

Autonomía de la inteligencia artificial: Una reflexión crítica desde la filosofía kantiana

Autonomy of artificial intelligence: A critical reflection from Kantian philosophy

Flavio Suárez-Muñoz 

Instituto Nacional para el Desarrollo de la Ciberseguridad, México

Alma Azucena Bautista Tejada 

Universidad Michoacana de San Nicolás de Hidalgo, México

Correspondencia: milenioflavio@hotmail.com, alma.bautista@umich.mx

Resumen. La inteligencia artificial ha avanzado significativamente en el procesamiento de grandes volúmenes de datos, permitiendo la toma de decisiones, la predicción de conductas futuras basadas en el comportamiento humano y la generación focalizada de información, creando una apariencia de autonomía. Sin embargo, al analizar la autonomía desde una perspectiva kantiana, se identifican los requisitos esenciales para su existencia. Surgen las preguntas sobre si es verdaderamente autónoma la inteligencia artificial?, o si, ¿la autonomía simplemente sirve como estrategia discursiva para legitimar acciones automatizadas sin consideraciones éticas?. Es crucial destacar que la ética y moralidad en las acciones de la inteligencia artificial no son inherentes, sino que dependen en gran medida de quienes la desarrollan y supervisan. La moralidad y ética relacionadas con las acciones automatizadas solo pueden evaluarse en relación con sus creadores, ya que en ello se basa el juicio de valor detrás de la tecnología. Por lo tanto, es evidente la necesidad de establecer sólidas regulaciones éticas y morales, para evaluar y sancionar de manera justa las intenciones de aquellos que utilizan esta tecnología y explotan datos personales sin considerar los posibles efectos negativos en la vida humana.

Palabras clave: Inteligencia artificial, autonomía, discurso, responsabilidad, ética de la inteligencia artificial, IA.

Abstract. Artificial intelligence has advanced significantly in the large volume of data processing that allows a better decision-making, the prediction of a future conduct based on human behavior and the focused generation of information, creating an appearance of autonomy. However, analyzing autonomy from a Kantian perspective, the essential requirements for its existence are identified. The questions emerge like Is artificial intelligence truly autonomous? or If does the autonomy simply serve as a discursive strategy to legitimize automated actions without ethical considerations? Crucially, the ethics and morality in the actions of artificial intelligence are not inherent, but depend heavily on those who develop and supervise it. The morality and ethics related to automated actions can only be evaluated in relation to their creators, this is based on the value judgment behind the technology. Therefore, the need to establish strong ethical and moral regulations is evident, to evaluate and sanction the intentions of those who use this technology and exploitation of personal data without considering the possible negative effects on human life.

Keywords: Artificial intelligence, autonomy, discourse, responsibility, ethics of artificial intelligence. AI.

Recibido: 19/10/2023 Aceptado: 17/05/2024

1. INTRODUCCIÓN

La inteligencia artificial es una tecnología que ha logrado grandes avances en los últimos años, al grado de realizar actividades que antes solo las personas podían hacer. Entre sus capacidades, se encuentra el procesamiento de lenguaje natural, que permite a las máquinas comprender, interpretar y generar texto de manera automatizada. Sin embargo, aun cuando parece actuar por sí misma, la inteligencia artificial carece de autonomía en el sentido humano, es decir, no posee libre albedrío ni voluntad propios.



Respecto a la autonomía, esta se define como la capacidad de los individuos para gobernar su propia vida y tomar decisiones sin influencias externas. Involucra la libertad de elegir el propio plan de vida y asumir la responsabilidad de las consecuencias. En cambio, la inteligencia artificial actúa siguiendo algoritmos programados por humanos, por lo que, su supuesta autonomía es solo de carácter técnico-operativo.

Por tanto, aunque algunas máquinas puedan buscar sus propios insumos de energía o generar texto de forma automatizada, sus acciones responden a instrucciones dadas, y sus decisiones se basan en los datos que se les proporcionan. En otras palabras, la inteligencia artificial como creación humana carece de una intencionalidad o conciencia propia. Por ello, no se les puede atribuir una ética o moralidad autónoma, sino que la moralidad de sus actos depende de la ética de sus creadores humanos.

No obstante, la inteligencia artificial está siendo incorporada en ámbitos que afectan la vida de las personas, como la generación de noticias, la focalización de información y la determinación de la credibilidad, entre otras, con base en información digital. Sus efectos deben analizarse considerando que, aunque no sea autónoma, sus acciones tienen consecuencias en el mundo real. Sus sesgos y falta de discernimiento ético pueden derivar en difusión de desinformación, manipulación de las audiencias y afectación de derechos humanos.

Por ejemplo, el uso de la inteligencia artificial por medios de comunicación permite generar contenidos personalizados de manera automatizada, según el perfil de cada usuario. Esto, sumado a la falta de supervisión humana, puede provocar la propagación de noticias falsas o tendenciosas, creando brechas sociales. También se han dado casos donde este tipo de desinformación deriva en pánico colectivo y acciones violentas.

Ante esto, es importante la regulación ética de la inteligencia artificial para prevenir daños. Se requiere transparencia sobre el uso de estos algoritmos y mecanismos para detectar y corregir sesgos. Asimismo, es clave la participación humana en la supervisión de los sistemas automatizados para garantizar la calidad de la información, así como para prevenir los posibles daños derivados.

