

Universidad Politécnica Salesiana

Robert Fernando Bolaños Vivas (Coordinador)

BIOÉTICA

Inteligencia artificial, cambio climático
y muerte asistida





Carrera de Educación

Grupo de Investigación de Filosofía
de la Educación (GIFE)

Si se considera que la vida es el regalo más valioso para los seres humanos, hoy enfrenta problemas y desafíos que intentan modificarla, limitarla o incluso aniquilarla. La bioética reflexiona sobre estas tensiones que acompañan a la existencia, en su sentido más amplio, con el propósito de encontrar soluciones pertinentes e inclusivas.

La inteligencia artificial y sus implicaciones en la vida humana, junto con las consecuencias del cambio climático y la complejidad de la muerte asistida, requieren un abordaje riguroso, abierto e interdisciplinario que involucre a expertos en ética, derecho, antropología, biología y otras ciencias. Iniciar un diálogo reflexivo sobre la vida y sus desafíos teórico-prácticos resulta urgente, pues las consecuencias de estos problemas repercuten, directa o indirectamente, en la calidad de vida de las personas y en el equilibrio ecosistémico del planeta. Los trabajos aquí reunidos surgen de investigadores que invitan a la reflexión crítica y responsable.



ISBN: 978-9942-52-004-3



9 789942 520043



BIOÉTICA

Inteligencia artificial, cambio climático
y muerte asistida

Robert Fernando Bolaños Vivas
(Coordinador)

BIOÉTICA

Inteligencia artificial, cambio climático
y muerte asistida



ABYA
YALA

2025

BIOÉTICA

Inteligencia artificial, cambio climático y muerte asistida

© Robert Fernando Bolaños Vivas (Coordinador)

© Autores: William Orlando Cárdenas-Marín, Cristian Barreto Calle, Diana Calle Ramírez, Alex Darío Estrada García, Jandry Geomar Chuni Gaona, Joe Llerena Izquierdo, Raquel Ayala Carabajo, Sara Eugenia Madera Gómez, Johan Manuel Méndez Reyes, Fredi Portilla Farfán y Jorge Iván Rengel Maldonado

1ra edición: © Universidad Politécnica Salesiana
Av. Turuhuayco 3-69 y Calle Vieja
Cuenca-Ecuador
P.B.X. (+593 7) 2050000
e-mail: publicaciones@ups.edu.ec
www.ups.edu.ec

CARRERA DE EDUCACIÓN

Grupo de Investigación de Filosofía
de la Educación (GIFE)

ISBN impreso: 978-9942-52-004-3

ISBN digital: 978-9942-52-005-0

DOI: <https://doi.org/10.17163/abyaups.141>

Imagen de portada: Shutterstock

Tiraje: 300 ejemplares

Diseño, diagramación
e impresión: Ediciones Abya-Yala
Quito-Ecuador

Impreso en Quito-Ecuador, octubre de 2025

Publicación arbitrada de la Universidad Politécnica Salesiana

El contenido de este libro es de exclusiva responsabilidad de los autores y las autoras.



Presentación	7
Nuevos horizontes de la filosofía ante el auge de la(s) inteligencia(s) artificial(es) <i>William Orlando Cárdenas-Marín</i>	13
Ética empresarial en la era de la inteligencia artificial: una reflexión desde la filosofía <i>Cristian Barreto Calle y Diana Calle Ramírez</i>	35
Implicaciones éticas de la inteligencia artificial <i>Alex Darío Estrada García y Jandry Geomar Chuni Gaona</i>	61
Hacia una perspectiva ética e integral del uso de la inteligencia artificial aplicada a la educación: apuntes para la reflexión <i>Joe Llerena Izquierdo y Raquel Ayala Carabajo</i>	85
Inteligencia artificial y ética del lenguaje: del atomismo lógico a las redes neuronales artificiales <i>Sara Eugenia Madera Gómez</i>	111
Transhumanismo y la inteligencia artificial: un análisis de sus implicaciones éticas para la humanidad <i>Johan Manuel Méndez Reyes</i>	125
La bioética y el cambio climático: uso humano de plantas medicinales contaminadas <i>Fredi Portilla Farfán</i>	157
La eutanasia en Ecuador: análisis de la sentencia 67-23-In/24 de la Corte Constitucional <i>Jorge Iván Rengel Maldonado</i>	177

Presentación

Los problemas y desafíos actuales relacionados con la bioética son numerosos y urgentes. Dichos aspectos, vinculados a la irrupción de la inteligencia artificial en la vida humana, las consecuencias del cambio climático y la complejidad de la muerte asistida, precisan ser abordados interdisciplinariamente mediante la reflexión crítica desde la ética, la jurisprudencia, la filosofía y las humanidades en general. Esta necesidad reflexiva responde a que las consecuencias prácticas de estos problemas repercuten, directa o indirectamente, en la calidad de vida de las personas y en el equilibrio ecosistémico del planeta. Los trabajos que aquí se presentan provienen de destacados expertos internacionales y locales que abordan esta problemática y someten sus artículos a la capacidad analítica y crítica de los lectores.

En primera instancia, William Orlando Cárdenas Marín presenta algunos horizontes de la filosofía ante el auge de la inteligencia artificial. En concreto, el autor aborda “la posibilidad de una inteligencia no biológica”, explorando las posturas filosóficas en torno a la inteligencia, la consciencia y los “límites conceptuales que han condicionado el debate”. También se reflexiona sobre el papel de la filosofía en el ámbito de las inteligencias artificiales y destaca “su capacidad para generar claridad conceptual, fomentar el pensamiento crítico y contribuir al desarrollo de herramientas como el diseño de *prompts*”; todo, sin olvidar la finalidad práctica de la filosofía que es la consecución de “una vida buena en un entorno donde la IA transforma nuestras formas de existencia y de comprensión del mundo”.

A continuación, Cristian Fernando Barreto Calle y Diana Cristina Calle Ramírez, presentan un análisis filosófico que explora cómo la filosofía puede proporcionar una base sólida para reflexionar sobre las implicacio-

nes éticas de la IA en el entorno empresarial. Los autores afirman que el mundo empresarial y su estrecha relación con los recursos de inteligencia artificial, enfrenta desafíos que van “desde la discriminación algorítmica hasta la invasión de la privacidad”, insistiendo en que “la IA puede generar consecuencias no deseadas si no se gestiona de manera ética”.

El trabajo de Alex Darío Estrada García y Jandry Geomar Chuni Gaona, tiene como objetivo “analizar y describir las implicaciones éticas de la IA y las repercusiones en los contextos sociales”. En efecto, para tener ideas claras sobre la inteligencia artificial, con sus potencialidades y sus riesgos, los autores se ponen cuestiones como: “¿qué es la IA? ¿qué creemos que es la IA? ¿en qué puede convertirse, de acuerdo con las proyecciones humanas, la IA? ¿las máquinas necesitan tomar decisiones éticas? ¿la ética en el diseño de la inteligencia artificial es fundamental? ¿cómo definir cánones éticos para el desarrollo, implementación y evaluación de programas de IA?”. Los autores concluyen que con “adecuadas prácticas de transparencia y responsabilidad [...] para mitigar los sesgos, se podrá crear sistemas de IA justos, inclusivos y capaces de mejorar la vida de las personas, sin importar su origen o condición”.

En la misma línea de la inteligencia artificial, la investigación de Joe Llerena Izquierdo y Raquel Ayala Carabajo explora, “de forma reflexiva y crítica, las implicaciones, fortalezas, riesgos, dificultades y desafíos de esta integración de la IA a la educación, a partir de los hallazgos, conceptualizaciones y reflexiones de autores”. Los investigadores concluyen que la IA, “en lugar de reemplazar a los profesores, tiene el potencial de fortalecer su rol al apoyar y enriquecer la enseñanza”. Por lo tanto, se subraya el papel fundamental del educador, descartando “la sustitución de la intervención humana por la tecnología”. Además, se afirma que es urgente que los docentes se adapten a los recursos de la IA de manera significativa, para lo cual se precisa de una renovada “preparación del profesorado y en la formación del estudiantado”, con la finalidad “que los sistemas educativos diseñen estrategias para integrar la tecnología de manera que fomente la honestidad académica y el pensamiento crítico”.

Sara Eugenia Madera Gómez sostiene la tesis de que “los problemas éticos fundamentales asociados con la IA generativa no son primariamente un problema de algoritmos, sino que derivan directamente de los conjuntos de datos (*datasets*) utilizados para su entrenamiento”. La autora, dejándose ayudar por el atomismo lógico de Bertrand Russell (el lenguaje debe reflejar la estructura de los hechos que describe), afirma que los problemas éticos derivados de la inteligencia artificial “en su esencia”, son problemas del lenguaje. Los datos que se utilizan en el entrenamiento de las IA funcionan como un *lenguaje* que intenta representar la realidad”; pero, así como el lenguaje convencional es “impreciso o sesgado”, los datos pueden contener errores, sesgos y limitaciones que se transmiten a los modelos de IA, lo que genera “consecuencias éticas significativas”. La autora no deja de mencionar el “desafío ético de crear sistemas de IA que no solo sean eficientes y precisos, sino también justos, inclusivos y respetuosos de la diversidad humana”.

La investigación de Johan Manuel Méndez Reyes propone analizar críticamente “las implicaciones éticas que tiene para la humanidad el transhumanismo y la inteligencia artificial”. La propuesta transhumanista busca superar las limitaciones inherentes a la condición humana mediante la manipulación artificial de la evolución corporal, vía farmacéutica y tecnológico-digital; tales modificaciones tecnocientíficas se esfuerzan por alcanzar la inmortalidad, la autonomía, el dominio y hasta la sustitución de la memoria humana por inteligencia artificial. Naturalmente que estos intentos cuestionan la imparcialidad y “la validez científica y la aplicabilidad universal de sus propuestas” y conllevan debates éticos en torno a la dignidad humana y su mismo porvenir; de ahí que es importante y necesario “estudiar sus alcances, efectos y los peligros que su utilización en el ámbito humano”. El autor subraya uno de los principales retos éticos que consiste en distinguir entre lo que es científicamente factible y lo que podría representar un salto hacia un nuevo tipo de ser humano o post-humano. El autor sugiere que, no se trata de una oposición ciega a los avances tecnocientíficos que aseguran una mejor calidad de vida a las personas, sino que es preciso “asegurar que los beneficios se equilibren

con los posibles riesgos y desafíos que estas transformaciones puedan conllevar”. Finalmente, el autor concluye que es urgente “establecer límites morales claros que preserven principios esenciales como la autonomía, la justicia, la seguridad y la transparencia, asegurando que estas tecnologías no comprometan la dignidad humana ni exacerben desigualdades sociales”.

El trabajo de Fredi Portilla Farfán presenta la preocupante situación que resulta del cambio climático y los altos niveles de contaminación. El autor expone la inestable situación social caracterizada por el miedo, la violencia, el consumo y descarte exagerado, la inequidad e injusticia en la participación de los ciudadanos sobre el uso de recursos naturales y la insuficiente cobertura de los servicios sociales básicos como educación y salud. Se afirma que “los sistemas imperantes imprimen su lógica productiva en desmedro de la ética del desarrollo humano, a quien consideran un tornillo más de su enorme fábrica”. Según el autor, hasta por medio de la dimensión antropológica de la religiosidad la religión conspiraría “con los poderes de turno para consolar a los necesitados, sobre invisibles prados floridos de futuros mundos distantes del sufrimiento actual”.

En medio este sombrío panorama, se presenta las enormes potencialidades de la utilización de plantas medicinales en la salud humana; pues “la flora medicinal [es] fuente importante de fármacos”. Sin embargo, el autor constata que la contaminación afecta “la cobertura vegetal y la atmósfera [por la presencia] de metales pesados [que llegan a las] plantas y pueden ingresar a la cadena alimenticia”. Sigue el trabajo de Abad y Poluche (2018) que detecta plomo y dióxido de carbono por sobre los valores límite que establecen la norma técnica en cuatro especies de plantas (manzanilla, toronjil, romero y ortiga) cultivadas en la ciudad de Cuenca. Si bien las plantas medicinales contribuyen a la lucha contra el cambio climático porque capturan gases de efecto invernadero, los resultados muestran que todas las especies, tanto cultivadas como las adquiridas en los mercados, estarían contaminadas y, “desde la bioética, no deberían ser comercializados para el consumo humano”.

El trabajo de Jorge Iván Rengel Maldonado, analiza la situación de la eutanasia en el Ecuador, tomando como punto de partida la Sentencia 67-23-IN/24 de la Corte Constitucional, dictado como consecuencia del pedido que formuló Paola Roldán, quien padecía de una enfermedad neurológica progresiva conocida como esclerosis lateral amiotrófica. Las posturas en torno a la eutanasia en el país han estado tanto a favor como en contra. El trabajo revisa la reflexión de la Corte Constitucional en sus aspectos más relevantes, previa a la decisión adoptada por la mayoría de los jueces, que abre en Ecuador un debate en torno a “la posibilidad de acceder a la eutanasia a las personas quienes cumplan con las especificidades contenidas de la resolución”. Según el autor, el fallo de la Corte exhorta al legislador a regular “de manera adecuada y asertiva el procedimiento y las condiciones específicas para la eutanasia a través de la Ley”. Además, se especifica que “no se trata de un hecho indiscriminado, sino una medida extrema frente un padecimiento o dolor que no puede ser tratado de otra manera”. Finalmente, el autor considera luego del fallo de la Corte “quedan más preguntas que certezas y, muchas de ellas, estarán siempre circunscritas al fuero interno de cada persona”.

