

Inteligencia artificial: retos y desafíos de la ética laboral en la sociedad tecnológica

Artificial Intelligence: Challenges of Work Ethics in the Technological Society

Recibido: 24 de junio de 2022

Aprobado: 25 de noviembre de 2022

Forma de citar: D.A. Duarte Arias, O.O. Chacón, "Inteligencia artificial: retos y desafíos de la ética laboral en la sociedad tecnológica", *Mundo Fesc*, vol 12 no. S3 pp. 265-280 de 2022

Daniel Augusto Duarte Arias* 

Magíster en Filosofía Contemporánea
danielduarte@ustavillavicencio.edu.co
Universidad Santo Tomás seccional de Villavicencio
Villavicencio, Colombia.

Orlando Ortega Chacón 

Magíster en Filosofía Contemporánea
oortega@usbog.edu.co
Universidad de San Buenaventura sede Bogotá
Bogotá, Colombia.

***Autor para correspondencia:**

danielduarte@ustavillavicencio.edu.co



Inteligencia artificial: retos y desafíos de la ética laboral en la sociedad tecnológica

Resumen

La inteligencia artificial (IA) ha sido objeto de múltiples definiciones y enfoques desde su concepción en la conferencia de Dartmouth College en 1956. Los modelos de IA se dividen en enfoques basados en la similitud con el desempeño humano y otro centrado en la toma de decisiones racionales. El test de Turing es fundamental en la evaluación de la IA, planteando escenarios donde humanos interactúan con máquinas sin distinguir entre ellos. Estos modelos abordan la capacidad de las máquinas para realizar tareas que van desde la comunicación hasta la resolución de problemas complejos, como los que se presentan en los entornos empresariales. No obstante, el desarrollo de la IA plantea desafíos éticos y sociales, especialmente en el ámbito empresarial, en una sociedad que avizora el impacto tecnológico. El problema radica en que la intervención tecnológica ha acelerado la transformación de las labores humanas, y con ello la desaparición o emergencia de nuevos trabajos. La tesis que aquí se quiere sostener es que se requiere de un análisis crítico ante estos desafíos y una propuesta de estrategia ética para la integración de la IA en el entorno empresarial. El presente es un artículo de reflexión que examina el concepto de la Inteligencia Artificial y su impacto en la condición humana; posteriormente, se expondrán los peligros subyacentes en la técnica moderna y finalmente, se indicará la serenidad [Gelassenheit] como el modo de relación libre frente a la tecnología como manifestación humana.

Palabras clave: Ética, Inteligencia Artificial, regulación tecnológica, transformación laboral.

Artificial Intelligence: Challenges of Work Ethics in the Technological Society

Abstract

Artificial intelligence (AI) has been the subject of multiple definitions and approaches since its conception at the Dartmouth College conference in 1956. AI models are divided into approaches based on similarity to human performance and one focused on rational decision making. The Turing test is central to the evaluation of AI, posing scenarios where humans interact with machines without distinguishing between them. These models address the ability of machines to perform tasks ranging from communication to complex problem solving, such as those encountered in business environments. However, the development of AI raises ethical and social challenges, especially in the business environment in a society that is anticipating technological impact. The problem lies in the fact that technological intervention has accelerated the transformation of human labor, and with it the disappearance or emergence of new jobs. The thesis of this paper is that a critical analysis of these challenges and a proposal for an ethical strategy for the integration of AI in the business environment is required. This is a reflective article that examines the concept of Artificial Intelligence and its impact on the human condition; subsequently, the underlying dangers of modern technology will be exposed and finally, serenity [Gelassenheit] will be indicated as the mode of free relationship in the face of technology as a human manifestation.

Keywords: Ethics, Artificial Intelligence, technological regulation, labor transformation.

Introducción

El progreso de la Inteligencia Artificial (IA), desde su concepción en la década de 1950 hasta los desarrollos más recientes, ha estado marcado por los cambios acelerados en sus alcances. Esto ha hecho difícil que pueda ser definida. La apuesta de este artículo de reflexión está en considerar la IA como la disciplina que tiene por objetivo crear programas, máquinas o agentes inteligentes basados en la imitación humana o el ideal racional. Así las cosas, dado el empeño de la IA, sus avances representan desafíos éticos y existenciales para el ser humano, sobre todo en sus diversas aplicaciones laborales.

La filosofía de la técnica, como la expuesta por Martin Heidegger en el siglo XX, es aún vigente porque ofrece una perspectiva sobre la esencia de la técnica y su impacto en la comprensión del ser humano en relación con el mundo. De este modo, la relación entre la esencia de la técnica y los avances de la IA revelan peligros existenciales inherentes a la dominación tecnológica a la que se está expuesto en la actualidad. Este problema es ampliamente elaborado por Francis Fukuyama cf. [2], Jürgen Habermas cf. [3] y Tamar Sharon cf. [4] porque pone en peligro incluso la condición humana. La serenidad [Gelassenheit] como se verá, será el modo adecuado de relación con la IA como producto de la técnica y ofrecerá el camino para la consolidación de una relación con la técnica que permite discernir entre el uso adecuado de la tecnología y su carácter absoluto, en otras palabras, una relación de libertad.