En definitiva, la inteligencia artificial plantea complejos dilemas éticos dada su capacidad de afectar vidas humanas, sobre todo cuando su uso se hace con intenciones de obtener beneficios económicos, y no con el objetivo de mejorar la calidad de vida de las personas. De modo que, aunque no tenga autonomía moral, sus efectos deben valorarse bajo consideraciones éticas por quienes la crean y aplican. Esto obliga a los humanos a hacerse responsables de las consecuencias de su uso en bien de la sociedad.

2. METODOLOGÍA

El presente trabajo de investigación tiene un enfoque cualitativo de corte documental, en el cual se emplearon principalmente el método inductivo y la hermenéutica. La elección de estas aproximaciones metodológicas responde a la necesidad de construir conocimiento sobre la autonomía de la inteligencia artificial a partir del análisis crítico de fuentes documentales. El método inductivo resulta pertinente, dado que permite derivar conclusiones generales a partir de premisas particulares. En este caso, se estudió el concepto de autonomía en seres humanos, retomando principalmente las aportaciones de Kant, para luego analizar si la inteligencia artificial cumple con los mismos elementos, en cuyo caso se le podrá considerar autónoma.

Por su parte, la hermenéutica resulta de utilidad para la interpretación contextualizada de los documentos consultados, mediante la aplicación de estos métodos se busca argumentar de forma sólida la falta de autonomía en la inteligencia artificial. El proceso conlleva la revisión crítica de fuentes primarias y secundarias, así como el razonamiento lógico para analizar las prácticas tecnológicas en concordancia con los postulados teóricos de la filosofía kantiana.

3. RESULTADOS

Acercamiento al concepto de autonomía

Cuando se habla de autonomía vienen a la mente principalmente dos conceptos: la autonomía de las instituciones, que corresponde a la libertad para el ejercicio y obtención de recursos, es decir, que no depende de los recursos que otra institución, el gobierno, personas o empresas le proporcionan para llevar a cabo la actividad para la que fue creada y; la autonomía como sujetos o individuos, que se refiere a la capacidad de los individuos para pensar y decidir de manera libre, sin influencias ni manipulaciones externas, sin la necesidad de recibir instrucciones de cómo debe pensar, a donde debe ir y como se debe comportar, guiado solo por las normas morales y jurídicas para una sana convivencia social, por ello, esta cualidad se considera un valor importante en la ética de las personas.

En este contexto, los individuos pueden actuar según su conveniencia, siempre dentro de un marco normativo. Pueden obtener recursos económicos para su bienestar y satisfacción personal y decidir cómo llevar a cabo estas acciones. Sin embargo, las acciones inmorales están sujetas a sanciones de acuerdo con las leyes, ya sean morales o jurídicas. En México, a partir de los 18 años se entiende que la persona ya es mayor de edad, por tanto, las personas son consideradas autónomas y deben asumir la responsabilidad de sus acciones. No obstante, según Galindo (2012), la autonomía depende del desarrollo cognitivo y se alcanza entre los 17 y 25 años, y el objetivo de dar un carácter de independencia es formar individuos morales, donde la moralidad se considera el fin último.

En ese sentido, la autonomía es un concepto más complejo de lo que parece. Las personas, en el ejercicio de su libertad, realizan acciones morales de manera autónoma, lo que requiere un análisis filosófico. Kant (2013) afirma que la autonomía es el principio fundamental de la moral. Él establece el principio de autonomía de la voluntad y el imperativo categórico, que coloca por encima de la voluntad individual; una ley universal de máximas. La autonomía permite que las personas actúen según su voluntad, pero deben considerar las consecuencias de sus actos y seguir el imperativo categórico para garantizar la moralidad y evitar dañar a otros.

Kant se pregunta: "¿Qué es la libertad de la voluntad sino autonomía, es decir, la propiedad de la voluntad de ser una ley para sí misma?". Para evitar transgredir las máximas, la voluntad debe someterse a la ley moral y actuar de manera razonable. La voluntad no puede ser ejercida libremente, sino que está regulada por reglas morales que los individuos establecen por su propia voluntad. Esto implica tratar a las personas como fines en sí mismos y no como medios para alcanzar un fin.

A pesar de estos ideales, la universalización de la ética puede parecer ingenua, ya que asume que las personas actuarán correctamente guiadas solo por la razón. Esta idea de moral universal es respaldada por Galindo (2012), quien sugiere que la moral kantiana busca que las acciones se ajusten a la ley universal. En la actualidad, es más razonable considerar la autonomía como un derecho alcanzado a medida que se demuestra la responsabilidad en el ejercicio de la libertad. En este enfoque, la moralidad no se deja al arbitrio de la razón, sino que se basa en la responsabilidad moral. Así, los actos autónomos son deliberados, responsables y reflejan la libertad de la persona.

Según Badillo (2003), la autonomía significa ser dueño de uno mismo y establecer las normas para guiar la propia vida. Solo en este contexto se logra la libertad individual. Martínez (2007) sugiere que la autonomía también se relaciona con la autodeterminación. Todo esto arroja luz sobre el concepto de autonomía y nos lleva a cuestionar la autonomía de la inteligencia artificial. ¿Pueden las máquinas ser dueñas de sí mismas y, por lo tanto, autónomas y autodeterminadas? Al parecer, las máquinas son propiedad de los humanos y carecen de los requisitos para la autonomía, pero algunas tecnologías muestran ciertos rasgos de autonomía, por ello, es crucial evaluar la moralidad de la interacción entre estas tecnologías y las personas.