Abrigamos la esperanza de que los valiosos aportes reflexivos de los distintos autores en torno a los desafíos que implica la inteligencia artificial, el cambio climático y la muerte asistida para la vida humana, contribuyan para que el lector se enriquezca con herramientas conceptuales críticas que le permitan comprender con claridad dichos aspectos y esté en capacidad de configurar una postura personal, debidamente justificada y argumentada.

Robert Fernando Bolaños Vivas

(Coordinador – Miembro del grupo GIFE, coordina el proyecto Educación y pensamiento crítico-creativo en la era de la Inteligencia Artificial)

Nuevos horizontes de la filosofía ante el auge de la(s) inteligencia(s) artificial(es)

William Orlando Cárdenas-Marín
Universidad Politécnica Salesiana
wcardenas@ups.edu.ec
<https://orcid.org/0000-0003-4696-457X>

La inteligencia artificial (IA) es, sin duda, alguna uno de los temas que al momento recibe mayor atención por parte del público y no es para menos. Esta tecnología repercute —y seguirá haciéndolo— en todas o casi todas las dimensiones de nuestra cotidianidad. En estos dos últimos años, la IA se ha integrado a teléfonos inteligentes, ordenadores personales —que tenemos a nuestro alcance en casa u oficina—, vehículos, instrumentos médicos, herramientas de investigación, entre muchos otros dispositivos, incluso algunos que tal vez jamás nos hubiésemos imaginado como refrigeradores o lavadoras de ropa.

Esa integración en nuestras vidas plantea un panorama cotidiano en el que no es posible eludir esas tecnologías y ante el que tenemos que hacer continuas evaluaciones críticas acerca de los beneficios y posibles riesgos de su presencia. La tecnología es lo suficientemente atractiva y satisfactoria para las tareas cotidianas y se ha integrado en nuestras vidas de manera, relativamente, natural para nosotros; por esto, no podemos pasar por alto que la IA también atravesará esa adaptación, asumiendo que no ha ocurrido aún.

A continuación, desarrollaré algunas ideas en torno a la inteligencia artificial y cómo la filosofía podría decir algo en un mundo en constan-

te cambio. Claro está que la filosofía no es algo que se mueva con vida propia, sino que se hace visible en la figura de quienes la estudian, en el quehacer de aquellos que podríamos llamar filósofos; por esto, pretendo esbozar algunas ideas para la labor filosófica en la era de la IA, ideas que no son definitivas y puede que no resulten del todo novedosas, pero que fortalecerán el argumento en favor de una filosofía que aborde problemas actuales y se relacione con otras ramas del conocimiento.

Abordaremos, en primer lugar, la posibilidad de una inteligencia no biológica, explorando las posturas en torno a la inteligencia y la consciencia, así como los límites conceptuales que han condicionado el debate. Posteriormente, examinaremos el papel de la filosofía en este contexto, destacando su capacidad para generar claridad conceptual, fomentar el pensamiento crítico y contribuir al desarrollo de herramientas como el diseño de *prompts*, entendido como una nueva forma de mayéutica en la interacción con sistemas de IA. Finalmente, veremos que la filosofía, además de articularse con la ciencia y la tecnología, debe retomar su dimensión práctica, orientándose hacia la búsqueda de una vida buena en un entorno donde la IA transforma nuestras formas de existencia y de comprensión del mundo.

La literatura acerca de la IA ha aumentado notablemente en el último par de años. El tema está latente en la mayoría de las disciplinas, permeando los campos de la filosofía e incluso de la educación; podría revisarse, por ejemplo, la guía propuesta por la UNESCO (2021) para la formulación de políticas educativas ante la IA. Las posturas ante estas tecnologías son variadas, difícilmente podríamos decir que el optimismo tecnológico gana al pesimismo, aunque también tenemos a quienes prefieren mantenerse neutrales ante el tema. Si bien en ocasiones podríamos sentir simpatía hacia los optimistas tecnológicos, la realidad arroja datos que debemos analizar. Tal como ha sucedido a lo largo de la historia en los problemas filosóficos, no podemos establecer una postura definitiva que posea la verdad absoluta; por ello, nos queda evaluar los diferentes argumentos para tomar la postura que mejor se adecúe a la realidad hu-

mana y que ocasione el menor impacto negativo o, lo que es igual, que obtenga los mejores beneficios posibles.

Para empezar, consideremos que el concepto de IA no es algo tan nuevo como parece; si bien ha sido noticia casi diaria en tiempo reciente, los planteamientos iniciales de una inteligencia artificial se pueden rastrear hasta los años 50 (Martínez-Freire, 2015; Russell y Nori, 2004). Hablamos de más de setenta años de existencia de un concepto y, luego, un área de investigación con procesos de desarrollo con altibajos. En todo este lapso de desarrollo las nociones de IA han llegado al público por medio de diferentes vías como estudios y desarrollos científicos y tecnológicos para el público académico, pero también por medio de películas, cómics, novelas de ciencia ficción, videojuegos. Desde los inicios del estudio de la IA ha existido controversia, no ha habido una respuesta unánime acerca de su posibilidad, de sus beneficios, de sus riesgos o de sus aplicaciones. Por esto pasemos un momento a desmenuzar la primera pregunta que surge en el público acerca de la IA.

¿Es posible una *inteligencia artificial*?

Esta es la pregunta básica que nos hacemos ante la IA y la respuesta depende del lado de donde provenga. Por un lado, tenemos a los humanistas, quienes niegan la posibilidad de que haya una inteligencia fuera de la humana; por otro, podríamos nombrar a los transhumanistas o a los posthumanistas, para quienes toda capacidad humana puede ser emulada e incluso superada por una máquina o por el mismo humano con modificaciones genéticas o tecnológicas (Diéguez, 2021). Menciono estos tres grupos porque es la división de pensamiento que mayor difusión tiene en la actualidad, no porque sean los únicos (me disculpo si dejo de lado alguna otra forma de comprensión con la que el lector se identifique). No es la idea de este apartado declarar un ganador entre las comprensiones humanista, transhumanista y posthumanista, pero debemos tener algunas consideraciones al debatir sobre las capacidades *puramente humanas*.

A lo largo de la historia hemos tenido múltiples argumentos que tienden a favorecer los atributos del animal humano por encima de las demás criaturas; lo contrario sería extraño, ya que si es el humano quien enuncia los argumentos, difícilmente menoscabará su valía ante otras especies. Esos argumentos en favor del animal humano y en desmedro de los demás seres ha servido de justificación para diferentes abusos; esta analogía la podemos notar, por ejemplo, en la validación de la esclavitud por parte de los poseedores del privilegio de decisión acerca de la condición de humano o animal en otras *razas*, como es el caso de los indios o los negros, bajo el argumento de la posesión o no de un *alma*.

En la actualidad difícilmente entramos en debate acerca de la posesión o de un alma, incluso de un alma en los animales; es más, ni siquiera entramos en debate acerca de la posesión de inteligencia por parte de los animales, debido a la cantidad de estudios que confieren niveles de inteligencia en diferentes especies de animales no humanos (Graziano, 2019; Llinás, 2001), estudios que han llevado, afortunadamente, a la prohibición del abuso de animales en zoológicos y en experimentación en laboratorios de algunos países (Diéguez, 2021). Sin embargo, esta concesión de inteligencia aún se ve limitada, por humanistas un poco más flexibles, a seres de carbono; presumo que esto ocurre por el hecho de coincidir en ciertas raíces biológicas comunes.

Pero cuando se trata de seres de silicio la situación cambia; parece que nuestro límite de concesión llega a aquellas especies que guardan, por lo menos, un mínimo de cercanía a nuestro origen basado en carbono, pero nuestra capacidad de pensamiento no nos permite conferir posibilidad de inteligencia a seres de silicio. Esto no es novedad, Turing (1950) ya cuestionó esta situación en su trabajo acerca de la posibilidad de pensamiento (o inteligencia) en la máquina. ¿Será acaso que nuestra comprensión se ve limitada cuando nos topamos con algo que lleva el apelativo de artificial? Hay una situación bastante interesante en nuestra cotidianidad, una tendencia hacia lo que consideramos *natural*. El concepto de lo natural se encuentra a lo largo y ancho de nuestro día a día,

propaganda que promueve el consumo de lo natural, el llamado hacia un estilo de vida *más natural*, incluso estudios que buscan defender la estabilidad de una naturaleza humana. Pero ¿A qué llamamos natural? ¿Qué es lo natural en el animal humano? Cualquier respuesta apresurada a estas interrogantes podría generar contradicción, por ello es prudente revisarlas un momento.

No pocos estudios a lo largo de la historia han procurado cimentar las bases de lo que podría llamarse *naturaleza humana*¹, pero estos han quedado lejos de ser definitivos y, más bien, han levantado controversia; es más, existen argumentos bastante radicales que contradicen varios de esos argumentos (Diéguez, 2021). Ante la defensa de un humano natural podríamos preguntar: ¿natural en qué época? No es lo mismo el animal humano natural del año 5000 a. C. que el del año 2024 y, si no queremos abrir una brecha tan amplia en el tiempo, no es lo mismo el ser humano natural en el año 2024 que el que lo fue en el año 1924. Ha habido modificaciones en su forma de ver el mundo, en las posibilidades de entender su realidad, en la tecnología que emplea, en sus modos de relación con los demás. Cuando tomamos perspectiva parece que esa noción de *naturaleza humana* alcanza un estatus más cercano al absurdo que al nivel *esencial* que se le intenta atribuir.

Solo demos un vistazo a nuestra vida diaria, nos guarecemos en una casa o departamento con cuatro paredes, empleamos cocinas para la cocción de nuestros alimentos, guardamos esos mismos alimentos en refrigeradores y los calentamos en microondas, usamos ropa de diferentes modelos y marcas, nos bañamos en duchas de agua caliente, nos avocamos al ocio con televisores conectados a internet y que funcionan con energía eléctrica; todo esto ¿es lo natural a lo que nos referimos? Si aun así defendemos una inmutabilidad de lo natural humano, entonces pareciera que estiramos la sábana para que el concepto ampare cualquier cosa que el animal humano haga, por el simple hecho de ser humano.

1 Ver Hobbes, 2008; Hume, 1984; Leibniz, 1983.

Pero la situación cambia cuando se trata de conferir atributos humanos a seres de silicio. Si lo artificial en inteligencia nos incomoda, ¿un cambio de concepto facilitaría la concesión de inteligencia a organismos que no sean de carbono? En la búsqueda de realizar una mejor distinción en el concepto de inteligencia, se va haciendo común dejar de lado el concepto de lo artificial —en gran parte por las contradicciones inherentes al concepto que vimos antes— para dar paso a la adecuación conceptual de inteligencias biológicas e inteligencias no biológicas. Es claro que este cambio en la forma de nombrar la inteligencia de organismos de silicio no elimina por completo el debate acerca de su posibilidad, pero el simple hecho de eliminar del medio el término *artificial* lima varias asperezas y contradicciones propias de las confusiones del lenguaje.

Luego de este cambio conceptual, aún no hemos dado respuesta a la pregunta de este apartado que, reacomodada, diría: ¿es posible una inteligencia no biológica? Esto depende del marco conceptual en el que nos desenvolvamos. Hablar de inteligencia solo a nivel biológico resulta amplio y para nada cerrado, ahora imaginemos llevar esos debates al campo de lo no biológico, ¿nos metemos en un gran lío! Pero si hacemos una reducción hasta dar con lo más básico que podríamos considerar como inteligencia, entendiéndola como el proceso de ingreso de información, planteamiento del problema, revisión de la información en función de la resolución del problema, obtención de la respuesta para el problema planteado²; entonces, ya contamos con modelos de inteligencia no biológica que satisfacen con poca complicación estos requisitos.

Es obvio que no podemos ser muy optimistas con los modelos de inteligencia no biológica con los que contamos al momento ya que, como varios críticos han puntualizado, sus funciones de resolución de problemas se enfocan en tareas bastante limitadas y repetitivas, como jugar al ajedrez o modelar grandes cantidades de lenguaje —Large Lan-

2 Ver Newell y Simon, 1972; Kahneman, 2011; Marr, 1982; Minsky, 1986; Dewey, 1910; Popper, 1963.

guage Models o LLM³—, pero no debemos olvidar que el avance de la tecnología se da a una velocidad mucho más acelerada que la evolución de los organismos biológicos (Kurzweil, 2007), así que no nos debería sorprender la ampliación de las limitaciones de las inteligencias no biológicas dentro de unos años.

Los temores y críticas a la IA en muchas ocasiones versan sobre la posibilidad de que esta adquiera consciencia propia. Este temor ha crecido, en parte, debido a los mundos postapocalípticos planteados por Hollywood en no pocas películas. Cuestionemos por un momento la posibilidad de consciencia en organismos no biológicos.

¿Inteligencia implica consciencia?