La serenidad, como la entiende Heidegger, invita a mantener una tensión constante entre el «sí y el no» frente a la tecnología, reconociendo su utilidad, pero también limitando su dominio absoluto sobre todas las dimensiones de la vida humana. En última instancia, la relación entre la IA y la serenidad heideggeriana resalta la importancia de las dimensiones técnicas (en su sentido griego) de la vida humana que la noción de técnica moderna oculta.

¿Qué es inteligencia artificial?

El término “inteligencia artificial” tiene origen en la conferencia Dartmouth College llevada a cabo en Hannover en el año 1956. Desde entonces, las definiciones de IA han sido abordadas a partir de dos enfoques principales: el primero se basa en la similitud a la actuación o desempeño humano, mientras que el segundo se fundamenta en el ideal racional de hacer o decidir lo correcto. Además, existen perspectivas que consideran la inteligencia como una propiedad interna del pensamiento y otros que consideran que es una característica externa centrada en el comportamiento. Estas orientaciones han sido objeto de estudio y en ocasiones pueden entrar en conflicto o no resultar compatibles entre sí cf. [5, pp. 19-20]. Un estudio actual del estado del arte sobre la inteligencia computacional y artificial puede verse en cf. [6].

John Preston [7] presenta una visión general de los modelos en el campo de la IA. Según él, ambos enfoques tienen su origen en las ideas de Alan Turing. El primer modelo se denomina "clásico" o "simbólico" y se enfoca en la creación de inteligencias centradas en la demostración de teoremas matemáticos basados en lógica, geometría y álgebra. La aplicabilidad de estos teoremas resulta interesante cuando se ponen a prueba en juegos como el ajedrez. El segundo modelo denominado conexionista, tal como lo expone Preston, argumenta que la IA, además de llevar a cabo actividades inteligentes de naturaleza simbólica, también debe ser capaz de resolver problemas que involucren un amplio espectro de elementos, como la visión, la comprensión del lenguaje natural, la planificación y el "aprendizaje automático", entre otras funciones cognitivas cf. [7, pp. 11-13]. De este modo, tanto el modelo que se centra en actividades abstractas de la IA como el modelo que busca la resolución de problemas por parte de una máquina se consideran los más influyentes en la búsqueda de inteligencias artificiales capaces de abordar diversas tareas y objetivos. Un ejemplo es la labor del ascensorista, que se vio reemplazada rápidamente por el surgimiento de herramientas automáticas como los elevadores capaces de desplazarse por las estructuras e instalaciones de un edificio.

Antes de caracterizar los diferentes enfoques de la IA, es relevante explicar en qué consiste el test de Turing. Esta prueba tiene su origen en el desarrollo de las ciencias computacionales y las matemáticas de Alan Turing de donde recibe el nombre el test. Su objetivo es crear un escenario en el que un ser humano se relaciona con un interlocutor y, a medida que avanza en esta actividad, es incapaz de distinguir si está interactuando con un humano o una máquina. La razón de esta incapacidad de distinción radica en que el rasgo de inteligencia de la máquina será equivalente a la humana. A partir de este punto, los enfoques en IA se basan en la semejanza con el ser humano o en una razón ideal.

Una clasificación de la inteligencia artificial, que refleja el desarrollo de diversos enfoques, se fundamenta en modelos que buscan imitar al ser humano y en modelos que responden a ideales de racionalidad. El enfoque humano se divide en aquellos basados en la actuación humana y en los basados en el pensamiento humano. Por otro lado, el enfoque racional se divide en modelos basados en el pensamiento racional y en la actuación racional. En cualquiera de estos casos, el avance de las inteligencias artificiales se vincula con la necesidad de crear máquinas capaces de imitar o incluso superar al ser humano.

El primer modelo humano es el enfoque basado en la actuación humana. Aquí se propone que los ordenadores o máquinas inteligentes superen rigurosamente la prueba del test de Turing [1] la cual consiste en alcanzar capacidades como la comunicación exitosa en el lenguaje humano, el almacenamiento de información, la resolución de preguntas, la extracción de conclusiones y el aprendizaje automático para adaptarse a nuevas circunstancias. Algunos investigadores añaden otras capacidades para que un ordenador demuestre inteligencia en una prueba de Turing total. Esta hace referencia a capacidades de los sentidos como la visión, el reconocimiento de voz, la manipulación

de objetos y el desplazamiento. A pesar de que los investigadores recurren al test de Turing para probar la inteligencia de las máquinas, consideran que el objeto de estudio no es imitar la inteligencia, sino centrarse en examinar los principios subyacentes de la inteligencia y con ello lograr emular rasgos de actuación e inteligencia humana en una máquina [5, p. 20].

