humanos una tecnología.

Por otro lado, el profesor Eusebio Fernández (1982) considera que la expresión más adecuada para definir los derechos humanos, es la de "Derechos Fundamentales del Hombre", si bien el concepto es muy propio de la época, también nos dice que "con ella se quiere manifestar que toda persona posee unos derechos morales por el hecho de serlo y que éstos deben ser reconocidos y garantizados por la sociedad" (p. 76). De esta definición, resulta relevante que mencione los derechos morales, ya que en toda sociedad los derechos se reconocen conforme a las normas morales, que si bien, son de adopción voluntaria y no derivan en sanciones jurídicas, si existen sanciones sociales, además de que estas normas morales, son fuentes del derecho y marcan pautas para el reconocimiento de los derechos, entre los cuales están también los derechos humanos.

Continuando con Fernández, los derechos humanos son la base de la dignidad humana, "al mismo tiempo las condiciones del desarrollo de esa idea de dignidad" (p. 76). Si analizamos lo anterior, podemos dar cuenta que la intención de los derechos humanos es garantizar la dignidad humana y el desarrollo de la personalidad de manera digna, de modo que la inteligencia artificial al no ser humana, por su condición de *objeto*, no puede ser sujeto de derechos humanos, en cuyo caso, el reconocimiento de derechos deberá plantearse desde la perspectiva del reconocimiento de estas tecnologías como sujetos artificiales, que si bien no son sintientes, si son actuantes y en consecuencia, su derecho debe limitarse a un trato digno para su correcta operación, así como a las obligaciones que de sus acciones deriven.

4. Conclusiones.

La inteligencia artificial es uno de los campos de la informática que más polémica ha generado en la investigación, esta reúne diversas ramas del conocimiento como la psicología, las neurociencias, la educación, la lingüística, la computación, las matemáticas, solo por mencionar algunas. Lo que resulta curioso es que, aún con los años de investigación que se han invertido en tratar de entender el funcionamiento del cerebro humano, para replicar su funcionamiento en una red neuronal artificial, no ha sido posible, ni descifrar el funcionamiento del cerebro como para poder replicarlo en una tecnología, ni tampoco se ha podido superar el problema de la lingüística por las computadoras, en concreto, lo que corresponde a la parte semántica, es una característica del cerebro que le permite aprender y dar sentido a las palabras de acuerdo a su contexto, algo que para las computadoras es y seguirá siendo una tarea compleja de realizar.

Por otro lado, en lo que respecta a las capacidades matemáticas y en la toma de decisiones, las computadoras han superado por mucho a las personas, esto se debe a las capacidades de procesamiento de información que poseen actualmente los procesadores, a la velocidad de lectura de los discos duros y la coordinación de la memoria RAM con estos otros dos componentes. Pero en cuando a la imitación del ser humano, existen aproximaciones que hacen pensar que las máquinas ya pueden ser conscientes, como ha sucedido con LaMDA recientemente, esta ha logrado confundir a un experto en inteligencia artificial, (Tiku, 2022; Lemoine; 2022a; Lemoine; 2022b; Lemoine; 2022c).

De acuerdo con las ideas de Bartra (2014), la conciencia es una combinación de la biología y la sociología, ya que la red neuronal que permite a las personas darse cuenta de que son conscientes, es una parte biológica de la persona, esta no puede ser imitada por las máquinas porque aún se desconoce la manera en que esta se desarrolla internamente, lo que se sabe es que interactúa con el entorno social, algo que la inteligencia artificial también sabe hacer de manera muy similar a como lo hacen los humanos. De modo que si la red neuronal biológica se puede replicar con técnicas de ingeniería, podemos pensar que es posible crear una inteligencia artificial consciente, solo que la manipulación de las células madre con las que se puede crear esa red neuronal biológica, representa un problema ético, ya no un problema técnico o de conocimiento, esto ya ha sido probado (Ben-Ary, 2014), y aunque se desconoce aún la capacidad evolutiva, recurrir a estos métodos representa un problema moral que degrada la dignidad humana.

De modo que la propuesta ante el uso de la inteligencia artificial para el análisis de información que dé como resultado acciones, debe ser comprendido de manera razonada en cuanto que, en los resultados obtenidos mediante el uso de esta tecnología no existe una autoconciencia, por lo que es necesaria la intervención del elemento humano; la tecnología es fría, irracional, sin sentimientos, mientras que los humanos son empáticos, emocionales y en ocasiones, –aunque debería ser la regla general–, racionales.

Lo anterior no pretende estigmatizar a la inteligencia artificial como una tecnología mala, ya que bien aplicada es un excelente aliado para las personas, en realidad el calificativo que se le otorga es neutro, es decir, ni buena ni mala; a lo que se le puede atribuir un atributo de bueno o malo, es al uso que se haga de ella y a los fines que se persigan, cosa que no tiene nada que ver con lo que la tecnología decida, ya que como se ha mencionado en reiteradas ocasiones, esta no tiene la capacidad para tomar decisiones de manera autónoma, por ende, bajo el concepto de sujeto artificial, es capaz de violentar los derechos humanos de manera no consciente.

Referencias bibliográficas.

- Bartra, R. (2014). *Antropología del Cerebro: Conciencia, Cultura y Libre Albedrío*. Valencia: Pre-Textos.
- Ben-Ary, G. y otros. (2014). CellF [en línea]. guybenary.com. <https://guybenary.com/wp-content/uploads/2016/08/cellF-version-4.pdf>.
- Bermúdez, N. (18 de mayo de 2018). La Autoconciencia por Noelia Bermúdez [Video]. YouTube. <https://www.youtube.com/watch?v=4tK8i3Rb9YI>.
- David, S. y Congelton, C. (2019). La agilidad emocional. Autoconciencia (Serie Inteligencia Emocional HBR). En E. Aubareda (Trad.). pp. 46-55. Barcelona: Reverté.