Tengamos en cuenta que el término acerca del que trataremos no es *conciencia* sin *s*, que hace referencia a juicios morales o la distinción de bien o mal; sino, *consciencia* que, por decirlo de la forma más simple, refiere a procesos de atención —en inglés resulta más clara la distinción de las palabras *conscience* y *consciousness*—, aunque esta definición se ha visto ampliada por diferentes autores. El tema de la consciencia ha sido trabajado de forma extensa por filósofos y científicos y, si queremos entrar a comprender el debate, podríamos partir de las comprensiones del dualismo, materialismo o incluso el pansiquismo⁴ y en cada una de estas comprensiones podemos encontrar argumentos en pro y en contra de la posibilidad de consciencia en organismos no biológicos, pero de momento no es nuestro interés profundizar en cada una de ellas. Otra distinción que debemos considerar es que inteligencia no implica consciencia; algunos autores plantean la posibilidad de organismos inteligentes que no necesariamente poseen consciencia (Graziano, 2019; Seth, 2021), en este último grupo alcanzamos una mejor comprensión de las inteligencias no biológicas.

3 Ver Dennett, 2006; Martínez-Freire, 2015; Seth, 2021.

4 Ver Dennett, 2006; Denton, 2009; Goff, 2019; Graziano, 2019; Martínez-Freire, 2015; Seth, 2021.

Dadas las diversas comprensiones del término consciencia es evidente el dilema al que nos enfrentamos cuando hablamos de IA. Tenemos a quienes plantean la posibilidad del desarrollo de consciencia a nivel no biológico, mientras existe otro bloque de expertos, que no son pocos, quienes niegan la posibilidad de consciencia en la máquina. No dejemos de lado a los optimistas tecnológicos que postulan la posibilidad de fusión humano-máquina a través de la transferencia de consciencia, entre quienes tenemos como principal promotor a Kurzweil (2007).

Si algo queda claro en esta discusión acerca de inteligencia y consciencia es que no hay una respuesta absoluta, sino múltiples posibilidades de abordaje de la temática. Así mismo, debemos tener claro que, como apreciamos en la historia humana, los avances tecnológicos no esperan a que lleguemos a consensos acerca de su desarrollo, peor aún esperan a que haya una visión unánime acerca de sus usos; pareciese que las humanidades —y me atrevería a decir que también las ciencias sociales— han quedado al margen de esos avances, convirtiéndose en recolectores de los restos que el avance de la ciencia y la tecnología dejan a su paso. Lo anterior nos da una bofetada acerca del rol de las humanidades en el mundo contemporáneo, nos da la sensación de que ya no tienen qué decir y sólo se les permite ser comentaristas del juego científico y tecnológico que se está desarrollando sin ellas, sólo espectadoras, pero nunca partícipes, porque han perdido toda capacidad de incidencia en un entorno copado por teorías científicas y lenguaje técnico.

¿Qué le queda a la filosofía?

Cuando escucho a colegas de filosofía hablar sobre el rol de ésta en el mundo contemporáneo noto la idea de que la filosofía debe tener un puesto privilegiado en este *per se*, como si hubiese en ella un halo sagrado o de grandeza *esencial* hacia la que todos están en la obligación de hacer referencia por el simple hecho de existir o, tal vez, por ser *la madre de todas las ciencias*. Creo que esta forma de pensar es la que ha llevado, en muchos espacios académicos a nivel mundial, a que la filoso-

fía no sea tomada en cuenta más que como una serie de asignaturas que deben revisarse en algunos colegios y universidades porque los jóvenes requieren *formación humanista*, únicamente presentes para rellenar una malla académica —asignaturas que a la primera oportunidad se podrían sustituir por algo de mayor *valor académico*—.

Si consideramos a la filosofía en el sentido al que hago referencia en el párrafo anterior, es más que válida la sentencia de muerte a la filosofía que declararon Hawking y Mlodinow (2010), porque es evidente que en muchos aspectos ha perdido su capacidad de respuesta a los problemas contemporáneos, por sumirse en tradiciones que la tienen desfasada en nuestra época.

Crisis de la filosofía enciclopédica o historicista

Quienes hemos tenido clases de filosofía alguna vez en el colegio recordaremos que la forma en que recibimos *clases* de filosofía fue a través de algún recorrido histórico o enciclopédico. Recuerdo un poco de mis clases de filosofía en el colegio (en cuarto y quinto curso) y lo que viene a mi memoria es únicamente haber escuchado al profesor hablar de Platón y Aristóteles, también recuerdo las consultas que nos enviaba a realizar acerca de estos dos pensadores. No será raro decir que la filosofía no permeó en absoluto las fibras de mi existencia en ese entonces.

Cualquiera podría pensar que por haberse tratado de una filosofía a nivel de colegio entonces el docente debía seguir un currículo rígido organizado por un organismo rector de la educación, entonces escapaban de sus manos los contenidos que debía impartir; pero, si avanzamos hasta el nivel universitario la realidad no es diferente. A nivel de universidad aún encontramos catedráticos que consideran que la filosofía se *enseña* como un recorrido histórico a través de las vidas, obras y lo que dijo cada pensador a lo largo de aproximadamente 2 500 años. Esto transforma la filosofía en una repetición *al pie de la letra* de esos contenidos, como si hacer filosofía se tratase de memorizar todo lo que filósofos anteriores han pensado, una filosofía de la repetición y de la memoria.

Cuando menciono ideas de este tipo en alguna charla o conferencia algunos colegas reaccionan de forma un tanto adversa, tal vez atento contra el canon filosófico o contra alguna tradición autosostenida entre quienes continúan defendiendo esa forma de *dictar clases* de filosofía, o me estoy enfrentando a la *filosofía perenne*. En defensa de mi argumento, considero importante aclarar que no estoy en contra de los clásicos, aunque a veces pudiese usar de manera un tanto jocosa la frase *dejemos que Aristóteles descanse en paz*. Acepto que existe una amplia riqueza conceptual filosófica que no podemos darnos el lujo de descartar, mucho menos menospreciar o pretender que de nada ha servido; mi punto se centra en la forma en que hemos encerrado el quehacer filosófico, ya no *hacemos* filosofía, parece que únicamente *aprendemos* o *memorizamos* lo que los filósofos han dicho. Cuando profundizamos en el pensamiento, en la obra, de un pensador captamos casi de inmediato que su mundo conceptual se esboza alrededor de un problema, su intención es dar respuesta a un problema filosófico y esto es lo que hemos perdido con la creencia de que la filosofía se hace desde la memoria, el historicismo y de manera enciclopédica.

El primer horizonte que nos plantea el mundo contemporáneo en la era de la IA es enfocarnos en los problemas actuales, los problemas que aquejan al humano hoy. Desde ese punto de partida, emplear nuestra capacidad de reflexión y el contenido filosófico de larga tradición permitiría explorar alternativas a estos problemas. Pensemos por un momento lo absurdo que sería empezar un curso IA con la pregunta: ¿qué diría Aristóteles acerca de la IA? La entrada obvia, dependiendo del enfoque que tenga ese curso, sería abordar las problemáticas que salen a la luz a partir de la IA y, desde ese punto, empezar el abordaje con teorías que podrían llegar hasta pensadores antiguos, siempre que ayuden a dilucidar de mejor manera la comprensión del problema.

Lo anterior nos lleva a otra situación, pareciera que la filosofía se ha cerrado en sí misma. A pesar del golpe al ego filosófico que nos pueda causar, debemos aceptar que, desde la modernidad con la ramificación

de diversas disciplinas científicas, estas han solventado mucho mejor los problemas del mundo natural que lo que la filosofía podría haber hecho (Hawking y Mlodinow, 2010), me refiero aquí a la filosofía especulativa, que habla mucho, pero que dice poco. En la actualidad es indudable que la ciencia ha tomado la posta en la explicación del mundo y, también, son los científicos los que se han convertido en una especie de nuevos humanistas, considerando la comprensión de la tercera cultura que hizo Brockman (2007) en su libro que lleva un nombre similar.

Tenemos hoy muchos ejemplos de científicos que llegan a preguntas filosóficas a partir de las diversas áreas como la psicología científica (Pinker, 1995), biología (Dawkins, 1976; Maturana y Varela, 2003), informática (Minsky, 2007), neurología (Dehaene, 2011; Ramachandran, 2008), entre muchas otras. Entonces, ¿los filósofos ya no son necesarios? Esta pregunta me lleva a formular otra: ¿a qué filósofo hacemos referencia? Si nos referimos a los filósofos que consideran la filosofía como un recorrido histórico, memorístico o enciclopédico, me inclinaría por la respuesta negativa, porque no habría discusión, solo repetición. En cambio, si nos referimos al filósofo que se enfrenta a los problemas contemporáneos desde una comprensión crítica y emplea la tradición filosófica para buscar alternativas de respuesta, entonces mi posición cambiaría, ese filósofo es el que requerimos en el contexto actual.

Lo anterior implica que el filósofo se interne en el diálogo con las diferentes ciencias. Necesitamos una filosofía que reflexione los problemas actuales en línea con los postulados científicos, no únicamente desde una perspectiva externa, porque eso nos transformaría nuevamente en especuladores del conocimiento, sino desde dentro de diferentes campos del conocimiento. Mi idea podría chocar con los filósofos que piensan que existe una pureza en la filosofía que les impide entrar en relación con otros dominios, para esos filósofos *puros* el único camino sería cerrarse a lo abstracto o inmaterial, casi a un nivel dogmático. Propongo, por el contrario, la formación de filósofos *impuros*, filósofos que programen sistemas informáticos, que estudien biología, que experimenten en laboratorios de

física, que formen parte de comités de bioética, una generación de filósofos que dialoguen con la ciencia con conocimiento de causa y no únicamente desde fuera, desde la especulación o desde la superficialidad. Por cierto, mi propuesta no es novedosa en el sentido en que ya existen, y han existido, este tipo de filósofos; son los que han ayudado a que las preguntas filosóficas sigan vigentes en diversos espacios y momentos históricos.

El horizonte que he abordado puede alterar algunas susceptibilidades, pero creo que es hacia donde debería dirigirse nuestra atención desde la filosofía si anhelamos que se mantenga vigente. Ahora bien, como he mencionado, no es mi afán oponerme a los clásicos, porque considero que hay valor en la tradición, pero ese valor se hará notable en la medida en que podamos emplear esas filosofías para el análisis de los problemas contemporáneos. A continuación, postularé unos cuantos ejemplos en los que puede hacerse discusión filosófica actual.

El *prompt* como nueva mayéutica

De Sócrates recordamos casi de inmediato en primer lugar su frase *sólo sé que no se nada*; en un segundo momento, su método basado en preguntas, la mayéutica. Pero hace unas líneas atrás realicé una crítica al historicismo ¿Por qué recurrir a un filósofo antiguo ahora? Porque el método de este filósofo nos ayudará a comprender una realidad muy actual frente a la IA, el *prompt*.

La mayéutica socrática, a breves rasgos, es una metodología basada en preguntas que guiaba al interlocutor a descubrir la respuesta *por su propia cuenta*. Consistía en un ir y venir de preguntas y respuestas con la finalidad de aumentar el conocimiento a partir de las respuestas dadas por el mismo interlocutor, causando la percepción de que éste es quien tenía en sí mismo la respuesta a su duda. Esta metodología nos ha llegado a través de los famosos diálogos platónicos en los que Sócrates entra en escena.

A simple vista pareciese que el filósofo simplemente lanza preguntas conforme va fluyendo la conversación, pero nada más opuesto a la

realidad. El arte de la mayéutica consiste, precisamente, en el dominio de un campo del conocimiento por parte del filósofo, quien con su práctica conoce con bastante determinación cómo guiar al interlocutor hacia la respuesta que desea obtener. Podríamos decir que este arte se basa en hacer la pregunta adecuada, en el momento preciso, en la forma correcta. Más adelante en la historia tuvimos otros pensadores que recalcaron la importancia de la pregunta filosófica, como Popper (1979), para quien el valor de la filosofía no se contempla en la respuesta, sino en la pregunta que se realiza. También, en ciencia ficción podemos ver el valor de la pregunta, en la película *Yo, robot* (2004), basada en las historias de robots de Isaac Asimov, en varias escenas cuando el detective dialoga con el holograma del fallecido científico, el primero realiza series de preguntas a las que la máquina responde: “Lo siento, mis respuestas son limitadas, debes hacer las preguntas adecuadas...”, pero cuando estructura de mejor manera su pregunta ya no recibe un mensaje como los anteriores, simplemente le dice: “Esa, detective, es la pregunta correcta...”.

Ahora revisemos por un momento la noción de *prompt* que ha ganado rápida acogida en los últimos dos años en el auge de la IA. Un *prompt* es, en pocas palabras, una cadena ordenada de comandos que permiten al usuario de una IA obtener la respuesta a un problema. Esto parece algo sencillo, pero en una red con millones de datos una ligera diferencia en el comando puede marcar una gran diferencia en el resultado obtenido. Esto ya lo vivíamos de alguna manera en los motores de búsqueda por internet, por ello la necesidad de conocer ciertos conectores booleanos para facilitar los resultados de búsqueda.

Ante esta situación, cuando nos movemos en el mundo de grandes datos, se hace necesario el dominio del planteamiento de *prompts*; lo que explica la gran cantidad de ofertas de cursos de formación en diseño de *prompts*. Al hacer una búsqueda rápida de estas propuestas de formación encontramos lo siguiente: la plataforma *Coursera* nos ofrece, hoy en día, 1708 propuestas, en *Platzi* encontramos 68 resultados y en *Udemy* tenemos la desbordante cantidad de 5079 cursos. La curiosidad me llevó a buscar

la existencia de alguna oferta relacionada con mayéutica y no puedo decir que me sorprendieron los resultados: en *Coursera* encontré 1 curso y en *Udemy* 8 cursos, en *Platzi* ni siquiera se menciona el término.