Esto nos lleva a preguntarnos sobre la inteligencia artificial que toma decisiones sin intervención humana, y si las máquinas pueden hacerse responsables de sus acciones. También debemos considerar el impacto de las máquinas que requieren recarga de energía por intervención humana, o la influencia del desarrollo y entrenamiento en su autonomía. Surgen entonces las preguntas de si ¿Las máquinas pueden ser conscientes?, o si ¿la autonomía de las

personas es un proceso consciente? Además, debemos analizar si el uso de inteligencias artificiales que generan contenido informativo o elabora predicciones sin razonamientos éticos, afecta a la sociedad.

Según la UNESCO (2018), las máquinas carecen de autonomía moral, ya que no tienen voluntad propia y siguen los objetivos que les imponemos. Esto sugiere que la autonomía de la inteligencia artificial es puramente técnico-operativa y que su intencionalidad depende de los desarrolladores. Esto está estrechamente relacionado con la ética de las personas, a pesar de las propuestas para otorgar derechos a los robots por parte del Parlamento Europeo (2017).

Considerando los elementos que constituyen la autonomía de los seres humanos, se puede observar que algunas inteligencias artificiales ya exhiben características de autonomía. Por ejemplo, el EcoBot III, desarrollado por la Universidad de Bristol en el Laboratorio de Robótica de Bristol, es la tercera generación de la serie de robots EcoBot I y II, y se destaca por su mejora en autonomía. El EcoBot III es un prototipo avanzado que puede buscar su propio alimento para generar la energía necesaria para su funcionamiento. Este robot se alimenta de desechos orgánicos, que procesa utilizando una celda microbiana. Su principal utilidad radica en la limpieza de alcantarillas, lugares inaccesibles o desagradables para los humanos. La idea detrás de estos robots es permitirles operar con la menor intervención humana y durante largos periodos de tiempo (Aradas, 2012; Ieropoulos, Greenman, Melhuish y Horsfield, 2010).

Esta autonomía alimentaria le permite operar durante un período prolongado, incluso hasta "20 o 30 años", lo que significa que una vez en funcionamiento, puede realizar tareas programadas sin requerir intervención humana adicional (Aradas, 2012). La versión EcoBot IV, desarrollada en 2015, se caracteriza por ser más compacta y reproducible. Para lograrlo, se fabricaron unidades a menor escala con diferentes materiales, el EcoBot-IV ha demostrado ser capaz de controlar sus niveles de líquido, desplazarse hacia la fuente de alimento más cercana y deshacerse de sus propios residuos. Luego de recolectar alimento, procesa los alimentos y el agua, distribuyéndolos de manera equitativa a todas las celdas (Laboratorio de Robótica de Bristol, 2015; Papaharalabos, Stinchcombe, Greenman, Melhuish y Ieropoulos, 2014).

Si la autonomía de la inteligencia artificial se limitara a estas tecnologías utilizadas para realizar tareas en lugares inhóspitos o desagradables para los humanos, no habría razón para preocuparse por sus efectos en las personas. Sin embargo, es esencial analizar el uso de la inteligencia artificial en el procesamiento de lenguaje natural, elaboración de predicciones y la toma de decisiones. En el contexto del procesamiento de lenguaje natural, desde 2014 se utiliza para la generación automatizada de noticias. Por ejemplo, *The Washington Post* empleó *Heliograf* para redactar notas en los Juegos Olímpicos de Río de Janeiro en 2016, y en 2019, produjo alrededor de 850 artículos con esta herramienta. *Associated Press* generó "más de tres mil historias cada trimestre" en 2019, y *Le Monde*, el periódico francés, utilizó *Syllabs* para crear 150,000 notas web que abarcaron 34,000 municipios y 2,000 cantones en las elecciones de 2015 (Torres, 2020; Túñez-López, Toural-Bran y Cacheiro-Requeijo, 2018).

Estos ejemplos de generación automatizada de noticias demuestran la autonomía de la inteligencia artificial en la creación de contenidos, pero adicionalmente requiere acceso a información y a la red de internet para desplegar los contenidos. Sin embargo, también se ha observado que la combinación de tecnologías, como las utilizadas para generar texto y plataformas de distribución que personalizan contenidos, puede alimentar la desinformación (Flores-Saviaga, 2020; Pariser, 2017). Además, al adaptar noticias específicamente al perfil de cada individuo, es más probable que las personas crean en ellas y las consideren verídicas si coinciden con sus opiniones, sin cuestionar las fuentes o considerar opiniones opuestas (Flores-Saviaga, 2020).

Por tanto, surge la necesidad de examinar la moralidad de las acciones de la inteligencia artificial. Aunque la tecnología no puede razonar ni dar explicaciones de sus acciones, es posible emitir juicios de valor sobre su conducta en relación con sus creadores. Las personas y las empresas que desarrollan estas tecnologías a menudo persiguen objetivos económicos o políticos, pero utilizan la tecnología para persuadir a las personas, se apoyan en principios éticos y directrices para legitimar su uso, a pesar de que puedan operar al margen de la ética y violar los derechos humanos (Kaiser, 2019; Vercelli, 2018; Zuboff, 2020).