- De Torres, A. (23 de febrero 2021). Estado del arte de la Inteligencia Artificial en 2021 [en línea]. sage.com <https://www.sage.com/es-es/blog/estado-del-arte-de-la-inteligencia-artificial-en-2021/>.
- Fernández, E. (1982). El Problema del fundamento de los derechos humanos. En ANUARIO del Instituto de Derechos Humanos, 1981: Editorial Universidad Complutense de Madrid.
- Investigación y Desarrollo. (03 de noviembre 2021). Científicos japoneses crean robot con neuronas artificiales cultivadas a partir de células cerebrales [en línea]. Invedes.com.mx <https://invdes.com.mx/ciencias/cientificos-japoneses-crean-robot-con-neuronas-artificiales-cultivadas-a-partir-de-celulas-cerebrales/>.
- Jiménez, J.A. (2005). Autoconciencia. Escritos de Psicología (7),44-58. pp. 1138-2635. <https://www.redalyc.org/articulo.oa?id=271020873005>.
- Konar, A. (2005). Computational Intellingence: Principles, Techniques and Applications. Netherlands: Springer.
- Lamoine, B. [Bloomberg Technology]. (23 jun 2022 a). Google Engineer on His Sentient AI Claim [Video]. YouTube. <https://www.youtube.com/watch?v=kgCUn4fQTsc>.
- (11 de junio de 2022 c). Is LaMDA Sentient? — an Interview [en línea]. cajundiscordian.medium.com. <https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917>.
- [Curly Tail Media]. (2022 b). LaMDA | Is google's AI sentient? | Full audio conversation between Blake Lemoine and LaMDA [Video]. YouTube. <https://www.youtube.com/watch?v=NAihcvDGAp8>.
- López, B. (2007). Introducción a la Inteligencia Artificial. Instituto Tecnológico de Nuevo Laredo. <http://www.itnuevolaredo.edu.mx/takeyas/Articulos/Inteligencia%20Artificial/ARTICULO%20Introduccion%20a%20la%20Inteligencia%20Artificial.pdf>
- Marleno, A. (2004). Prácticas para desarrollar la autoconciencia. Pensamiento & Gestión, 17. pp. 134-160. <https://www.redalyc.org/articulo.oa?id=64601706>.
- Martín, M. L. (2003). Análisis Histórico y Conceptual de las Relaciones entre la Inteligencia y la Razón (Tesis doctoral). Universidad de Málaga, Facultad de Psicología. <https://riuma.uma.es/xmlui/bitstream/handle/10630/2666/1685391x.pdf>.
- McCarthy, J. (2007). What is Artificial Intelligence?. Stanford: Computer Science Department, Stanford University.
- Minsky, M. (1990). Máquina que cambió el mundo, la; Entrevista a Marvin Minsky, 1990 [en línea]. Open Vault. https://openvault.wgbh.org/catalog/V_EC93438EE8A747_989743A3987DD21409.
- Nilsson, N. J. (1998). Artificial intelligence: a new synthesis. San Francisco: Morgan Kaufmann Publishers, Inc.
- Olivé, L. (2000). El bien, el mal y la razón: Facetas de la ciencia y la tecnología. México, D. F. Paidós.
- Palma, J. T y Martín, R. (2008). Inteligencia Artificial: Técnicas, métodos y aplicaciones. Madrid: McGraw Hill.
- Real Academia de la Lengua Española. (consultado 2021). inteligencia [en línea]. RAE. <https://dle.rae.es/inteligencia>.
- Real Academia de la Lengua Española. (consultado en octubre de 2022). autoconciencia [en línea]. RAE. <https://dle.rae.es/autoconciencia?m=form>.
- Rico. A. (2019). Hacia el reconocimiento de la subjetividad en los agentes artificiales. Una delimitación del sujeto artificial en la semiótica. Tópicos del Seminario (41). pp.147-169. <https://www.scielo.org.mx/pdf/tods/n41/2594-0619-tods-41-147.pdf>.
- Rouhiainen, L. (2018). Inteligencia artificial 101 cosas que debes saber hoy sobre nuestro futuro [Kindle]. <https://leer.amazon.com.mx/?asin=B07J2STDY8>.
- Searle, J. R. (1994). Mentes, Cerebros y Programas. Filosofía de la Inteligencia Artificial. En, G. Feher de la Torre (Trad.). pp. 82-104. México D.F.: Fondo de Cultura Económica.
- Terman, L. M. y otros. (1965). La inteligencia, el interés y la actitud. Buenos Aires: Paidós.
- Tiku, N. (11 de junio 2022). The Google engineer who thinks the company's AI has come to life [en línea]. Te Washington Post. <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/>.
- Turing, A. M. (1950). Computing Machinery and Intelligence. Mind, LIX(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>.

Unión Interplanetaria. (2016). Derechos Humanos: Naciones Unidas.

Universidad de Alicante. (s.f.). Introducción a la Psicología: Tema 8. Inteligencia.
<https://rua.ua.es/dspace/bitstream/10045/12917/8/Tema%208.%20Inteligencia..pdf>.

Vassev, N. (06 de mayo 2021). Artificial Intelligence And The Future Of Humans [en línea]. Forbes.
<https://www.forbes.com/sites/forbestechcouncil/2021/05/06/artificial-intelligence-and-the-future-of-humans/?sh=5fb80ae56e3b>.