Lejos de lamentarnos por la poca importancia que se da al concepto mayéutica, esta situación nos da luces hacia los nuevos espacios de acción de la filosofía, campos donde la pregunta retoma su importancia no únicamente en el diálogo con interlocutores humanos, sino en la resolución de problemas con el apoyo de inteligencias no biológicas. Este es un ámbito bastante llamativo al que podríamos prestar atención, pero que requiere que demos mayor valor a una poderosa herramienta filosófica que en algunos espacios académicos del campo de las humanidades, sobre todo en Latinoamérica, se ha dejado de lado: la lógica.

A breves rasgos, lógica es la herramienta (filosófica en primera instancia) que nos ayuda a distinguir el argumento correcto del incorrecto.⁵ La ventaja que tenemos desde la filosofía es que la lógica puede ayudarnos como puente entre filosofía y ciencia, así como entre filosofía y tecnología. Haré aquí un breve paréntesis, la lógica a la que hago referencia es a la simbólica o matemática, que en su evolución ha potenciado el desarrollo de la lógica informática. No me refiero a la lógica silogística, formal o aristotélica que aún se sigue enseñando en colegios y universidades como un camino clásico, pero que de poco o nada sirve para entender los nuevos lenguajes de la ciencia y la tecnología (no dudo que quienes siguen haciendo énfasis en esa lógica son los mismos que ven a la filosofía desde el historicismo).

Si tenemos la lógica a la mano, nos resultaría mucho más práctico involucrarnos en el mundo de las nuevas tecnologías y, por ende, adentrarnos en el diseño de *prompts*, todo esto sin dejar de lado el pensamiento crítico que caracteriza a la filosofía. Veo aquí una riqueza de reflexión al entender la mayéutica y al desarrollar *prompts* a partir del pensamiento filosófico, estoy convencido de que las reflexiones que saldrían de ahí ge-

5 Ver Copi y Cohen, 2014; Gamut, 2002; Suppes y Hill, 1988.

nerarían diálogos interdisciplinarios y propuestas mucho más actualizadas para la comprensión del animal humano en el mundo contemporáneo.

Tenemos que traer de regreso la lógica a los espacios formativos en el área de las humanidades —hago esta aclaración porque en ámbitos de ingenierías y tecnología esta es punto de partida obligado—; no porque esta tenga un puesto privilegiado por sí misma, si defendiendo esto nada me diferenciaría de los colegas de filosofía hacia los que hice referencia párrafos atrás, sino porque es evidente que proporciona al filósofo las herramientas necesarias para estructurar adecuadamente sus ideas y también lo conecta con otros campos disciplinarios. Lo anterior abre ámbitos de reflexión y evita que nos cerremos al diálogo internista en filosofía; es decir, filósofos que leen únicamente a otros filósofos, dialogan solo con ellos y escriben para ellos. Creo que de esta manera la filosofía podría tener espacios donde sea escuchada, no porque se lo *merezca*, sino porque enuncia cosas con sentido y que responden a los problemas que el humano actual se plantea.

El poder del pensamiento crítico

Luego de lo abordado en el apartado anterior quedan ciertas ideas que podríamos profundizar por separado, ideas que involucran al pensamiento crítico. Este es un término tan recurrente, pero a la vez tan ambiguo, en gran parte por la gran cantidad de literatura en la que se dan indicaciones para su desarrollo. No diré que toda esa literatura está errada (aunque sí tenemos que ser selectivos al momento de escogerla), pero considero que debemos hacer una depuración conceptual para entender la base de lo que podemos denominar pensamiento crítico.

Para explicar mejor las aclaraciones que realizaré, partiré de una anécdota. Hace unos pocos años tuve la oportunidad (más bien fue una disposición, pero oportunidad suena más sutil) de desarrollar una asignatura en la institución donde colaboro, esta llevaba por nombre *Pensamiento Crítico*. Cuando revisé los contenidos que posiblemente se debían

abordar estos no terminaron de convencerme, en gran parte porque he desempeñado por algunos años la cátedra de lógica, además mi formación es filosófica. Opté por centrar la asignatura en dos pilares: en un primer momento lógica formal, específicamente lógica simbólica y, en un segundo momento, la lógica informal, en la que se revisen falacias y sesgos del pensamiento. Sentí cierta satisfacción en la forma en que quedó diseñada la asignatura, pero muchos colegas no pensaron igual y generó (y sigue generando hasta la actualidad) cierta controversia. Acontece que el programa que diseñé debía ser impartido por diferentes catedráticos en todos los programas académicos de la institución; la crítica recurrente que he recibido por ese programa ha sido del tipo: “¡Eso no es pensamiento crítico!”, de donde me empecé a preguntar: ¿qué entenderán por pensamiento crítico muchas personas?, no únicamente mis colegas.

En secuencia con la crítica en mención, suelen mencionar frases como: “¡La lógica no es la única forma de pensamiento crítico, también existen otras formas como el pensamiento creativo!” y no estoy en total desacuerdo con esto, pero, a mi forma de ver y no únicamente mía, muchos autores tienen una comprensión similar. Existe una base para el desarrollo del pensamiento crítico y esa base no es otra que la lógica formal e informal (Pinker, 2021; Talbot, 2017). Considero pertinente un trabajo de claridad conceptual.

La palabra *crítica* es entendida de muchas maneras, pero lo más preciso es entenderla como sinónimo de análisis, en filosofía tenemos la conocida crítica kantiana (Kant, 2004; 2007) que nos ubica en la comprensión del término. Si hablamos de pensamiento crítico, entonces estamos haciendo referencia a enfocar nuestras facultades intelectuales en analizar una situación, un problema, un argumento o lo que deseamos comprender. Es por esto por lo que el inicio en el desarrollo del pensamiento crítico no puede ser otro que la lógica formal, que proporciona estructura al pensamiento y al lenguaje, así como la lógica informal, que nos ayuda a protegernos de los engaños del lenguaje cotidiano bajo la figura de falacias y sesgos. Luego, cuando tenemos una base sólida sobre la cual cimentar

nuestro análisis, podemos dar cabida a cualquier pensamiento creativo, emotivo, actos imaginativos, entre otros.

Ahora bien, ¿por qué menciono al pensamiento crítico dentro de los nuevos horizontes de la filosofía? Porque la filosofía, desde sus inicios, constituye la base de desarrollo del pensamiento crítico y hoy más que nunca la crítica se hace necesaria para el abordaje de los problemas del humano contemporáneo en su relación con el entorno. Cuando hablamos del uso adecuado de la IA usualmente recurrimos a decir que debemos fortalecer el pensamiento crítico ante esta herramienta, pero ¿cómo podríamos fortalecer el pensamiento crítico si ni siquiera tenemos clara su distinción?

Si consideramos la lógica formal e informal la base del pensamiento crítico, ampliamos las posibilidades de reflexión que podemos tener ante el auge de las IA, porque podremos evaluar de mejor manera diferentes situaciones, a la vez que esa misma lógica nos servirá de herramienta para conectar con lenguajes científicos y tecnológicos, como vimos en el apartado anterior.

Filosofía y claridad conceptual

Este apartado es clave porque detalla la labor que, realizada desde el inicio de esta argumentación, la aclaración de conceptos. Quisiera recalcar que en ningún momento consideraré a la filosofía como la reducción de los problemas a un orden lingüístico como dieron a entender los filósofos del lenguaje y que varios autores criticaron posteriormente (Popper, 1979); pero es una realidad que muchos problemas de comprensión, sobre todo cuando trabajamos con diferentes disciplinas, se deben a la falta de una aclaración conceptual previa.

¿A qué nos referimos con claridad conceptual? Es el proceso de depuración de un concepto, de un término, con la finalidad de entenderlo con la mayor objetividad posible. Fue que hicimos al aclarar conceptos como inteligencia, consciencia o pensamiento crítico, una indagación que permita comunicarnos en términos aceptados en un mismo sentido por

todos los interlocutores. La crítica aquí puede ser que no todos nos vamos a poner de acuerdo en los significados para un término, pero no es una generalización hacia todo el mundo, sino para las personas involucradas en una discusión específica, es la definición clara de un término en una trama argumentativa, como la definición de variables en un programa informático.

Con el avance de la ciencia y la tecnología ha existido una explosión de nueva terminología que, en muchas ocasiones, sobrepasa las posibilidades de retención y comprensión del público; por esto la aclaración conceptual es necesaria, porque facilita el diálogo entre partes, caso contrario nos ahogaríamos en un mundo de términos y comprensiones subjetivas, lo que generaría más debate por la falta de acuerdo en el significado del término antes que por las implicaciones de tal o cual problema en nuestro entorno.

La claridad conceptual conecta con los puntos anteriores como uno de los horizontes de la filosofía porque nos llama a entender los términos en los que se mueven las nuevas tendencias en IA, es nuestro mecanismo para adentrarnos en otros campos disciplinarios y analizarlos sin el riesgo de cometer falacias, sino entendiendo a qué hace referencia lo que se dice en determinado campo del conocimiento.

Filosofía para la buena vida

La pregunta por la posibilidad de una inteligencia no biológica que tenga consciencia y que pueda hacer cosas que el humano pensó por mucho tiempo que únicamente él podía, nos deja con una especie de sin sentido. Imagino que, en parte, es por esta razón que buscamos defender a capa y espada que el humano es el único que puede realizar esos procesos racionales y cognitivos en sentido estricto, el resto no es más que una imitación *barata* de la grandeza del ser humano.

Pero, ¿por qué no tomar otra postura? Parece que a nada nos lleva esa defensa enardecida de la imposibilidad de un organismo no biológico para ejecutar acciones que pensábamos propias de los organismos bioló-

gicos. Entonces, ¿por qué seguir aferrados a eso? Una postura diferente podría ser volver la vista hacia nosotros, a esas especificidades que como humanos dejamos de lado cuando nos centramos en pesar que lo que nos fundamenta como especie es nuestra extrema productividad, nuestra capacidad racional *superior* a las demás especies, la capacidad —o al menos el deseo— de tener todo bajo control.

Parece que la llegada de la IA nos desestabiliza porque cambia el sentido de lo que pensábamos que nos fundamentaba, nos coloca una suerte de espejo en el que vemos que en lo que nos creíamos únicos en realidad podía replicarse fuera de nosotros. Es claro que existen muchos otros factores relacionados con la IA que generan tensión como son el futuro de los trabajos, el aumento de la desigualdad social, entre otros, pero por la línea del argumento no extenderé mi análisis a esos temas.

Ante esta inestabilidad generada por un mundo tecnológico que nos sobrepasa es conveniente hacer una revisión hacia nosotros y dar otro rumbo a preguntas del tipo: ¿por qué esa máquina no se compara con el ser humano?, ¿qué es lo que hace que yo sea humano? Considero que el espejo que nos pone la IA nos ayudará en gran medida para esta revisión de fundamentos. La búsqueda por entender lo que hace se seamos humanos no es reciente, apareció en los inicios de la filosofía con preguntas básicas como *¿quién soy?* Y la cantidad de respuestas ha sido tan amplia y variada como la cantidad de filósofos y líneas del pensamiento que hemos tenido a lo largo de la historia. Entonces, *¿es otro camino sin salida?* No necesariamente. Recordemos que la versatilidad y valor de la filosofía radica en que no es una ciencia y no es algo cerrado, sino que se abre a todas las dimensiones en que el humano se desarrolla, es por esto por lo que otro de los horizontes de la filosofía contemporánea es cuestionar al individuo para que éste realice revisión constante acerca de lo que es, una actualización constante de las preguntas *¿quién soy?*, *¿de dónde vengo?* y *¿hacia dónde me dirijo?*

El auge de la IA constituye una invitación a pensarnos como especie y como individuo. Aquí entra el trabajo de la filosofía en sentido

práctico, porque es un campo para el que la ciencia y la tecnología no tiene respuesta más allá de lo que nos pueda decir acerca de nuestro funcionamiento biológico. La pregunta filosófica no es sólo pregunta por cómo mi organismo funciona, sino, también, pregunta por el sentido de mi existencia individual ante tantas otras existencias individuales biológicas y no biológicas.

Desde la antigüedad podemos ver que la filosofía además de la búsqueda de la comprensión del funcionamiento del universo, también se estructura como la búsqueda de la vida buena para el animal humano en relación con su entorno. La filosofía no solo pretende conocer el entorno, sino que permite transitar por esta vida de la forma más armoniosa en la que podamos; por esto, considero relevante retomar, además del diálogo con otras disciplinas científicas y tecnológicas, el rol de la filosofía que acompaña la búsqueda de sentido por parte de los individuos.

Existen varios caminos filosóficos para la búsqueda de una vida armoniosa, algunos pueden seguir la lectura y práctica del pensamiento clásico como el estoicismo, epicureísmo, la práctica de las virtudes en sentido aristotélico, entre otros (Pigliucci, 2018). Hay quienes encuentran un camino en la meditación (en sentido filosófico) acerca de la existencia y las acciones que ejecutan en su vida cotidiana; otros pueden encontrar el camino en la consultoría filosófica (Guerra Bravo *et al.*, 2007), donde un profesional de la filosofía dialoga con nosotros para acompañar el desenredo de problemas existenciales empleando herramientas filosóficas (como la claridad conceptual que revisamos antes).