El segundo enfoque del primer modelo está basado en el pensamiento humano. Este sostiene que se puede aprender del pensamiento humano a partir de la introspección y captura de nuestro propio pensamiento, la observación de una persona en acción y la observación del cerebro en acción. A diferencia del enfoque basado en la actuación humana, este aprendizaje puede expresarse en un programa informático, y si el input y output del comportamiento del programa coinciden con el humano, entonces se considera prueba de que el comportamiento del programa puede funcionar de manera similar a la cognición humana. Algunos de los defensores de esta modalidad de inteligencia artificial son Herbert Simón y Allen Newell en *Human problem solving* en 1972. Allí los autores argumentan a favor de la posibilidad de modelar computacionalmente la inteligencia y la toma de decisiones basadas en el ser humano. Para esto, Newell y Simon proponen el "Teorema de la Completitud", que consiste en programar una computadora de tal forma que, apoyada en una buena capacidad de memoria y tiempo de cálculo, logre resolver problemas cf. [8]. Si bien, en este enfoque no necesariamente se debe recurrir a capacidades sensoriales, el enfoque basado en pensamiento humano sí requiere de la capacidad de una máquina para reconocer el contenido semántico y así arrojar datos sobre cómo funciona la cognición humana. Algunos, como Marvin Minsky [9] argumentan que, si un modelo logra replicar los procesos mentales y cómo se interconectan, entonces se podrán crear inteligencias artificiales más sofisticadas cf. [9].

Por otra parte, del segundo modelo se presenta el primer enfoque basado en el pensamiento racional. Este intenta codificar el pensamiento correcto a partir de razonamientos irrefutables. Un modelo de este tipo es el desarrollado por John McCarthy denominado LISP. McCarthy pretendía, a partir del procesamiento de listas, programar inteligencias artificiales capaces de interactuar con el lenguaje natural, reglas lógicas y manipulación de símbolos cf. [10]. Este tipo de estructuras inferenciales no se puede refutar debido a su validez lógica y evidencia indiscutible como en las matemáticas. A partir de este modelo, los programadores pretenden crear programas con sistemas inteligentes. El límite de la lógica es que hay conocimiento sobre el mundo que no se reduce a reglas aritméticas como, por ejemplo, la política o la guerra. Para suplir este vacío surgen las teorías de probabilidad para construir un modelo de pensamiento con información incierta que traslade la información perceptiva básica hasta la comprensión del funcionamiento del mundo y las predicciones sobre el futuro. Sin embargo, este enfoque tiene como límite que no genera comportamiento inteligente, pues carece de una teoría de la acción racional. Dicho de otra manera, para generar un comportamiento inteligente, no es suficiente procesar información para predecir sucesos futuros; además, se requiere saber qué hacer con esa información, es decir, actuar de manera racional ante la información procesada [5, p. 21].

La segunda caracterización del modelo racional también denominado “enfoque de actuación racional” es en el que se argumenta que un agente racional “es aquel que actúa para conseguir el mejor resultado o, cuando hay incertidumbre, el mejor resultado esperado” [5, p. 22]. De este modo, al operar de forma autónoma y adaptarse al cambio, logran alcanzar los objetivos para los que son programados. Esta perspectiva presenta ventajas sobre el modelo basado en el ser humano y en el pensamiento racional, ya que utiliza métodos de teorías de la probabilidad y actúa de forma racional tomando la mejor decisión posible y, a partir de estos, la manipulación y el procesamiento de datos. Así las cosas, los agentes inteligentes toman decisiones incluso en casos de incertidumbre. El objetivo proporcionado al agente es lo que determina qué es lo correcto o cuál acción es la adecuada.

Un agente racional actúa de manera adecuada cuando no hay tiempo suficiente para realizar cálculos y tomar decisiones complejas. Un ejemplo puede ser el de una máquina que debe encontrar una serie de objetos en un espacio determinado; sin embargo, dicho espacio está minado por obstáculos que no le permiten el paso sin ser esquivados o superados. La máquina deberá procesar la información que logre detectar mediante un rastreo del espacio, incluso de aquellos lugares a los que no tenga acceso por los límites visuales o locativos. En estos casos procesará la información y, ante algunas incertidumbres, hará procesos probabilísticos para actuar de forma racional. A este tipo de paradigma se le denomina “modelo estándar” y se aplica en la creación de la IA cf. [5, pp. 21-22].