Surge entonces la pregunta de si ¿la inteligencia artificial puede desarrollar una autonomía similar a la de los humanos? o, ¿Es posible que exista una inteligencia artificial general capaz de tomar decisiones libres y razonadas sin influencias externas, más allá de las decisiones automatizadas como las de los algoritmos de generación de texto?. Desde una perspectiva humanista, es difícil concebirlo, ya que la electrónica difiere significativamente de la biología en términos de libertad y autoconciencia. A pesar de ello, existen discusiones y argumentos que defienden lo contrario (Bartra, 2019; Floridi y Sanders, 2004).

Cuando se examina la propuesta de Benjamin Franklin (1972), se puede trazar un paralelismo entre el funcionamiento de una inteligencia artificial y el proceso de toma de decisiones morales humanas. Franklin, en su respuesta a la carta de su amigo Joseph Priestley, quien le solicitó orientación en una decisión de gran relevancia, plantea una perspectiva interesante. La disyuntiva en cuestión involucra la renuncia al cargo de ministro en la iglesia Mill Hill Chapel para aceptar una oferta como asistente bibliotecario y educador de los hijos de Lord Shelburne. A pesar de reconocer la falta de premisas para aconsejar directamente a su amigo, Franklin introduce lo que se denomina *álgebra moral o prudencial*.

Este enfoque propuesto por Franklin implica un método peculiar para la toma de decisiones cruciales. Consiste en dividir una página en dos, en una de las secciones se anotan los aspectos favorables (*pros*) y en la otra, los desfavorables (*contras*). Durante un período de tres o cuatro días, se agregan argumentos en ambas secciones de la página. Una vez que se ha recopilado toda la información relevante, se otorgan pesos a los argumentos en función de su importancia. Cuando se identifican dos argumentos con pesos equivalentes en lados opuestos de la página, se eliminan de la consideración. Si se encuentra que un argumento favorable es igual en peso a dos argumentos desfavorables, se descartan los tres. De manera similar, si dos argumentos en *contra* tienen el mismo peso que tres argumentos a favor, se eliminan los cinco. Este proceso de equilibrio continúa hasta que se alcanza una decisión. Si después de uno o dos días no surgen nuevos elementos, se toma una determinación en concordancia con el resultado. La metodología de Franklin representa una aproximación interesante para abordar decisiones complejas, lo que suscita una conexión con el modo en que las inteligencias artificiales emprenden sus propias elecciones morales.

Esta observación plantea la posibilidad de atribuir una dimensión moral a las acciones de la inteligencia artificial. No obstante, persiste la incógnita: ¿puede la inteligencia artificial alcanzar la autonomía? Desde una perspectiva lógica, la respuesta a esta interrogante tiende a ser negativa, esto se debe a que su funcionamiento y ejecución de acciones siguen patrones similares a los del álgebra moral, ya que opera siguiendo instrucciones y no ejerce su propia libertad individual. En esencia, la inteligencia artificial continúa respondiendo a directrices humanas en lugar de actuar de manera autónoma y voluntaria.

Es evidente que la toma de decisiones ya sea para seres humanos o inteligencias artificiales, depende en gran medida de la información disponible, cuanto más completa y clara sea la información, más sólidas serán las bases para tomar decisiones acertadas. En el caso de las inteligencias artificiales, la calidad de la información es un factor determinante, ya que una mayor calidad reduce la inserción de sesgos y contribuye a que sus elecciones sean más coherentes con principios éticos y morales.

Derechos humanos y desarrollo de la inteligencia artificial

Es de vital importancia distinguir entre autonomía y automatización. La autonomía, tal como se ha definido previamente, abarca la posesión del libre albedrío, la libertad y la voluntad en la toma de decisiones, un rasgo que, hasta la fecha, parece inalcanzable en el ámbito de la inteligencia artificial. La automatización, por su parte, se refiere a la ejecución de tareas repetitivas. En el contexto de los algoritmos de procesamiento de lenguaje natural, esto significa analizar grandes cantidades de información para generar nuevos textos. No obstante, es imperativo destacar que esta acción no se debe a una decisión propia de la máquina, sino a la programación de un algoritmo que opera mediante cálculos matemáticos. Esto no implica la creación de nuevo conocimiento, sino más bien la reinterpretación de texto existente para construir contenido nuevo, como si fuera información completamente original.

La libertad, cuando se considera como un elemento fundamental de la autonomía individual, se ve amenazada en la sociedad actual, ya que, en muchos casos nuestros actos están influenciados por las decisiones tomadas por algoritmos. Estamos atrapados en una especie de cárcel invisible de la libertad, donde consumimos lo que los algoritmos determinan y actuamos de acuerdo a las decisiones que estos algoritmos han tomado previamente.

Un ejemplo de esto se encuentra en los algoritmos de procesamiento de lenguaje natural, que redactan noticias con el propósito de mantener a las personas informadas. Sin embargo, estas noticias a menudo carecen de lógica y reflejan fuertes intereses políticos, comerciales y de control social. Las plataformas como Facebook también se encargan de perfilar a los usuarios para ofrecer servicios de contenidos personalizados, lo que permite que los algoritmos de generación de noticias produzcan contenido que se ajuste a las ideologías de los lectores. Esta convergencia entre el perfil de los usuarios y el procesamiento de lenguaje natural crea un entorno digital que resulta propicio para ejercer control sobre las masas, a menudo de manera silenciosa y desapercibida.

En este sentido, la inteligencia artificial ha logrado emular ciertos aspectos de la autonomía humana. Por ejemplo, robots como el EcoBot III pueden buscar, procesar y desechar su propio alimento, lo que representa un significativo avance en términos de autonomía operativa. No obstante, es fundamental comprender que estos robots siguen dependiendo de la programación realizada por seres humanos. No han desarrollado una autonomía en el sentido de tomar decisiones por iniciativa propia. Su autonomía se reduce a la operación y se encuentra determinada por las instrucciones de programación.