Existen muchos espacios para el desarrollo de la filosofía en la era de la IA, estos espacios constituyen horizontes hacia donde podríamos apuntar si queremos que la labor filosófica no se pierda en la historia y, más bien, se reivindique como un espacio desde el que podemos dialogar con la ciencia y la tecnología y, a la vez, responder las preguntas humanas que siguen siendo tan vigentes como antaño. Comprendo que he pasado por alto muchos otros espacios en los que la filosofía puede desarrollarse actualmente ante los avances del mundo, pero el objetivo

de mi argumentación es plantear unas cuantas nociones que sin duda se seguirán expandiendo, bien sea para coincidir con estas ideas, así como para contraargumentar, pero de la mejor forma en que podemos, haciendo filosofía desde los problemas con los que nos encontramos hoy.

Referencias bibliográficas

- Brockman, J. (Ed.). (2007). *El nuevo humanismo y las fronteras de la ciencia*. Kairós.
- Copi, I. y Cohen, C. (2014). *Introducción a la lógica*. Limusa.
- Dawkins, R. (1976). *The selfish gene*. Oxford University Press.
- Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. Oxford University Press.
- Dennett, D. (2006). *Dulces sueños. Obstáculos filosóficos para una ciencia de la conciencia*. Katz.
- Denton, D. (2009). *El despertar de la consciencia. La neurociencia de las emociones primarias*. Paidós.
- Dewey, J. (1910). *How we think*. D.C. Heath & Co.
- Diéguez, A. (2021). *Cuerpos Inadecuados. El desafío transhumanista a la filosofía*. Herder.
- GAMUT, L. T. F. (2002). *Introducción a la lógica*. Editorial Universitaria de Buenos Aires.
- Goff, P. (2019). *Galileo's error. Foundations for a new science of consciousness*. Pantheon Books.
- Graziano, M. (2019). *Rethinking consciousness. A scientific theory of subjective experience*. W.W. Norton & Company.
- Guerra Bravo, S., Chacón Padilla, G., Montalvo Zumárraga, A. y Villamarín Carrascal, M. (2007). *El arte de vivir con sentido. Manual de consultoría filosófica latinoamericana*. Abya Yala.
- Hawking, S. y Mlodinow, L. (2010). *El gran diseño*. Crítica.
- Hobbes, T. (2008). *Human nature and De corpore politico* (J. Gaskin, Ed.). Oxford University Press.
- Hume, D. (1984). *Tratado de la naturaleza humana* (F. Duque, Ed.). Orbis.
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kant, I. (2004). *Crítica de la razón pura*. Tecnos.
- Kant, I. (2007). *Critique of Judgement*. Oxford University Press.
- Kurzweil, R. (2007). La singularidad. En J. Brockman (Ed.), *El nuevo humanismo y las fronteras de la ciencia* (pp. 271-290). Kairós.

- Leibniz, G. (1983). *Nuevos ensayos sobre el entendimiento humano* (J. Echeverría Ezponda, Ed.). Editora Nacional.
- Llinás, R. (2001). *I of the vortex. From neurons to self*. The MIT Press.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman.
- Martínez-Freire, P. (2015). *La nueva filosofía de la mente*. Gedisa.
- Maturana, H. y Varela, F. (2003). *El árbol del conocimiento*. Lumen.
- Minsky, M. (2007). ¿Qué vendrá después de las mentes? En J. Brockman (Ed.), *El nuevo humanismo y las fronteras de la ciencia* (pp. 251-270). Kairós.
- Minsky, M. (1986). *The society of mind*. Simon & Schuster.
- Newell, A. y Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.
- Pigliucci, M. (2018). *Cómo ser un estoico. Utilizar la filosofía antigua para vivir una vida moderna*. Ariel.
- Pinker, S. (2021). *Racionalidad. Qué es, por qué escasea y cómo promoverla*. Paidós.
- Pinker, S. (1995). *The language instinct. The new science of language and mind*. Harper Perennial.
- Popper, K. (1979). Cómo veo la filosofía. En C. Bontempo y J. Odell (Eds.), *La lechuza de minerva ¿Qué es filosofía?* (pp. 57-72). Cátedra.
- Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge*. Routledge & Kegan Paul.
- Proyas, A. (2004). *Yo, robot (I, Robot)* [Video recording]. 20th Century Fox.
- Ramachandran, V. S. (2008). *Los laberintos del cerebro*. Libros La Liebre De Marzo.
- Russell, S. y Norvig, P. (2004). *Inteligencia artificial. Un enfoque moderno*. Pearson.
- Seth, A. (2021). *Being you. A new science of consciousness*. Dutton.
- Suppes, P. y Hill, S. (1988). *Primer curso de lógica matemática*. Editorial Reverté Colombiana.
- Talbot, M. (2017). *Critical reasoning. Romp Through the Foothills of Logic for Complete Beginners*. Metafore.
- Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460.
- UNESCO. (2021). *Inteligencia artificial y educación. Guía para las personas a cargo de formular políticas*. UNESCO.

<https://doi.org/10.17163/abyaups.141.2>

Ética empresarial en la era de la inteligencia artificial: una reflexión desde la filosofía

Cristian Barreto Calle
Universidad Politécnica Salesiana
cbarreto@ups.edu.ec
<https://orcid.org/0009-0000-7528-8705>

Diana Calle Ramírez
Universidad Politécnica Salesiana
dcalle@ups.edu.ec
<https://orcid.org/0009-0007-1672-2779>

Introducción

En la actualidad, la IA (inteligencia artificial) se considera una fuerza omnipresente en la sociedad de la información, con aplicaciones que van desde asistentes virtuales hasta sistemas de conducción autónoma. Al volverse más indispensable en nuestras vidas, surge la necesidad de examinar, críticamente, cómo las empresas abordan los desafíos éticos asociados con su desarrollo y despliegue. Ha emergido como una tecnología transformadora en el mundo empresarial que ofrece oportunidades para mejorar la eficiencia, la productividad y la innovación. Sin embargo, su adopción plantea importantes cuestiones éticas que las empresas deben abordar de manera proactiva. Desde la discriminación algorítmica hasta la invasión de la privacidad, la IA puede generar consecuencias no deseadas si no se gestiona de manera ética.

Según un estudio de Accentur, solo el 35 % de los consumidores globales confía en la forma en que las organizaciones implementan la tecnología de IA y el 77 % cree que las organizaciones deben rendir cuentas por el mal uso que hacen de la IA. En este sentido, es fundamental la reflexión ética para alentar a las organizaciones que desarrollan IA en sus procesos a propiciar un marco de IA sólido y transparente e impulsar un desarrollo orientado desde una perspectiva ética.

La ética empresarial, entendida como una conceptualización encaminada a presentar principios y valores que guían las acciones y decisiones de las empresas, se enfrenta a nuevos desafíos y oportunidades en esta era de la IA. En este artículo de investigación y análisis filosófico, se explora cómo la filosofía puede proporcionar una base sólida para reflexionar sobre las implicaciones éticas de la IA en el entorno empresarial. Se analizan cuestiones fundamentales como la responsabilidad moral, el sesgo algorítmico, la privacidad y la seguridad desde una perspectiva filosófica, con el objetivo de plantear *insights* y recomendaciones para promover una ética empresarial sólida y fundamentada en la era de la inteligencia artificial desde este planteamiento teórico-reflexivo. En cuanto a los resultados, se analizan una serie de datos relacionados con el auge de la IA en el ámbito económico y empresarial, para finalmente establecer conclusiones desde el ámbito filosófico que orienten el devenir del proceso interrelacional de la ética empresarial y el uso de la inteligencia artificial bajo el paradigma de valores y principios orientados desde la responsabilidad empresarial.

Ética de la inteligencia artificial

Iniciamos con una reflexión de la ética desde el ámbito filosófico orientada como el estudio sistemático de valores, principios y acciones morales y éticas que guían el comportamiento humano, así como las decisiones relacionadas con lo correcto y lo incorrecto, lo bueno y lo malo. Según Velásquez (2018), la ética implica reflexionar sobre las normas y valores que orientan nuestras acciones y evaluar críticamente las consecuencias éticas de nuestras elecciones. Además, la ética busca proporcionar

un marco conceptual para entender y abordar dilemas morales en diversos contextos individuales y sociales.

La inteligencia artificial (IA) refiere la facultad de las máquinas y sistemas computacionales para realizar tareas que normalmente requieren inteligencia humana. Según Russell y Norvig (2016), la IA comprende una amplia gama de técnicas y enfoques, incluidos algunos algoritmos de aprendizaje automático, implementa redes neuronales artificiales y procesamiento del lenguaje natural, que permiten a las máquinas realizar funciones cognitivas como el razonamiento, la percepción y el aprendizaje; en consecuencia, la IA busca simular la inteligencia humana para resolver problemas, tomar decisiones y mejorar la eficiencia en múltiples campos, desde la medicina hasta la ingeniería.

La intersección entre ética y la inteligencia artificial (IA) se presenta como una línea de investigación y análisis fascinante y complejo que ha capturado la atención de filósofos, científicos tecnólogos por igual. En la actualidad, también ha cobrado gran relevancia en el ámbito empresarial, se aclara que este análisis no se plantea solo desde una reflexión netamente filosófica, sino que también se oriente desde un planteamiento en términos de reputación corporativa. En el corazón de esta interacción se encuentran cuestiones fundamentales sobre el papel de la moralidad en la creación, implementación y uso de sistemas inteligentes aplicados al mundo de los negocios. Este estudio explora algunas de las principales preocupaciones éticas relacionadas con la IA y su impacto en la sociedad y la repercusión que se presenta cada vez en mayores ámbitos de la cotidianidad de la vida empresarial.

Una de las cuestiones éticas más debatidas en relación con la IA es la responsabilidad moral. Como señala Floridi (2010), los sistemas de IA pueden actuar de manera autónoma y tomar decisiones que tienen un considerable suceso en las vidas de las personas. Sin embargo, la asignación de responsabilidad en caso de que estas decisiones resulten perjudiciales puede ser complicada. ¿Deberíamos responsabilizar a los desarrolladores de IA, a los propietarios o usuarios de los sistemas, o a

los propios sistemas? Esta pregunta plantea dilemas éticos profundos sobre la naturaleza de la responsabilidad moral en un mundo cada vez más dependiente de la tecnociencia.

Otro aspecto crucial es el sesgo algorítmico en los sistemas de IA. Como apunta Russell y Norvig (2016), los algoritmos de IA pueden influenciar sesgos inherentes a los conjuntos de datos con los que son entrenados. Esto puede conducir a decisiones discriminatorias o injustas, lo que plantea serias preocupaciones éticas sobre la equidad y la justicia. Abordar este sesgo algorítmico no solo implica cuestiones técnicas, sino también reflexiones profundas sobre cómo definimos y mitigamos la discriminación en un contexto algorítmico. Además, la privacidad y la seguridad son preocupaciones éticas fundamentales en el desarrollo y la implementación de la IA. Generalmente los sistemas de IA requieren acceso a ciertas cantidades de datos personales para su funcionamiento, lo que plantea preguntas sobre el consentimiento informado, la protección de la privacidad y el potencial abuso de datos. Como señalan Müller y Bostrom (2016), es crucial establecer un equilibrio entre la utilidad de la IA y la protección de los derechos individuales de privacidad y seguridad para garantizar un desarrollo tecnológico ético y responsable.

El planteamiento de la inteligencia artificial implica la incorporación de principios éticos desde el inicio, incluyendo la fase de diseño y desarrollo de sistemas de IA. Este enfoque proactivo puede ayudar a prevenir posibles consecuencias no deseadas y garantizar que la IA se utilice para promover el bienestar humano. Según Floridi y Chiriatti (2020), las empresas deben adoptar un enfoque centrado en el ser humano al diseñar sistemas de IA, teniendo en cuenta el impacto que estos tendrán en la sociedad en general.

A partir de lo expuesto, es necesario plantear algunas consideraciones éticas en el despliegue de la IA. Junto con su desarrollo ético, también es crucial establecer cómo se implementan y utilizan estos sistemas en el mundo real. Las decisiones sobre cómo se despliega la IA pueden tener importantes implicaciones éticas y sociales, por ejemplo, el uso de algo-

ritmos de IA en la toma de decisiones judiciales ha generado preocupaciones sobre la equidad y el sesgo algorítmico (Diakopoulos, 2016). Las empresas tienen la responsabilidad de ser transparentes sobre cómo se utilizan los algoritmos de IA y garantizar que se implementen de manera justa y equitativa, puesto que:

En el ámbito económico se produce un proceso de algoritmización en las tomas de decisiones. Existe un reconocido elenco de consejeros y directivos algorítmicos en instituciones, organizaciones y empresas, se coloca a un algoritmo frente a determinadas secciones o dentro de ellas. Xerox, Google, Unilever, L’Oreal o Amazon tienen un algoritmo dentro o al frente de la dirección de Recursos Humanos y también se multiplican en los medios de comunicación. (Cortina Orts, 2019, p. 380)

Se presenta una diferencia abismal entre hacer uso de sistemas inteligentes (sean máquinas, algoritmos, robots) a la hora de tomar decisiones y delegar en esos sistemas decisiones significativas para la vida de las personas y de la naturaleza. Esto plantea una interrogante, ya que ¿una ética de la inteligencia artificial es la que deben practicar los sistemas inteligentes desde sus propios valores, o es la que los seres humanos deberíamos adoptar para servirnos de los sistemas inteligentes? (Cortina Orts, 2019, p. 381). La respuesta a esta interrogante deberá consolidarse como guía de la reflexión en el futuro para las empresas.