Desde esta última modalidad de inteligencias artificiales surge un modelo basado en multiagentes. Bárbara Grosz y David Sarné [11] argumentan a favor de un mecanismo que permite a los agentes inteligentes mejorar sus estimaciones probabilísticas de interacción con entornos dentro de un sistema en el que las observaciones directas de las actuaciones de otros agentes juegan un papel fundamental en la recolección de datos y aprendizaje. Este modelo de interacción muestra que incluso los agentes inteligentes, al enfrentarse a datos dispersos, son capaces de aprender en entornos dinámicos. De este modo, con el desarrollo multiagente se superan las estimaciones probabilísticas basadas exclusivamente en las propias observaciones de un agente cf. [11].

Con los enfoques hasta ahora abordados se puede definir a la Inteligencia Artificial como el campo o disciplina que tiene como objetivo crear programas, máquinas o agentes inteligentes basados en la imitación humana o el ideal racional. Esta definición puede ser ampliamente aceptada en los diferentes campos que se dedican al estudio de la IA.

El peligro subyacente en la técnica moderna

En su conferencia La pregunta por la técnica [12], Martín Heidegger plantea la posibilidad de establecer una relación libre entre el ser humano y la técnica [12, pp. 9; 9]. Para ello,

el filósofo elabora una distinción entre la esencia de la técnica y sus manifestaciones concretas, destacando la influencia dominante que la técnica moderna ejerce sobre la forma en que el ser humano se relaciona con la naturaleza. En este sentido, Heidegger recurre a la acepción griega de *Tecné* estableciendo una relación con la *poiesis* que amplía la comprensión objetivante de la técnica en la actualidad.

La distinción entre la técnica y la esencia de la técnica implica una diferencia fundamental entre el fenómeno humano de la técnica y su percepción como expresión concreta. En otras palabras, los avances técnicos, como la producción de IA, se diferencian de lo que esta actividad expresa acerca de la esencia de la técnica como proceso productor. El filósofo alemán, en su conferencia, hace referencia al sentido latino de la técnica como *instrumentum* [12, pp. 10; 10], destacando así la limitación de definir la técnica exclusivamente en términos instrumentales, ya que tal enfoque no captura la esencia misma de la técnica. Esta distinción resulta crucial en la estructura interna de la exposición de Heidegger, porque busca establecer una "relación libre" con las manifestaciones materiales de la técnica, como las formas contemporáneas de producción de energía y la relación con la naturaleza como recurso. En este sentido, Heidegger alude a una significación mucho más profunda de la *Tecné* en su sentido griego, con el propósito de establecer las condiciones para una relación genuinamente libre. En palabras del filósofo: "La relación es libre si abre nuestro ser a la esencia de la técnica". Si respondemos a ella, entonces somos capaces de experimentar lo técnico en su limitación" [12, pp. 9; 9].

La limitación de la técnica a la que alude Heidegger se encuentra en una comprensión restringida que se circunscribe únicamente a un enfoque moderno. En otras palabras, se limita a una concepción de la verdad [13, p. 63] que únicamente considera lo objetivo o, dicho de otra manera, aquello que puede ser objetivado, como real. Heidegger empieza su conferencia *La época de la imagen del mundo* con la siguiente afirmación: "en la metafísica se lleva a cabo la meditación sobre la esencia del ente así como una decisión sobre la esencia de la verdad" [13, p. 63]. El filósofo muestra así que cada época está definida por la concepción de verdad que subyace en su comprensión del mundo. En ese sentido, lo que hace que la época Moderna sea diferente a la Medieval o la Antigua es su interpretación de lo ente. Para el caso de la modernidad, solo se considera real lo que es objetivable. Desde la perspectiva de la ciencia moderna, todo lo que no pueda ser abordado desde esta óptica se considera un error metodológico y, por ende, se excluye de cualquier consideración que pueda ser validada bajo los criterios aceptados de veracidad.

En este contexto, Heidegger sostiene que la técnica es aquello que pone delante algo del ser humano Cfr. [12, p. 15]. En otras palabras, la técnica revela aspectos del ser humano que permanecían ocultos debido a la falta de un medio adecuado para su expresión. Por consiguiente, podría argumentarse que la técnica constituye una forma de ser del ser humano (no exclusiva) que se manifiesta a través de inventos técnicos. Los avances en inteligencia artificial en la actualidad pueden interpretarse como expresiones técnicas

del ser humano, revelando nada más que la capacidad humana de trascender su entorno y aspirar a la perfección. En Reglas para el parque humano, Peter Sloterdijk encuentra que las expresiones técnicas que buscan la perfectibilidad humana son posibles porque el ser humano es para sí mismo domesticable o, en palabras del filósofo, “el hombre representa para el hombre el poder superior” [14, p. 215].