Este principio se extiende igualmente a las inteligencias artificiales centradas en el procesamiento de lenguaje natural. Estas máquinas operan conforme a las instrucciones programadas para procesar texto, sin mostrar un aprendizaje autónomo para mejorar su comprensión del lenguaje natural. En lugar de eso, se basan en grandes cantidades de información y generan texto en función de los temas predefinidos por quienes poseen el algoritmo. Aquí, la tecnología y los algoritmos siguen instrucciones, y aunque los humanos reaccionan conforme a sus resultados, creyendo que son moralmente aceptados, estas máquinas carecen de la capacidad de atribuir significados propios a sus acciones.

Al sopesar si la inteligencia artificial puede alcanzar autonomía en sus decisiones, comparándola con la autonomía humana, es importante recordar las palabras de Coeckelbergh (2021): "Todavía hoy se puede objetar que somos seres vivos encarnados, conscientes y creadores de significado, cuya naturaleza, mente y conocimiento no pueden explicarse mediante la comparación con máquinas". En línea con esta perspectiva, las máquinas creadas por seres humanos, a pesar de poder emular el comportamiento humano, no pueden desarrollar una autonomía y conciencia análogas a las humanas, ya que sus acciones dependen por completo de la programación y los datos disponibles.

Díaz del Campo-Lozano y Chaparro-Domínguez (2018) han señalado que el surgimiento del Big Data, ha allanado el camino para el desarrollo de lo que hoy denominamos periodismo de datos, una evolución que ha dado lugar al *periodismo computacional*. Esta última encarnación del periodismo se centra en la automatización del proceso informativo, y su sustento recae en la inteligencia artificial. Sin embargo, esta automatización no está exenta de desafíos éticos significativos.

La ausencia de intervención humana en la creación de contenido no garantiza que la información resultante sea inherentemente precisa o veraz. La autonomía de la inteligencia artificial, en este contexto, depende en gran medida de la calidad de los algoritmos utilizados, la información con la que se alimenta y las intenciones de quienes operan esta tecnología. Los datos pueden ser interpretados de manera incorrecta, lo que, a su vez, puede resultar en la diseminación de información perjudicial y engañosa para los lectores (Díaz del Campo-Lozano y Chaparro-Domínguez, 2018, pp. 1142-1143; Ventura, 2021, p. 27). La necesidad de abordar estos problemas éticos se vuelve más apremiante a medida que la inteligencia artificial y el *periodismo computacional* continúan desempeñando un papel cada vez más relevante en la producción y distribución de información.

Ética y regulación en el uso de la inteligencia artificial

Una vez que hemos acotado la noción de autonomía en el contexto de la inteligencia artificial, enfocándola en aspectos técnicos y operativos, se hace evidente que etiquetar a la inteligencia artificial como ética o antiética debido a sus acciones autónomas es inapropiado. En cambio, las acciones, decisiones y predicciones de dicha tecnología, en relación con la ética humana, subrayan la necesidad de una regulación cuidadosa y directrices para guiar su desarrollo. Esto es fundamental para evitar que las consecuencias de las acciones de la inteligencia artificial afecten negativamente a las personas y erosionen los principios éticos fundamentales (Coeckelbergh, 2021, p. 59).

Entonces, ¿cómo podría influir la generación de noticias a través de la inteligencia artificial en las personas? Un ejemplo pertinente se encuentra en los trágicos eventos en Acatlán de Osorio, Puebla, México, donde un individuo que se hacía llamar "el Tecuanito" difundió un mensaje a través de WhatsApp, en el que alertaba sobre secuestros de niños para la extracción de órganos. En respuesta a esta alarma, los habitantes del lugar descubrieron que la policía había detenido a dos desconocidos, y en un acto de *pánico moral*, los sacaron de las instalaciones policiales y los quemaron vivos. Posteriormente, se reveló que las dos víctimas eran albañiles llamados Alberto y Ricardo. En este caso, el mensaje no fue generado por una inteligencia artificial, pero la tecnología sirvió como medio de difusión del texto. Este episodio ilustra cómo un mensaje amenazante puede desencadenar una reacción colectiva de *pánico moral* (Excélsior TV, 2018; Argemí, 2019, pp. 97–101).

Imaginemos ahora el escenario de un *pánico moral* alimentado por una inteligencia artificial, donde la ética de los medios de comunicación y los sesgos de la información utilizada para entrenar el algoritmo se suman a la ecuación. En esta situación, tendríamos un generador de noticias parcializado que distribuye contenido a personas cuyas creencias e ideologías se alinean con la narrativa difundida. El algoritmo reconoce a quiénes dirigir esta información, y estas personas, a su vez, desconocen que están siendo manipuladas por un algoritmo que las comprende mejor que ellas mismas. Se plantea entonces la cuestión de si ¿estas personas están tomando decisiones y pensando de manera autónoma?, la posible respuesta sugiere que la autonomía de una persona como dueña de sus propias decisiones y pensamientos podría estar en entredicho (Excélsior TV, 2018; Argemí, 2019, pp. 97–101).