Tomemos en consideración que nuestro mundo se presenta en un estado de digitalización y auge de las inteligencias artificiales. El paso es irreversible y, por lo tanto, no cabe preguntar si debemos darlo, sino cómo hacerlo para conseguir el mayor bien posible (Cortina Orts, 2019, p. 386). Es decir, al plantear diversos aspectos “gracias al *machine learning*, encontramos, en las grandes corporaciones, máquinas o robots inteligentes, que interactúan de una forma natural y precisa con las personas” (Goñi Sein, 2019, p. 59), esto genera algunas planteamientos interesantes, por ejemplo el hecho de que las empresas recurren cada vez más a la inteligencia artificial en la gestión de personas. Esto sucede, particularmente, en los procesos de selección que las empresas llevan a cabo sin ayuda humana

o con una intervención mínima a base de sistemas automatizados. Son los softwares especializados los que realizan el primer filtro de selección en el reclutamiento de personal parametrizando una serie de palabras clave relacionadas con la empresa o el puesto vacante y en muchos casos la selección completa (Goñi Sein, 2019, p. 320).

Pueden mencionarse ciertas situaciones como “el caso de Amazon que proporciona a los trabajadores un dispositivo digital, mediante el cual el trabajador es guiado y controlado en el tiempo que invierte en transportar paquetes de un sitio a otro, de manera que si emplea más de lo establecido se le envía un mensaje de forma automática” (Goñi Sein, 2019, p. 388). Así mismo, hay que tomar en cuenta que:

[En] la interacción de los humanos y los robots y/o máquinas inteligentes que empiecen a mostrar una conducta autónoma, se deben instaurar o crear códigos éticos que rijan o limiten las acciones de dichas máquinas. De igual manera, se deben implantar dichos códigos en los procesos de producción de otras máquinas inteligentes (Tello Castrillón y Rodríguez Córdoba, 2016, p. 15). Las tecnologías GPT digitales de algunos sistemas de información como chatbots interactivos están reemplazando “funciones administrativas, de publicidad y mercadeo y de generación de textos e imagen en las empresas. Económicamente, es rentable, genera ganancias y reduce problemas y, por lo tanto, no se desechará. (Gordon Graell, 2023, p. 144)

Al analizar esta realidad es factible apuntar que la IA “es capaz de hacer que la humanización de la vida progrese; pero también está siempre presente el peligro de que —en línea con la acertada intuición que habla de un *responsibility gap* (Matthias, 2004)— no todo redunde en beneficio de las personas y de la sociedad en su conjunto” (Fernandez, 2021, p. 117). El Instituto Tecnológico de Massachusetts desarrolló una plataforma llamada *Moral Machine* para analizar los problemas éticos que implica el uso de estas tecnologías autónomas. Se orienta según los veintitrés principios de Asilomar de inteligencia artificial, cuyo tema principal es que “el ser humano debe decidir cuándo y cómo delegar las decisiones cuando desarrolla y utiliza la IA. No es fácil responsabilizar a una máquina por los actos y

decisiones que genere ya que finalmente esta se encuentra desarrollada por los seres humanos” (Lenin y Satama, 2023, p. 104).

Desde estos planteamientos, sin duda es necesario concluir que las personas tendrían que “despejar las incógnitas de la ecuación atendiendo criterios éticos. [Es] difícil que una máquina [calibre] este tipo de decisiones, por eso es tan importante que la inteligencia artificial esté dotada de una ética o código de valores que condicione, a nuestra imagen y semejanza, la esencia de sus actos” (Villalba, 2020, p. 683). Ante estas consideraciones es factible argumentar que la ética de un individuo es importante porque los factores que la integran son prioritarios en tanto persiguen el buen uso de la IA. “[La] creación de marcos regulatorios-jurídicos son indispensables, para preservar las buenas prácticas y acciones empresariales” (Sequeira Cascante y Sánchez Sánchez, 2023, p. 152).

Entonces la sugerencia de someter a debate público las repercusiones de la IA en nuestras vidas nace de la necesidad de imaginar alternativas y tiene que ver con la asunción de la responsabilidad (Terrones Rodríguez, 2018, p. 168). Si hablamos de concientización debemos entender que “la ética de la IA abarca el diseño y el uso de sistemas autónomos, los prejuicios de las máquinas, la privacidad y la vigilancia, la gobernanza, el status de las máquinas inteligentes, la automatización y el desempleo, e incluso la colonización espacial” (Degli-Esposti, 2023, p. 35).

A partir de los argumentos presentados se deduce que las empresas también tienen la obligación y responsabilidad de garantizar que la IA se utilice de manera coherente y en delimitación ética. Esto implica no solo cumplir con las regulaciones y estándares éticos existentes, sino también anticipar y abordar activamente posibles consecuencias no deseadas de sus acciones. Por ejemplo, las empresas que desarrollan sistemas de IA para el reconocimiento facial deben considerar cómo se pueden utilizar estos sistemas de manera inapropiada, como la vigilancia masiva o la discriminación racial (Jobin *et al.*, 2019).

Al analizar la realidad derivada del continuo desarrollo de la IA se comprende que, desde siempre, ha sido importante plantearse interrogantes sobre las adecuadas prácticas empresariales. El referente ideal sería un contrato social que rija el comportamiento humano sobre una serie de acuerdos y conductas comunitarias que, a su vez, permita la interacción con pares permitiendo la consolidación del sector empresarial y productivo. Los cambios tecnológicos que conlleva la IA deben ser vistos como un potencial para que la empresa promueva acciones en beneficio de sus *stakeholders* internos (empleados y accionistas) y externos como los clientes, proveedores y sociedad en general.

Los especialistas y estudiosos de las ramas de las ciencias sociales y exactas discuten sobre en qué era se encuentra la humanidad, respecto del uso e implementación de la tecnología. Se acuerda en la clasificación de revolución industrial 4.0, conocida como la revolución de la digitalización, donde el plano físico se mezcla con el plano digital. En este contexto, el mundo físico de la humanidad se conecta con el mundo virtual en el metaverso, donde la identidad humana se convierte en avatar (Llanes y Lorenzo, 2021).

Surgen algunas estrategias para promover el uso ético de la inteligencia artificial, entre las que se pueden mencionar establecer políticas internas claras que fomenten el desarrollo y uso ético de la IA. Esto incluye la creación de comités éticos dedicados para revisar y supervisar proyectos de IA (Scherer, 2016). Además, las empresas pueden invertir en la formación y capacitación de su personal en ética y responsabilidad digital, para garantizar que comprendan las implicaciones éticas de su trabajo en el campo de la IA. En el ámbito económico y social —que tiene una importante repercusión en la consolidación del uso responsable de las IA— las empresas pueden adoptar diversas prácticas para promover un uso ético de la IA, tales como:

- Evaluación de riesgos éticos

Realizar evaluaciones de impacto ético para identificar y mitigar posibles riesgos asociados con los sistemas de IA (López, 2020). Esta

concientización parte de una reflexión ética y filosófica que forma parte de un marco lógico institucional.

- Diseño centrado en el usuario

Involucrar a las partes interesadas, incluidos los grupos marginados, en el diseño y la prueba de sistemas de IA para garantizar que se satisfagan sus necesidades y preocupaciones (Sánchez, 2021). Sin que esto signifique un choque con los objetivos puntuales de la empresa, el desarrollo enmarcado en una visión humanista.

- Auditoría algorítmica

Implementar procesos de auditoría para examinar sistemáticamente los sistemas de IA en busca de sesgos y otros problemas éticos (Fernández, 2019). Esta auditoría se presenta como una oportunidad de plantear estrategias clave para los planteamientos éticos.

Así pues, será relevante en esta época de ámbito empresarial, considerar el tema ético en términos de reputación corporativa. Es obligación de las empresas ser éticas y a su vez esto les genera beneficios porque hoy, más que nunca, los clientes reconocen y premian a las organizaciones que tienen una gestión administrativa y económica transparente. La inteligencia artificial (IA) entonces ha traído consigo una serie de implicaciones éticas que deben ser consideradas de manera crítica. Entre estas implicaciones, se destacan la discriminación algorítmica, la privacidad y la transparencia.

La discriminación algorítmica, como señala García (2020), se refiere a la tendencia de los algoritmos de IA a motivar y amplificar sesgos inherentes en los datos utilizados para su entrenamiento. Esto puede conducir a decisiones discriminatorias en diversos contextos, como la contratación, el otorgamiento de créditos y el sistema judicial. Por ejemplo, un algoritmo de selección de personal que se entrena con datos históricos puede reproducir sesgos de género o raza, lo que resulta en la exclusión injusta de ciertos grupos.

En cuanto a la privacidad, Pérez (2021) destaca que la recopilación masiva de datos necesaria para entrenar los modelos de IA plantea preocupaciones sobre la protección de la información personal. Los sistemas de IA pueden acceder, analizar y utilizar datos sensibles de individuos sin su consentimiento explícito, lo que afecta la privacidad y la autonomía de las personas. Por ejemplo, los asistentes virtuales y los sistemas de reconocimiento facial pueden recopilar y almacenar información biométrica sin el conocimiento del usuario, lo que aumenta el riesgo de abusos y violaciones de la privacidad.

Además, la falta de transparencia y claridad en los algoritmos de IA plantea desafíos éticos significativos. Como señala Rodríguez (2018), muchos sistemas de IA operan como *cajas negras*, lo que significa que es difícil entender cómo llegan a sus decisiones. Esta opacidad dificulta la supervisión y evaluación de los sistemas de IA, lo que puede generar desconfianza y socavar la legitimidad de estas tecnologías.

Para abordar estos desafíos, diversas organizaciones han propuesto principios éticos para guiar el desarrollo y la implementación de la IA:

- Equidad y justicia

Garantizar que los sistemas de IA no perpetúen ni amplifiquen sesgos y discriminaciones existentes (Smith, 2019). Este principio plantea garantizar que los sistemas de IA no amplifiquen sesgos existentes. Desde una perspectiva filosófica, este principio se relaciona con los planteamientos de la justicia social de John Rawls para quien las decisiones justas se toman cuando quienes las adoptan lo hacen sin saber cuál será su posición en la sociedad, esto aseguraría que las reglas beneficien equitativamente a todos.

En el contexto de la IA, este *velo de ignorancia* podría interpretarse como un enfoque para diseñar algoritmos que consideren diversas perspectivas sociales sin favorecer a ciertos grupos. Pero, este ideal enfrenta obstáculos prácticos, ya que los datos utilizados para entrenar sistemas

de IA suelen reflejar las desigualdades y los prejuicios históricos presentes en las sociedades. Si un modelo de IA se alimenta con datos históricos, puede replicar patrones de discriminación de género o raza al evaluar candidatos para un empleo. Esto puede llevar a considerar una perspectiva de justicia distributiva en términos aristotélicos (distribución justa de recursos) y justicia correctiva (rectificación de desequilibrios). En la IA, la justicia distributiva implicaría garantizar el acceso equitativo a los beneficios de la tecnología, mientras que la correctiva abordaría activamente los sesgos que perjudican a grupos marginados.

Estos planteamientos, permiten concluir que la equidad en la IA no presenta sólo una cuestión ética, sino un acto de reparación moral, puesto que las personas deben ser tratadas como fines en sí mismas y no como medios para un propósito.

- Transparencia

Proporcionar explicaciones claras sobre cómo se toman las decisiones algorítmicas y qué datos se utilizan (González, 2022). Este principio busca proporcionar explicaciones claras sobre las tomas de decisiones algorítmicas y qué datos se utilizan, eso se puede vincular con la fundamentación ética kantiana, fundamentada en la autonomía y las capacidades de decisión de las personas.

En el contexto de la IA implica a respetar el derecho de las personas a entender cómo se utilizan sus datos y cómo los sistemas afectan sus vidas, esto constituye un gran dilema filosófico y práctico debido a la complejidad de los modelos algorítmicos, como las redes neuronales profundas, dificultan el ofrecer explicaciones comprensibles para usuarios no especializados. Esto, en definitiva, presenta una tensión entre la transparencia como principio ético y las limitaciones epistemológicas inherentes a estas tecnologías, entonces tenemos una dicotomía puesto que —por ejemplo desde la fenomenología de Heidegger— la transparencia podría interpretarse como una forma de *desocultar* la tecnología, exponiendo sus fundamentos escondidos y mostrando cómo opera en relación con el

mundo humano. Desde una perspectiva utilitarista, la transparencia debe equilibrarse con otros valores como la seguridad y la eficiencia; revelar demasiada información sobre el funcionamiento interno de un sistema podría facilitar su manipulación o comprometer su funcionamiento óptimo. Este conflicto destaca la necesidad de encontrar un punto medio entre la claridad informativa y la protección de la funcionalidad tecnológica.

- Responsabilidad

Asignar responsabilidades claras para el desarrollo, implementación y supervisión de sistemas de IA (Martínez, 2023). Este principio se refiere a establecer responsabilidades claras para el desarrollo, implementación y supervisión de sistemas de IA. Se vincula con la ética de la virtud de Aristóteles, que enfatiza la importancia de las virtudes personales y las acciones responsables para construir una sociedad justa, esto significa que los desarrolladores, reguladores y usuarios deben actuar con prudencia, integridad y compromiso ético en todos los ámbitos de un ciclo de vida de los sistemas.