Para Heidegger las expresiones técnicas del ser humano se concretaron seguramente en las tensiones de la Guerra Fría, el primer alunizaje y los desafíos nucleares de oriente y occidente. Sin embargo, las preocupaciones del filósofo eran diferentes a las que la IA trae para la contemporaneidad. Ciertamente, no es la amenaza de extinción por radiación la que genera zozobra hoy sino la amenaza de la extinción de las condiciones de posibilidad de la vida humana en el capitalismo tales como el acceso a recursos naturales, vivienda y empleo. Es plausible pensar que la IA representa una amenaza para la existencia de trabajos profesionales como la administración de empresas, la contaduría, las especialidades médicas, o incluso la traducción de manuscritos, el periodismo, la docencia, la jurisprudencia. Más aun los trabajos técnicos y los no calificados como la operación aérea, conducción de vehículos y el servicio al cliente en general. Todos ellos están en riesgo de extinción tal como ocurrió en las anteriores revoluciones industriales. Lo particular de la IA en la actual revolución es su capacidad de suprimir el factor humano de todos los procesos técnicos. Sin embargo, esta amenaza no es externa pues ha surgido del mismo ser humano.

En el libro Crítica de la razón cínica el profesor Peter Sloterdijk hace una consideración sobre las bombas atómicas que puede parafrasearse para pensar las amenazas de la IA hoy. “La bomba no es ni un ápice más destructiva que la realidad ni una pizca más destructiva que nosotros. Es solo proyección, la representación material de nuestro ser. [...] La bomba no exige de nosotros ni lucha ni resignación, sino autoexperiencia. Nosotros somos ella” [15, p. 219]. Ciertamente, en esta afirmación de Sloterdijk subyace una comprensión de la técnica en sentido griego, como la que Heidegger expuso en la conferencia La pregunta por la técnica. Una preocupación similar la expresa Nick Bostrom cf. [16] en una breve fábula donde expone los riesgos que conlleva el desarrollo de inteligencias artificiales. La problemática radica en que si emerge la Superinteligencia Artificial (SIA) muy probablemente el ser humano pasará a ser una especie sometida por las máquinas. Así, tanto la fisión nuclear, a la que se refiere Sloterdijk, como las capacidades de la IA que estremecen el estado actual de cosas, la técnica vuelve a presentar hoy la antigua perplejidad de lo que significa el hombre para sí mismo porque, siguiendo a Heidegger, “en todas partes estamos encadenados a la técnica sin que nos podamos librar de ella, tanto si la afirmamos apasionadamente como si la negamos” [12, pp. 9; 9].

Una vez más surge el interrogante antiguo de la filosofía sofística del hombre como enigma para sí mismo. En verdad, la técnica no es un mero medio [instrumentum] sino “un modo de salir de lo oculto” [12, pp. 16; 15]. En ese sentido, lo que la técnica trae

consigo es un momento de autoexperiencia del ser humano. Este panorama es el que permite hacer comprensible la relación entre Tecné y Poiesis que establece Heidegger en la conferencia citada. En efecto, la esencia de la técnica es «trae-ahí-delante» [12, pp. 16; 15] para el ser humano lo que es el ser humano mismo; la tecné “saca de lo oculto algo que no se produce a sí mismo y todavía no se haya ahí delante y por ello puede aparecer y acaecer de este modo o de este otro” [12, pp. 16; 17]. Así las cosas, la particularidad de la técnica en la modernidad, como se verá a continuación, es emplazar la realidad.

El hacer salir de lo oculto que domina por completo a la técnica moderna tiene el carácter de emplazar, en el sentido de la provocación. Éste acontece así: la energía oculta en la naturaleza es sacada a la luz, a lo sacado a la luz se lo transforma, lo transformado es almacenado, a lo almacenado a su vez se lo distribuye, y lo distribuido es nuevamente conmutado [12, pp. 18; 20].

La cita anterior pareciera hacer referencia al proceso que se da en la producción de energía de una hidroeléctrica. Ciertamente, la técnica provoca a la naturaleza para que produzca lo que le interesa al modo de proceder objetivador. Es decir, el modo de acercarse de la hidroeléctrica al agua solo toma en consideración lo que es susceptible de convertirse en movimiento y este a su vez en energía que pueda ser almacenada y distribuía. Así, el proceso técnico emplaza una dimensión mecánica de la realidad y desplaza todo lo que no obedezca a ese modo de proceder. En ese sentido, siguiendo a Heidegger la técnica moderna tiene como objetivo “sacar a la luz, transformar, almacenar, distribuir, conmutar” [12, pp. 18-19; 20]. Para el caso de la hidroeléctrica lo que se saca a la luz es la dimensión mecánica del agua. Así las cosas, de la cita anterior lo que nombra el carácter propio de la técnica moderna es su capacidad de emplazar la realidad.

La palabra «emplazar», en el rótulo estructura de emplazamiento, no mienta solamente el provoca, al mismo tiempo tiene que conservar la resonancia de otro «emplazar» del que deriva, a saber, de aquel pro-ducir y representar que, en el sentido de la Poiesis hace que venga a darse lo presente” [12, pp. 23; 24].