Es fundamental considerar la supervisión humana en los procesos de generación automatizada de contenidos, especialmente en lo que respecta a los sesgos de dichos contenidos, ya que cuando estos son producto de un proceso automatizado, pueden tener la capacidad de propagar información errónea, poniendo en peligro el derecho a la información y la integridad de las personas, tal como ocurrió en el caso de las personas linchadas en Puebla. La automatización y la desinformación están estrechamente relacionadas, lo que resalta la importancia de supervisar y regular esta tecnología (Ventura, 2019, p. 20; 29).

En cuanto a los robots inteligentes, sostiene Bringsjord (2007) que estos no pueden ser considerados agentes morales, ya que sus acciones están determinadas por sus desarrolladores. Incluso si intentáramos hacer que sus acciones fueran aleatorias, seguirían siendo determinadas por la aleatoriedad programada, lo que plantea interrogantes sobre la autonomía. Esta idea se alinea con la moral kantiana, que requiere la presencia de voluntad y libertad propias para la autonomía. Los robots, al igual que los algoritmos de inteligencia artificial, actúan en función de la intencionalidad humana y carecen de autonomía real (Kant, 2013).

Limitaciones de la autonomía de la inteligencia artificial

En relación con el tema anterior, Sullins (2006) destaca que la agencia moral se encuentra dentro de una red de relaciones que configuran una comunidad de interacción (p. 26). No obstante, su conclusión de que incluso los robots más modestos pueden ser considerados agentes morales con derechos y obligaciones, se enfrenta a un análisis crítico en este apartado. Los argumentos previamente presentados refutan esta postura y no permiten respaldar la idea de que la inteligencia artificial posea autonomía y moralidad en un sentido pleno.

Para que un agente se considere humano, Floridi y Sanders (2004) sugieren que debe cumplir con ciertas condiciones, como interactividad, autonomía y adaptabilidad. Al incluir la condición de autonomía, también concuerda

con la concepción de la conciencia humana propuesta por Bartra (2014), que enfatiza la necesidad de una red neuronal que permita a un ser humano tomar conciencia de su entorno y actuar en consecuencia de manera autónoma. A pesar de que las máquinas se aproximan a esta teoría de la conciencia artificial de Bartra (2019), los argumentos presentados anteriormente, fundamentados en la autonomía de Kant (2013), sugieren que las máquinas aún carecen de verdadera autonomía. Por lo tanto, la posibilidad de que las máquinas posean una moral limitada, como proponen Floridi y Sanders, se desafía.

Siguiendo la línea de pensamiento de Floridi y Sanders (2004), se puede confirmar que la moralidad de la inteligencia artificial no está vinculada a su autonomía, sino más bien al significado que los seres humanos asignan a los artefactos culturales. Cuando una empresa desarrolla una inteligencia artificial que lleva a cabo actividades anteriormente realizadas por personas, se "humaniza" a esta inteligencia artificial y se le atribuye una moralidad. Se considera moralmente buena si respeta las normas y moralmente mala si las viola. Sin embargo, es importante destacar que tanto las normas morales como las acciones de la inteligencia artificial son creadas por seres humanos. En consecuencia, se puede atribuir moralidad a las acciones de la inteligencia artificial, pero, según los argumentos previamente presentados, estas acciones carecen de ética autónoma.

En suma, los humanos han desarrollado diversas tecnologías y se adhieren al discurso de la intención de mejorar la calidad de vida de las personas, la inteligencia artificial es parte de estos desarrollos y pretende ofrecer "[...] la emocionante perspectiva de mejorar la condición humana. Pero hay un lado oscuro potencial de la IA que es difícil de ignorar" (Deloitte, 2019, p. 6).

En lo que refiere a los algoritmos de inteligencia artificial para el procesamiento de lenguaje natural, ya son usados por los medios informativos para la generación de noticias, y según Lindén (2017) Facebook fue acusado de emplear a periodistas para capacitar a sus algoritmos para seleccionar historias para el servicio de noticias que podría interesar a la gente. Para Dörr y Hollnbuchner (2017) esto produce "asimetrías de información" e impiden el cumplimiento de la responsabilidad moral de la audiencia, por no distinguir el contenido creado por algoritmos, y dado que la audiencia controla la demanda y la forma del periodismo algorítmico al concatenarse con Facebook —como ejemplo—, también decide los temas sobre los que se ha de difundir información, lo que supone una burbuja limitada de contenido de interés real.

4. CONCLUSIONES

La inteligencia artificial ha avanzado de manera significativa en los últimos años y ha demostrado su capacidad para realizar una amplia gama de tareas, desde la automatización de procesos industriales hasta la generación de contenido textual. Sin embargo, en medio de este avance tecnológico, es crucial recordar que la inteligencia artificial no puede ser considerada autónoma en el sentido en que lo son los seres humanos. Aunque puede parecer que las máquinas están tomando decisiones por sí mismas, en realidad, sus acciones y decisiones son el resultado de algoritmos y programación diseñados por humanos.

La noción de autonomía implica la posesión de libre albedrío, libertad y voluntad propias, características que son inherentes a los seres humanos pero que no pueden ser replicadas en las máquinas. La inteligencia artificial, por más avanzada que sea, sigue dependiendo en última instancia de las instrucciones y la programación que le proporcionan los humanos. No tiene la capacidad de comprender el significado de la autonomía ni de tomar decisiones basadas en una comprensión ética o moral del mundo.