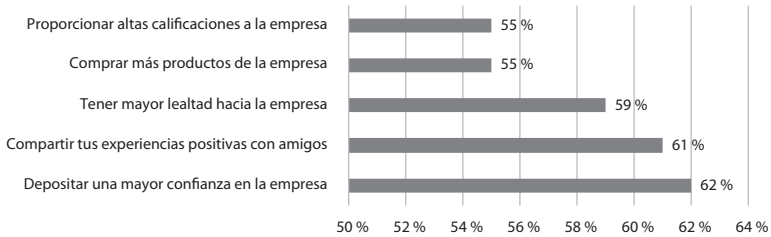
Un desafío filosófico fundamental es la dilución de la agencia moral. Cuando múltiples actores participan en el diseño y operación de un sistema, surge la pregunta: ¿quién es el responsable cuando algo sale mal? Es posible tomar como ejemplo, el hecho de que un automóvil autónomo cause un accidente, ¿es responsable el fabricante del vehículo, el programador del algoritmo, o la entidad que suministro los datos? Esto se compara con el problema de la ‘banalidad del mal’ de Hanna Arendt, que surge cuando las personas delegan su responsabilidad ética a sistemas burocráticos o tecnológicos, diluyendo así su implicación moral. Estos argumentos se relacionan con una perspectiva deontológica, así pues Kant enfatizaría que la asignación de responsabilidades debe basarse en el principio de deber: cada actor involucrado tiene el deber moral de prever las consecuencias de sus acciones y garantizar que sus decisiones respeten la dignidad humana, esto contrasta con un enfoque utilitarista, que podría justificar la asignación de responsabilidades basándose en

la maximización del bienestar colectivo, incluso si ello implica ciertos intereses individuales.

Si bien cada uno de estos principios aborda una dimensión específica de los desafíos éticos de la Inteligencia Artificial, su interdependencia es innegable. La equidad y la justicia no pueden garantizarse sin transparencia, ya que las decisiones injustas suelen ocultarse detrás de procesos opacos. Así mismo, la transparencia es insuficiente sin responsabilidad, ya que comprender un sistema no garantiza que los actores relevantes asuman su rol en corregir sus fallos. Entonces, una IA transparente podría revelar un sesgo en sus decisiones, pero sin un marco claro de responsabilidad, nadie podría ser obligado a corregirlo. Del mismo modo, una IA diseñada para ser justa podría fallar en su implementación si los responsables no actúan de manera ética.

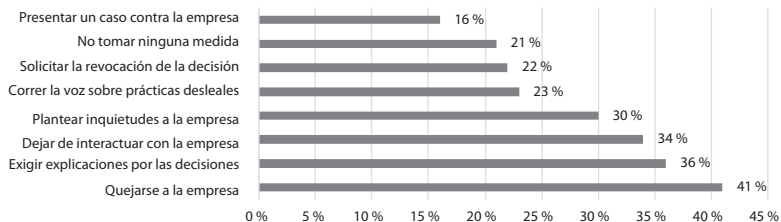
A partir de esta reflexión filosófica y ética, es relevante presentar y analizar cifras relacionadas con las implicaciones éticas de las acciones de las empresas frente al auge de la inteligencia artificial. Desde una orientación ética y social es fundamental establecer directrices de acción como respuesta a los resultados evaluados. En primera instancia según el estudio de Capgemini Research Institute “Por qué abordar cuestiones éticas en IA beneficiará a las organizaciones” la *inteligencia artificial ética* es la piedra angular sobre la que se construyen la confianza y la lealtad del cliente (Thieullent, 2023). Este criterio indudablemente es prioritario en el contexto actual del mundo empresarial cada vez más relacionado con la responsabilidad social.

En este estudio encuestaron a 1580 ejecutivos de 510 organizaciones y a más de 4400 consumidores a nivel mundial con el objetivo de conocer qué están haciendo las organizaciones para abordar las implicaciones éticas de la inteligencia artificial. Lo anterior permite tener una visión general para el análisis. Las principales conclusiones del estudio se resumen en las siguientes gráficas:

Figura 1*Percepción del cliente del uso la inteligencia artificial con ética***¿Qué harías si percibes que la interacción facilitada por la inteligencia artificial es ética?**

Fuente: Instituto de Investigación Capgemini, encuesta de consumidores sobre ética en inteligencia artificial. N = 4447 consumidores.

Este primer insumo nos demuestra que la ética en la inteligencia artificial genera reputación corporativa para las empresas. Los consumidores encuestados, el 62 % manifestó la confianza en una empresa cuyas interacciones con IA son percibidas como éticas, mientras que el 61 % señala compartiría experiencias positivas con amigos y familiares. Esto nos demuestra que los consumidores exigen un accionar ético de los procesos de IA de las empresas.

Figura 2*Comportamiento del cliente frente al actuar ético de la empresa ante el uso de IA***¿Qué es probable que haga en caso de que perciba que la empresa no actúa éticamente con sus sistemas de inteligencia artificial?**

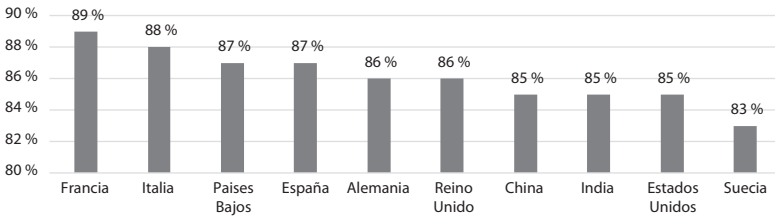
Fuente: Instituto de Investigación Capgemini, encuesta de consumidores sobre ética en inteligencia artificial. N = 4447 consumidores.

Este gráfico evidencia que un 34 % de los consumidores dejan de interactuar con las empresas si perciben problemas éticos en los sistemas de inteligencia artificial.

Figura 3

Problemas resultantes del uso de la IA

En los últimos dos o tres años, ¿le han llamado la atención los siguientes problemas resultantes del uso y la implementación de sistemas de IA (porcentaje de ejecutivos por país)



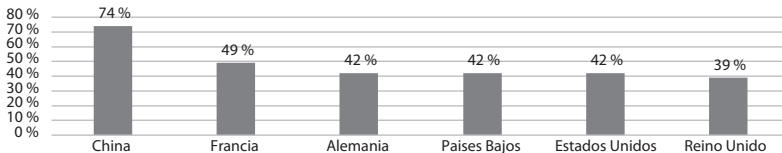
Fuente: Instituto de Investigación Capgemini, encuesta a ejecutivos sobre ética en Inteligencia Artificial. N = 1580 ejecutivos, 510 organizaciones.

Se evidencia que los ejecutivos de nueve de cada diez organizaciones consideran que los problemas éticos han resultado del uso de sistemas de inteligencia artificial desde los últimos 2 o 3 años. Los problemas que más se han identificado son: la dependencia excesiva de las decisiones dirigidas por máquinas en banca y seguros y la recopilación de datos personales de pacientes sin consentimiento en el cuidado de la salud.

Figura 4

Consumidores que han enfrentado problemas éticos en la IA

Proporción de consumidores que han experimentado al menos dos tipos de uso de IA que resultaron en problemas éticos en los últimos dos o tres años por país



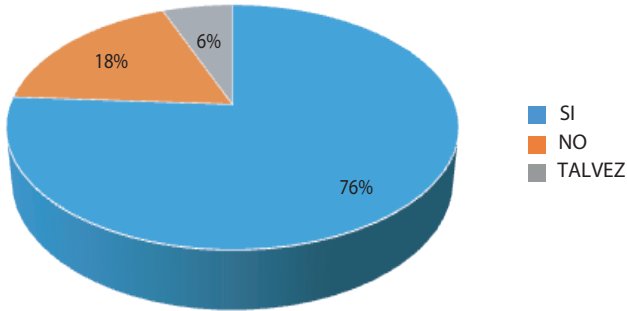
Fuente: Instituto de Investigación Capgemini, encuesta de consumidores sobre ética en inteligencia artificial. N = 4447 consumidores.

En esta figura podemos evidenciar que aproximadamente la mitad de los encuestados manifiestan haber experimentado al menos dos tipos de usos de la IA que generaron problemas éticos en los últimos dos o tres años siendo los consumidores chinos quienes han experimentado mayor porcentaje de tipos de uso de inteligencia artificial que desencadenaron en problemas éticos.

Figura 5

Desarrollo de regulaciones referentes al uso ético de la IA

¿Considera usted que debe haber nuevas regulaciones para el uso ético de la inteligencia artificial en las organizaciones?



Fuente: Instituto de Investigación Capgemini, encuesta de consumidores sobre ética en Inteligencia Artificial. N = 4447 consumidores.

Según estos resultados se evidencia que para el 76 % de los consumidores es importante que existan nuevas regulaciones sobre el uso de la IA.

Las implicaciones éticas de la IA en las organizaciones son diversas y complejas. Para abordar estos desafíos, es fundamental que las organizaciones adopten enfoques éticos y responsables en el desarrollo, implementación y regulación de la IA. Como evidencia el estudio realizado por Capgemini Research Institute, los clientes premian a las empresas cuyo accionar con los sistemas de IA es ético, además es importante mencionar que las organizaciones a través de sus directivos consideran importante y

necesario resguardar que los sistemas de inteligencia artificial sean éticos y transparentes, así lo afirmaron el 51 % de los ejecutivos consultados.

Para que la IA sea exitosa debe haber una prevención anticipada de riesgos lo que implica usar una perspectiva holística en la organización, es decir que la alta gerencia debe estar convencida del accionar ético en el uso de los sistemas de inteligencia artificial y transmitir esta convicción y formación a todos los colaboradores involucrados en implementar la IA en los diferentes ámbitos de las organizaciones.

Materiales y método

La presente investigación se desarrolló bajo la metodología hermenéutica que, como enfoque interpretativo en la investigación filosófica, se centra en la comprensión profunda de textos y discursos para desentrañar sus significados subyacentes en el contexto histórico y cultural. La metodología hermenéutica se fundamenta en los principios de comprensión reflexiva e interpretación dialógica y sigue los siguientes pasos.

- Selección del texto o tema

La investigación hermenéutica comienza con la selección de un texto filosófico o un tema de estudio relevante o, según Gadamer (2013), textos clásicos. En el caso de esta investigación el tema es relevante puesto que la ética como disciplina filosófica ha cobrado gran significado en la actualidad de las empresas.

- Comprensión preliminar

Previo al análisis detallado, es esencial adquirir una comprensión preliminar del texto o tema en cuestión. En este caso se ha realizado un análisis filosófico de la interrelación entre el accionar ético de las empresas frente a la inteligencia artificial. Como Heidegger (2002, p. 45) sostiene: “La comprensión previa no es sólo un acceso previo al texto, sino el horizonte a partir del cual se le puede leer”.

- Análisis hermenéutico:

El análisis hermenéutico implica una inmersión profunda en el texto o tema de estudio, explorando sus múltiples capas de significado. Según Ricoeur (1970, p. 73), “La interpretación consiste en mediar entre lo que el texto dice y lo que quiere decir”.

- Diálogo interpretativo:

La hermenéutica enfatiza el diálogo interpretativo entre el investigador y el tema de investigación. Como Gadamer (2013, p. 311) sugiere: “El prejuicio debe ser comprendido y no eliminado, porque es el medio a través del cual se lleva a cabo la comprensión”.

- Contextualización hermenéutica:

El proceso interpretativo se contextualiza dentro de la historia y la cultura en la que el tema emerge. En palabras de Dilthey (2002, p. 330): “La hermenéutica es el arte de entender la vida humana a través de la comprensión de los productos de la mente humana”.

- Reflexión filosófica:

La investigación hermenéutica culmina en una reflexión filosófica sobre los significados desvelados. Según Gadamer (2013, p. 282): “La aplicación de una interpretación es siempre la participación en un horizonte que permanece en disputa”.

- Presentación de resultados:

Finalmente, se presentan los hallazgos de la investigación de manera clara y coherente, reconociendo la provisionalidad de las conclusiones hermenéuticas. Ricoeur (1970) señala: “La tarea del texto y de su interpretación no es tanto ofrecer respuestas definitivas como abrir nuevas posibilidades de significado”.

Esta metodología hermenéutica proporciona un marco riguroso para la investigación filosófica, permitiendo una comprensión enriquecedora de textos y temas filosóficos significativos en su contexto histórico y cultural.

Resultados y discusión

Interpretación de los datos

A partir de los datos expuestos, es evidente que, en el panorama actual, la indudable evolución de la tecnología, especialmente en el campo de la IA, ha planteado nuevos desafíos éticos para las empresas. Como se ha expuesto a lo largo de este trabajo la integración de sistemas de IA en diversas áreas empresariales, como la toma de decisiones, el análisis de datos y el servicio al cliente, ha generado interrogantes sobre cómo garantizar un actuar ético en este contexto. Para abordar esta cuestión, resulta pertinente recurrir a la perspectiva ética de la filósofa española Adela Cortina, cuyo enfoque en la ética del discurso proporciona un marco sólido para analizar las implicaciones éticas de la IA en el ámbito empresarial.