El párrafo anterior ofrece una claridad sobre la relación entre tecné y poiesis en el modo de proceder emplazatorio de la técnica moderna. En efecto, la tecné pertenece al ámbito de lo poiético en la medida en que cumple la función de sacar de lo oculto [poiesis]. No obstante, ese sacar de lo oculto de la tecné, oculta otras dimensiones de la naturaleza. En ello emerge un peligro: La estructura del emplazamiento no solo pone en peligro al hombre en su relación consigo mismo y con todo lo que es. [...] la estructura del emplazamiento oculta sobre todo aquel salir lo oculto que, en el sentido de la poiesis, hace venir delante, deja aparecer a lo presente [12, pp. 29; 31].

Así las cosas, el peligro está en la relación que el ser humano establece consigo mismo cuando la técnica muestra una dimensión particular de la existencia humana y la reduce a esta. Más aún, el proceder de la técnica se extiende a todas las dimensiones del ser humano y establece su dominio sobre él. Puede pensarse por ejemplo en la relación que

se establece con las plataformas como Waze o Google Maps cuando definen las rutas menos congestionadas entre el punto A y el punto B en una ciudad. La credibilidad que tienen lleva a pensar en los cambios que la IA trae para la esencia de la verdad. Es decir, la relación que se establece, en el caso particular de las aplicaciones mencionadas, es una relación con la verdad.

La serenidad como modo de relación con la técnica.

Dos años después de que Heidegger pronunciara la conferencia *La pregunta por la técnica* [12] expone en Messkirch, el 30 de octubre de 1955 la conferencia *Serenidad* [Gelassenheit] [17]. En la conferencia *Serenidad* [17] Heidegger retoma el problema que había trabajado en *La pregunta por la técnica* [12] y se refiere, se advierte entre líneas, al proceder propio de la técnica moderna: el emplazamiento de la realidad.

La creciente falta de pensamiento reside así en un proceso que consume la médula misma del hombre contemporáneo: su huida ante el pensar. [...] Su peculiaridad consiste en que cuando planificamos, investigamos, organizamos una empresa, contamos ya siempre con circunstancias dadas. Las tomamos en cuenta con la calculada intención de unas finalidades determinadas. [...] El pensamiento calculador corre de una suerte a la siguiente, sin detenerse nunca ni pararse a meditar. El pensar calculador no es un pensar meditativo; no es un pensar que piense en pos del sentido que impera en todo cuanto es [17, p. 18].

Heidegger hace explícita la necesidad de las empresas de actuar con un proceder calculador. Lo cual resulta ampliamente aceptado e incluso incuestionable hoy. OpenAI, la empresa creadora de ChatGPT [18] expone en su visión que sus avances en Inteligencia Artificial General buscan beneficiar a toda la humanidad. Algunos exponentes como Antonio Diéguez cf. [19] suelen ser más conservadores y críticos de estos avances tecnológicos pues mira con reservas el uso indiscriminado de la IA. Ciertamente, cualquier proyecto que no especifique sus medios y fines no puede ejecutarse bajo los parámetros de la investigación actual. No se trata aquí de defender la irracionalidad y la ausencia de planeación, sino de hacer evidente un dogma técnico que impide el pensamiento. La expresión del filósofo es “la creciente falta de pensamiento reside así en un proceso que consume la médula misma del hombre contemporáneo: su huida ante el pensar” [17, p. 18]. En este sentido, el pensamiento calculador, al modo de las aplicaciones de IA no da lugar a objeciones porque hay una decisión incuestionable sobre la esencia de la verdad. En *Serenidad* Heidegger hace explícito el peligro que trae la técnica como emplazamiento de la realidad: Lo verdaderamente inquietante, con todo, no es que el mundo se tecnifique enteramente. Mucho más inquietante es que el ser humano no esté preparado para esta transformación universal; que aún no logremos enfrentar meditativamente lo que propiamente se avecina en esta época [17, p. 25].

Una vez enunciado el peligro: que la técnica se extienda a todos los ámbitos de la vida del hombre, formula su célebre «sí y no» el que se comprende el sentido del título de

su conferencia: Gelassenheit. El término ha sido traducido al español como serenidad, al francés como Sérénité y al inglés como Releasemet como indica el traductor de la versión castellana Yves Zimermann [17, p. 7]. Sin embargo, en el castellano existen algunos términos relacionados con Gelassenheit y gelassen tales como: dejar entrar, radicado en, comprometerse, desistir, confiar a, soltar, conceder, abandonar, tolerar [17, p. 10]. La traducción filosófica de Gelassenheit implica una relación directa con la pregunta por la técnica y a la relación meditativa que podría establecerse hoy con relación a la técnica. Dado que no hay un equivalente en el español (y en el alemán solo tiene sentido en su lugar de enunciación) la traducción del término podría ser, Gelassenheit: sí a la técnica, pero no a que se extienda a todos los ámbitos de la vida del hombre; «sí y no». En palabras de Heidegger:

Podemos decir «sí» al inevitable uso de los objetos técnicos y podemos a la vez decirles «no» en la medida en que rehusamos que nos requieran de modo tan exclusivo, que dobleguen, confundan y finalmente, devasten nuestra esencia. Pero si decimos simultáneamente «sí» y «no» a los objetos técnicos, ¿no se convertirá nuestra relación con el mundo técnico en equívoca e insegura? Todo lo contrario. Nuestra relación con el mundo técnico se hace maravillosamente simple y apacible. Dejamos entrar a los objetos técnicos en nuestro mundo cotidiano y, al mismo tiempo, los mantenemos fuera, o sea, los dejamos descansar en sí mismos como cosas que no son algo absoluto, sino que dependen ellas mismas de algo superior. Quisiera denominar esta actitud que dice simultáneamente «sí» y «no» al mundo técnico con una antigua palabra; la Serenidad (Gelassenheit) para con las cosas [17, p. 27].

Ante la pregunta por las implicaciones para la vida humana de los avances técnicos de la IA, la conferencia de 1955 pareciera tener una vigencia extraordinaria. En efecto, las posibilidades que parecían utópicas hace solo tres años, hoy son tecnologías puestas en el mercado. No obstante, ¿cómo establecer una relación con los avances de la IA desde una actitud serena que se establezca en el «sí y no»?

Los avances de la IA se materializan en el desplazamiento del factor humano de tareas en las que la precisión y el cálculo son absolutamente necesarias para el modo de proceder de la técnica hoy. No hay otra opción que decir sí a la técnica expresada como IA. Empero, habría que decir no al carácter absoluto de la forma de emplazamiento del mundo de la técnica moderna que desplaza al lugar del error metodológico lo que no es susceptible de ser calculado. Este principio se ve enfrentado a una perspectiva de la tecnología sostenida por la tesis de la mente extendida. Andy Clark consideraría que incluso, en este sentido, no se podría decir sí o no, sino que siempre es sí porque desde el principio el ser humano ha sido naturalmente un Cyborg y su relación con el mundo es tecnológica cf. [20]. En ese sentido, las religiones con sus búsquedas espirituales; el lenguaje poético con su capacidad de insistir en la dimensión no objetualizadora de las palabras y la labor misma de la filosofía de mantenerse en epojé frente a lo incuestionable, axiomático y categórico de las ciencias será lo que hará posible el «sí y no» en la relación

serena [Gelassenheit] con la expresión técnica del ser humano en la IA.

El término epojé es popularizado en el siglo pasado por el filósofo alemán Edmund Husserl. El filósofo construyó un método para fundamentar la filosofía como ciencia rigurosa. En ese contexto plantea tres tipos de reducción para poder acceder a los fenómenos de la conciencia. En la primera reducción (reducción fenomenológica) se plantea la necesidad de hacer epojé de los prejuicios. En palabras de Husserl: "se requiere, por tanto, una consecuente epojé por parte del fenomenólogo si quiere alcanzar su conciencia como fenómeno puro, singularmente, pero también como el todo de su vida pura" [21, p. 42]. Adolf Reinach fue uno de los discípulos más sobresalientes de la etapa temprana de la formulación de la fenomenología de Husserl. El discípulo escribe: "lo que vemos [del mundo] se rige en general por nuestras necesidades y nuestros fines. Sabemos cuán penoso es aprender a ver realmente" [22, p. 18]. Ciertamente, las interpretaciones axiomáticas que se tienen a la hora de acercarse a un fenómeno del mundo pueden ser obstáculo para su análisis certero. Por esa razón se hace necesario hacer epojé o, en otras palabras, poner entre paréntesis las creencias y prejuicios.

Conclusiones

La Inteligencia Artificial (IA) se ha desarrollado desde distintas perspectivas y modelos. Los más prometedores son aquellos que intentan imitar al ser humano o aquellos que generan producción a partir del ideal racional. Estos enfoques basados en el lenguaje simbólico y el conexionismo buscan crear máquinas capaces de resolver tareas y abordar objetivos mejor que un ser humano. Esto propone retos éticos por el riesgo al remplazo de los roles humanos por máquinas inteligentes que pueden cumplir sus labores.

Existe una relación entre la esencia de la técnica moderna y los avances en la IA. Para Heidegger la técnica moderna, al emplazar la realidad y reducir la esencia del ser humano a sus expresiones técnicas, plantea peligros existenciales. El peligro radica en que el ser humano, desprovisto de un andamiaje preparativo, no logra enfrentar los desafíos y las transformaciones que trae consigo la tecnología.