El discurso que busca atribuir responsabilidades a las inteligencias artificiales es problemático en este contexto. Si bien es cierto que la inteligencia artificial puede llevar a cabo tareas complejas de manera eficiente, no puede comprender ni asumir verdaderamente la responsabilidad por sus acciones. No experimenta remordimientos ni tiene la capacidad de evaluar las consecuencias éticas de sus decisiones. Por lo tanto, la noción de que esta tecnología puede tener responsabilidades morales, es fundamentalmente errónea.

En lugar de atribuir responsabilidades a las máquinas, es esencial reconocer que la responsabilidad recae en los seres humanos que diseñan, desarrollan y utilizan la inteligencia artificial. La ética y la moralidad en el contexto de dicha tecnología deben centrarse en las acciones humanas y en la toma de decisiones que afectan a la sociedad y a los derechos humanos. Es responsabilidad de los desarrolladores de esta tecnología y de las empresas que la utilizan asegurarse de que sus sistemas sean diseñados y utilizados de manera ética y responsable.

Uno de los desafíos más apremiantes es garantizar que la inteligencia artificial se utilice de manera que respete los derechos humanos. La automatización de procesos, la generación de contenido y la toma de decisiones basadas en algoritmos deben ser supervisadas cuidadosamente para evitar consecuencias negativas. La dependencia excesiva de algoritmos de inteligencia artificial en la generación de noticias, por ejemplo, puede llevar a la creación de burbujas de información y asimetrías de información, lo que socava la capacidad de la audiencia para formar opiniones informadas, y distorsiona la calidad de la información.

Para abordar estos desafíos éticos, es necesario implementar regulaciones efectivas que supervisen y controlen el uso de la inteligencia artificial en diversas industrias y aplicaciones. Estas regulaciones deben garantizar la transparencia en la toma de decisiones algorítmicas, la protección de datos y la equidad en el acceso y uso de la tecnología. Además, las empresas deben asumir la responsabilidad de capacitar a sus algoritmos de manera ética y evitar sesgos y discriminación injusta.

En última instancia, la conclusión es clara: la inteligencia artificial no puede ser autónoma en el sentido humano y no debe ser considerada responsable de sus acciones. La responsabilidad recae en los seres humanos que diseñan, controlan y utilizan los algoritmos. La ética y la moralidad en el desarrollo de esta tecnología deben centrarse en garantizar el respeto de los derechos humanos, así como en promover un uso ético y responsable de la misma. La regulación efectiva es esencial para proteger a la sociedad y garantizar que la inteligencia artificial beneficie a la humanidad en lugar de perjudicarla.

Referencias

- Aradas, A. (2012). El robot que tiene que ir al baño. *BBC News Mundo*.
https://www.bbc.com/mundo/noticias/2012/02/120213_tecnologia_robot_poo_aa
- Argemí, M. (2019). *Los Siete hábitos de la Gente Desinformada: Cómo informarse y Tomar Decisiones en las redes sociales*. Penguin Random House.
- Badillo, P. (2003). *Pluralismo, Tolerancia, multiculturalismo: reflexiones para un mundo plural*. Akal, Tres Cantos.
- Bajtín, M. (2011). *Las fronteras del discurso*. Las Cuarenta.
- Bartra, R. (2014). *Antropología del Cerebro: Conciencia cultura y libre albedrío*. Pre-Textos.
- Bartra, R. (2019). *Chamanes y Robots: Reflexiones sobre el efecto placebo y la conciencia artificial*. Anagrama.
- Bringsjord, S. (2007). Ethical Robots: The Future Can Heed Us. *Rensselaer Polytechnic Institute*.
<https://doi.org/10.1007/s00146-007-0090-9>.
- Bustamante, M. (2020). Un robot en la redacción: el impacto de la automatización en las rutinas de trabajo de los periodistas. Los casos de TN, Clarín y Olé. *Observatorio Latinoamericano y Caribeño* 4(1), pp. 81-100.
<https://publicaciones.sociales.uba.ar/index.php/observatoriolatinoamericano/article/view/5440>.
- Coeckelbergh, M. (2021). *Ética de la inteligencia artificial*. En L. Álvarez (Trad.). Cátedra.
- De Lara, A., García-Avilés, J.-A., & Arias-Robles, F. (2022). Implantación de la Inteligencia Artificial en los medios españoles: análisis de las percepciones de los profesionales. *Textual & Visual Media*, 1(15), 1-16.
<https://doi.org/10.56418/txt.15.2022.001>.