El concepto de Gelassenheit (serenidad) propuesto por Heidegger avizora un camino de relación con la IA. Al igual que ocurre con otras técnicas, el ser humano puede asumir una postura crítica frente a los diferentes usos de estas inteligencias artificiales y su dominio sobre la vida humana. Una tensión constante entre el sí y el no. La actitud serena invita a decir sí a la técnica, pero no a su carácter absoluto. Este equilibrio tiene como fin la preservación de la vida humana.

La relación entre la Inteligencia Artificial y la propuesta Heideggeriana de la serenidad expone la necesidad de dimensiones como la espiritualidad, la poesía y la filosofía para considerar el avance tecnológico. De no ser esto posible, el cálculo y la objetivación propia de la técnica moderna conducirán al declive humano. Valorar las demás dimensiones

humanas puede ser una herramienta útil para enfrentar los desafíos empresariales del surgimiento de la IA.

Agradecimientos

A la Universidad Santo Tomás seccional Villavicencio y el grupo de investigación Humanidades digitales, formación y construcción de paz territorial. ÁBA COL0142699. También a la Universidad de San Buenaventura sede Bogotá y el grupo de investigación Devenir COL0047323.

Referencias

- [1] A. M. Turing, "Computing Machinery and Intelligence", *Mind*, vol. LIX, n.o 236, pp. 433-460, oct. 1950. Doi: 10.1093/mind/LIX.236.433
- [2] F. Fukuyama, *Our posthuman future: consequences of the biotechnological revolution*, 1. Picador ed. New York: Picador, 2003
- [3] J. Habermas, *El futuro de la naturaleza humana: ¿Hacia una eugenesia liberal?* en Biblioteca del Presente, no. 20. Barcelona: Paidós, 2002
- [4] T. Sharon, *Human Nature in an Age of Biotechnology: The Case for Mediated Posthumanism*. New York, NY: Springer Berlin Heidelberg, 2013
- [5] S. J. Russell y P. Norvig, *Artificial intelligence: a modern approach*, Fourth edition. en Pearson series in artificial intelligence. Hoboken: Pearson, 2021
- [6] J. Ferney-Virgüez, H. F. Castro-Silva, y T. Velásquez-Pérez, "Tendencias de la inteligencia computacional", *Mundo FESC*, vol. 11, n.o 22, pp. 48-58, ene. 2021
- [7] J. Preston, «Introducción», en *Views into the Chinese room: new essays on Searle and artificial intelligence*, M. Bishop y J. Preston, Eds., Oxford: Clarendon Press, pp. 1-50. 2002
- [8] A. Newell y H. A. Simon, *Human problem solving*, 8. print. Englewood Cliffs, NJ: Prentice-Hall, 1972.
- [9] M. Minsky, *The society of mind*, 6. Pb-Pr. en A Touchstone book. New York: Simon and Schuster, 1988
- [10] J. McCarthy, "History of LISP", *ACM SIGPLAN Not.*, vol. 13, n.o 8, pp. 217-223, ago. 1978. Doi: 10.1145/960118.808387
- [11] D. Sarne y B. J. Grosz, "Sharing experiences to learn user characteristics in dynamic

environments with sparse data", en *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, Honolulu Hawaii: ACM, pp. 1-8, may 2007. Doi: 10.1145/1329125.1329176

- [12] M. Heidegger, "La pregunta por la técnica", en *Conferencias y artículos*, Barcelona: Odós, pp. 9-37, 1994
- [13] M. Heidegger, "Época de la imagen del mundo", en *Caminos de bosque*, Madrid: Alianza, pp. 63-90, 2014
- [14] P. Sloterdijk, "Reglas para el parque humano: Una respuesta a la carta sobre el humanismo de Heidegger", en *Sin Salvación: tras las huellas de Heidegger*, Madrid: Akal, pp. 197-220, 2011
- [15] P. Sloterdijk, *Crítica de la razón cínica*. Madrid: Siruela, 2019
- [16] N. Bostrom, *Superintelligence: paths, dangers, strategies*, First edition. Oxford: Oxford University Press, 2014
- [17] M. Heidegger, *Serenidad*. España: Odós, 1994
- [18] "ChatGPT". [En línea]. Disponible en: <https://openai.com/chatgpt>
- [19] A. Diéguez, *Transhumanismo*. S.l: Herder Editorial, 2021
- [20] A. Clark, *Natural-born cyborgs: minds, technologies, and the future of human intelligence*. Oxford; New York: Oxford University Press, 2003
- [21] E. Husserl, *Invitación a la fenomenología*, 1. Reimp. en *Pensamiento contemporáneo*, no. 27. Barcelona: Paidós, 1998
- [22] A. Reinach, *Sobre fenomenología*. Madrid: Encuentro, 2014