- Deloitte. (2019). *Ética en la Inteligencia Artificial: Un nuevo imperativo para empresas, juntas y alta gerencia*.
<https://www2.deloitte.com/content/dam/Deloitte/co/Documents/risk/col-ai-ethics-a-new-imperative-for-businesses-boards-and-c-suites-espREV.pdf>.
- Díaz del Campo-Lozano, J. y Chaparro-Domínguez, M-A. (2018). Los desafíos éticos del periodismo en la era del big data: análisis de códigos deontológicos latinoamericanos. *Palabra Clave* 12(4), pp. 1136-1163.
<https://doi.org/10.5294/pacla.2018.21.4.8>.
- Dörr, K. N. y Hollnbuchner, K. (2017). Ethical Challenges of Algorithmic Journalism. *Digital Journalism* 5(4), pp. 404-419. <https://doi.org/10.1080/21670811.2016.1167612>.
- Excélsior TV. (30 de agosto de 2018). *Hombres linchados en Puebla eran campesinos* [Vídeo]. YouTube.
<https://www.youtube.com/watch?v=YjGf-NJopPs>.
- Flores-Saviaga, C. (2020). Inteligencia artificial, aliada o adversaria de las fake news. En G. Orozco (Coord.), *Inteligencia Artificial en Contenidos Audiovisuales* (pp. 53–76). Productora de contenidos culturales.
- Floridi, L. y Sanders, J. W. (2004). On the morality of artificial agents. *Minds and machines* 14(3), pp. 349-379.
<https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- Founders Online. (1972). *From Benjamin Franklin to Joseph Priestley, 19 September 1772*.
<https://founders.archives.gov/documents/Franklin/01-19-02-0200>.
- Galindo, J. D. (2012). Sobre la noción de autonomía en Jean Piaget. *Educación y Ciencia* 15. pp. 23-33.
<https://dialnet.unirioja.es/descarga/articulo/7981961.pdf>
- García-Avilés, J.-A. (14 de febrero de 2023). Ética de la Inteligencia Artificial en el periodismo: nos jugamos el futuro de la profesión [en línea]. *Master en Innovación en Periodismo*. <https://mip.umh.es/blog/2023/02/14/etica-inteligencia-artificial-periodismo-futuro-profesion/>
- Ieropoulos, I., Greenman, J., Melhuish, C. y Horsfield, I. (2010). EcoBot-III: a robot with guts. *Alife XII Conference*.
<https://mitpress-request.mit.edu/sites/default/files/titles/alife/0262290758chap131.pdf>.
- Kaiser, B. (2019). La dictadura de los Datos. En C. R. Malavé (Trad.): HperCollins.
- Kant, I. (2013). *Fundamentación de la metafísica de las costumbres*. En M. García (Trad.). Centaur Edicións.
- Laboratorio de Robótica de Bristo. (2015). *EcoBot IV - Miniaturización y multiplicación* [en línea].
<http://www.brl.ac.uk/research/researchthemes/bioenergysustainable/ecobotiv.aspx>.
- Lindén, C. G. (2017). Algorithms for journalism: The future of news work. *The Journal of Media Innovations* 4(1), pp. 60-76. <http://dx.doi.org/10.5617/jmi.v4i1.2420>.
- Martínez, J. A. (2007). Autonomía. *Anuario Jurídico y Económico Escorialense*, XL, pp- 711-764.
<https://dialnet.unirioja.es/servlet/articulo?codigo=2267971>.
- Mazo, H. M. (2012). La autonomía: Principio ético contemporáneo. *Revista Colombiana de Ciencias Sociales* 3(1), pp. 115-132. <https://dialnet.unirioja.es/descarga/articulo/5123760.pdf>
- Papaharalabos, G., Stinchcombe, A., Greenman, J., Melhuish, C. y Ieropoulos, I. (2014). *EcoBot IV: A Robotic Energy Harvester*. <https://www.h2fcsupergen.com/wp-content/uploads/2014/01/Microbial-Abstracts.pdf>.
- Pariser, E. (2017). *El Filtro Burbuja: Como la web decide lo que leemos y lo que pensamos*. En M. Vaquero (Trad.). Ciudad de México: Taurus.
- Parlamento Europeo. (2017). *Normas de Derecho civil sobre robótica*. Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL)). https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.pdf.

- Pastor, M.A. (2005). Pablo Badillo o Farrell (coordinador): Pluralismo, tolerancia, multiculturalismo. Reflexiones para un mundo plural. Unia/Akal, Madrid 2003, 262 pp.: [Reseña]. *Anuario de Derechos Humanos*, 6, 774-790.
- Sánchez, J. L. M., y Ruiz, M. J. U. (2020). Inteligencia artificial y periodismo - Artificial intelligence and journalism: una herramienta contra la desinformación. *Revista CIDOB d'Afers Internacionals*, 124, 49-72.
<https://www.jstor.org/stable/26975708>.
- Sullins, P. (2006). When Is a Robot a Moral Agent?. *International Review of Information Ethics*, 6.
<https://doi.org/10.29173/irie136>.
- Torres, G. (2020). Inteligencia artificial, big data y analíticas: medición e interactividad de las audiencias. En G. Orozco (Coord.), *Inteligencia Artificial en Contenidos Audiovisuales* (pp. 21-38). Productora de contenidos culturales.
- Túñez-López, J-M., Toural-Bran, C., y Cacheiro-Requeijo, S. (2018). Uso de bots y algoritmos para automatizar la redacción de noticias: percepción y actitudes de los periodistas en España. *El profesional de la información* 27(4), pp. 750-758. <https://doi.org/10.3145/epi.2018.jul.04>.
- UNESCO. (2018). Inteligencia artificial: promesas y amenazas.
https://unesdoc.unesco.org/ark:/48223/pf0000265211_spa.
- Ventura, P. (2021). *Algoritmos en las redacciones: Retos y recomendaciones para dotar a la inteligencia artificial de los valores éticos del periodismo*. Fundación Consell de la Informació de Catalunya.
- Vercelli, A. (2018). Facebook Inc. - Cambridge Analytica: (des)protección de datos personales y campañas globales de desinformación. *Electronic Journal of SADIO* 18(2). pp. 57-70.
<http://sedici.unlp.edu.ar/handle/10915/135072>.
- Zuboff, S. (2020). *La Era del Capitalismo de la Vigilancia: La lucha por un futuro humano frente a las nuevas fronteras del poder*. En A. S. Mosquera (Trad.): Paidós.