Cortina (2017) sostiene que la ética del discurso se basa en el principio de la racionalidad comunicativa y la intersubjetividad, es decir, en la capacidad de los seres humanos para comunicarse y entenderse mutuamente a través del diálogo racional. Desde esta perspectiva, las empresas tienen la responsabilidad de garantizar que sus prácticas y decisiones estén sujetas a un proceso de deliberación ética que involucre a todas las partes afectadas. En el contexto de la IA, esto implica reflexionar sobre cómo se utilizan los algoritmos, los datos y si estos procesos son coherentes con los principios éticos fundamentales.

Así mismo uno de los principales desafíos éticos que enfrentan las empresas en relación con la IA es el sesgo algorítmico. Los algoritmos de IA, al basarse en datos históricos, pueden perpetuar y amplificar sesgos injustos presentes en la sociedad, como el racismo o el sexismo (Cortina,

2019). Esto plantea interrogantes sobre la responsabilidad de las empresas en la identificación y mitigación de estos sesgos. Desde la ética del discurso, abordar este problema implica promover la transparencia y la rendición de cuentas en el diseño y la implementación de sistemas de IA, así como garantizar la participación de diversas perspectivas en este proceso.

Otro aspecto ético relevante propuesto por Cortina es la seguridad de los datos, puntualmente la privacidad. Las empresas al utilizar IA tienen acceso grandes cantidades de información personal de los usuarios, lo que plantea preocupaciones sobre el uso indebido de estos datos y la vulneración de la privacidad (Cortina, 2018). Para abordar este desafío ético, las empresas deben adoptar políticas y prácticas que protejan la privacidad de los individuos y aseguren la seguridad de sus datos, respetando así su autonomía y dignidad.

En cuanto a la responsabilidad moral, se plantea el debate sobre si las empresas pueden asumir las que existen consecuencias de acciones y decisiones en el marco de la IA. Esto implica garantizar que los sistemas de IA sean utilizados de manera ética y responsable, y que se establezcan mecanismos adecuados de rendición de cuentas en caso de que surjan problemas o daños (Cortina, 2016). Desde la ética del discurso, la responsabilidad moral de las empresas radica en su capacidad para actuar de acuerdo con los principios éticos fundamentales y para responder ante las acciones en el contexto de la IA.

Desde la perspectiva de Adela Cortina y su ética del discurso, las empresas tienen la responsabilidad de garantizar que sus prácticas y decisiones en relación con la IA sean coherentes con los principios éticos fundamentales, como la justicia, la equidad y el respeto a la dignidad humana.

Conclusiones

El actuar ético de las empresas frente a la inteligencia artificial es crucial para mitigar los riesgos y maximizar los beneficios de esta tecnología. Al adoptar principios éticos, implementar prácticas empresariales

sólidas y fomentar la transparencia y la responsabilidad en las tomas de decisiones, las empresas desempeñan un protagonismo fundamental en la promoción de una IA ética y responsable.

En conclusión, el desarrollo y despliegue de la inteligencia artificial plantea una serie de desafíos éticos para las empresas. Sin embargo, con un enfoque proactivo en el desarrollo ético, consideraciones éticas en el despliegue de la IA y una responsabilidad ética continua, las empresas deben asegurar que la IA se utilice de manera responsable para promover el bienestar humano.

Si bien la inteligencia artificial puede resultar una herramienta beneficiosa para ahorrar costos y optimizar procesos en la gestión empresarial y en general para todo ámbito del quehacer humano, también trae consigo riesgos que si no se manejan adecuadamente podrían atentar al bienestar humano. Es evidente que ninguna máquina podrá reemplazar a la inteligencia humana, por lo tanto, los seres humanos debemos capacitarnos para que esta poderosa herramienta sea utilizada para el beneficio de la colectividad.

Insights y recomendaciones para promover una ética empresarial sólida y sostenible en la era de la inteligencia artificial

Luego del análisis de los datos proporcionados en el artículo se proponen las siguientes recomendaciones para promover una ética empresarial sólida y sostenible:

- La ética empresarial ha ampliado su campo de acción y es la raíz del desarrollo sostenible y la base para la sostenibilidad en el mundo de los negocios. Las empresas deben crear un comité de ética empresarial que esté liderado por la alta gerencia que tenga como responsabilidad garantizar que todas las acciones que la empresa emprenda estén enmarcadas en un accionar ético. Las empresas que cuenten con un programa integral y una estructura sostenible,

alcanzan operaciones eficientes, logran un excelente desempeño financiero, crean un gran ambiente de trabajo y clientes fidelizados.

- Establecer un equilibrio entre la utilidad de la IA y la protección de los derechos individuales de privacidad y seguridad para garantizar un desarrollo tecnológico ético y responsable.
- Mantener políticas claras que incluyan el consentimiento informado, la protección de la privacidad y el potencial abuso de datos.
- La IA debe estar siempre supervisada por humanos garantizando por un lado la protección de los derechos y también la dignidad humana basándose en principios fundamentales como la transparencia y la equidad.
- La IA debe ser confiable y transparente, las empresas deben establecer claramente quién entrena sus sistemas de IA, que datos se utilizan en este entrenamiento y que se incluyó en las recomendaciones de sus algoritmos. Para cumplir con este criterio las empresas deben considerar para sus algoritmos datos diversos y representativos, técnicas mitigación de sesgos, equipos de desarrollo diversos
- Toda empresa debe reflexionar críticamente y adoptar una regulación adecuada para garantizar que los beneficios de la IA se maximicen y los riesgos como la privacidad de los datos, el sesgo algorítmico y el impacto en el empleo se mitiguen.
- Las organizaciones deben manejar una IA responsable para generar confianza en las soluciones que la misma provee a las empresas propiciando un impacto social positivo y las medidas necesarias para alinear las prácticas de IA con los valores de las partes interesadas, las normas legales y los principios éticos.
- Realizar programas de concientización a los empleados, partes interesadas y los tomadores de decisiones sobre las prácticas de IA responsables.
- El análisis filosófico de los principios de equidad y justicia, transparencia y responsabilidad subraya la importancia de un enfoque ético integral para fomentar el desarrollo de la inteligencia artificial aplicada al ámbito empresarial. El análisis de los problemas filosóficos en esta área nos recuerda que la tecnología no es neutra:

refleja los valores, prejuicios y decisiones de quienes la diseñan y utilizan. Por ello garantizar que la IA contribuya al bienestar humano requiere una reflexión ética constante que trascienda las consideraciones éticas.

- Para avanzar hacia este objetivo, es esencial combinar enfoques normativos (que establezcan principios generales) con enfoques pragmáticos (que detallen cómo aplicarlos en contextos específicos). Solo así podremos desarrollar sistemas de IA que no solo sean técnicamente avanzados, sino también éticamente sólidos, reflejando lo mejor de nuestras aspiraciones humanas.

Referencias bibliográficas

- Cortina, A. (2019). Ética de la Inteligencia Artificial. *Anales de La Real Academia de Ciencias Morales y Políticas*, 379-394.
- Cortina, A. (2019). *Ética sin moral*. Tecnos.
- Cortina, A. (2018). *Aporofobia, el rechazo al pobre: un desafío para la democracia*. Paidós.
- Cortina, A. (2017). *Ética aplicada y democracia radical*. Tecnos.
- Cortina, A. (2016). *Ética de la empresa*. Trotta.
- Degli-Esposti, S. (2023). *La ética de la inteligencia artificial*. Editorial Consejo Superior de Investigaciones Científicas. <https://bit.ly/42nRWzg>
- Diakopoulos, N. (2016). *Algorithmic accountability: A primer*. Columbia Journalism School.
- Fernández, A. (2019). Auditoría algorítmica en inteligencia artificial. *Revista de Ética Empresarial*, 15(2), 45-58.
- Fernández, J. L. (2021). Hacia un Humanismo Digital *Disputatio. Philosophical Research Bulletin*, 10(17), 107-130.
- Floridi, L. (2010). *The Cambridge Handbook of Information and Computer Ethics*. Cambridge University Press.
- Floridi, L. y Chiriatti, M. (2020). AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 30(4), 1-36.
- Gadamer, H. G. (2013). *Verdad y Método: Fundamentos de una hermenéutica filosófica*. Sígueme.

- García, M. (2020). La discriminación algorítmica y sus implicaciones éticas. *Harvard Business Review*, 30(4), 112-125.
- García, M. y Sánchez, P. (2020). *Inteligencia Artificial y Ética Empresarial*. Harvard Business Review.
- González, R. (2022). *Decisiones algorítmicas y justicia social*. Editorial Ariel.
- González, R. (2022). *Transparencia y explicabilidad en sistemas de inteligencia artificial*. Editorial Tecnos.
- Goñi Sein, J. L. (2019). Innovaciones tecnológicas, inteligencia artificial y derechos humanos en el trabajo. *Documentación Laboral*, 57-72, II (117) <https://bit.ly/4mJXgUS>
- Jobin, A., Inenca, M. y Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.
- Lenin, F., y Satama, V. (2023). Inteligencia Artificial: El reto contemporáneo de la gestión empresarial. *Revista ComHumanitas*, 14, 94-111.
- López, J. (2020). Evaluación de riesgos éticos en la implementación de inteligencia artificial. *Revista de Tecnología y Ética*, 25(3), 78-91.
- Martínez, E. (2023). Responsabilidad corporativa en la era de la inteligencia artificial. *Revista de Ética Empresarial*, 18(1), 32-45.
- Müller, V. C. y Bostrom, N. (2016). Future progress in artificial intelligence: A survey of expert opinion. En K. R. Vincent (Ed.), *Ética de la inteligencia artificial* (pp. 1-19). Springer.
- Pérez, L. (2021). Ética y Responsabilidad en la Inteligencia Artificial. *Revista de Ética Empresarial*, 7(6), 2054-2076.
- Pérez, L. (2021). Privacidad y protección de datos en la inteligencia artificial. Editorial Ariel.
- Rodríguez, A. (2018). Principios éticos para el desarrollo de inteligencia artificial. *Revista de Tecnología y Ética*, 20(2), 56-68.
- Russell, S. J. y Norvig, P. (2016). *Inteligencia artificial: un enfoque moderno*. Pearson Educación.
- Sánchez, P. (2021). Diseño centrado en el usuario para sistemas de IA éticos. *Harvard Business Review*, 32(1), 89-102.
- Scherer, M. U. (2016). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, 29(2), 353-400.
- Sequeira Cascante, M. y Sánchez Sánchez, J. (2023). La ética y la moral en la inteligencia artificial empresarial; una aproximación epistemológica y jurídica. *Revista Latinoamericana de Ciencias Sociales y Humanidades*, 4(4 SE-Artículos), 1147-1168. <https://doi.org/10.56712/latam.v4i4.1289>

- Smith, J. (2019). *Ética en la inteligencia artificial*. Editorial Tecnos.
- Tello Castrillón, C. y Rodríguez Córdoba, M. D. P. (2016). Fundamentos éticos disciplinares de la responsabilidad social organizacional. *Revista Venezolana de Gerencia*, 21(75). <https://doi.org/10.31876/revista.v21i75.21895>
- Terrones Rodríguez, A. (2018). Inteligencia artificial y ética de la responsabilidad. *Artificial Intelligence and the Ethic of Responsibility*, 4(22), 141-170. <https://doi.org/10.0.74.109/01235095.v4.n22.2018.8311>
- Thieullent, A.-L. (2023). Why addressing ethical questions in AI will benefit organizations Executive Summary. *Capgemini Research Institute*, 5-33.
- Villalba, J. F. (2020). Algor-ética: la ética en la inteligencia artificial. *Anales de La Facultad de Ciencias Jurídicas y Sociales de La Universidad Nacional de La Plata*, 17(50 SE-Derecho y Nuevas Tecnologías), 62. <https://doi.org/10.24215/25916386e062>

<https://doi.org/10.17163/abyaups.141.3>

Implicaciones éticas de la inteligencia artificial

Alex Darío Estrada García
Universidad Nacional de Educación (UNAE)
alex.estrada@unae.edu.ec
<https://orcid.org/0000-0001-5278-8221>

Jandry Geomar Chuni Gaona
Universidad Nacional de Educación (UNAE)
jandry.chuni@unae.edu.ec
<https://orcid.org/0009-0006-5618-076X>

Introducción

La eclosión de la inteligencia artificial generativa (IA) insta a discutir sobre las prácticas históricamente asentadas en la cotidianeidad del ser humano, una de ellas son las cuestiones relacionadas con la ética. En este escenario, es trascendente establecer reflexiones críticas sobre las implicaciones éticas que podrían presentarse en el contexto social a partir de la utilización de la IA. Terminologías como responsabilidad, decisiones, transparencia, seguridad, privacidad, vulnerabilidad, sesgos cognitivos demandan un abordaje filosófico urgente para establecer diálogos que guíen a la dilucidación de explicaciones plausibles relacionadas con el diseño, uso y evaluación de programas de IA en la cotidianeidad del ser humano.

Dignum (2018, p. 2) propone tres niveles de integración de la ética en la IA: *ethics by desing* (integración técnica de la ética en el algoritmo),

ethics in desing (revisión y análisis de las implicaciones éticas durante el diseño de programas), *ethics for desing* (códigos de conducta y estándares éticos para programadores). Hablar de ética en el contexto de la IA demanda pensar el proceso de diseño de los programas, al igual que escuchar las opiniones e intereses de las partes en cuestión. Es preciso tener presente que en varias narrativas con respecto a las implicaciones de la IA en la vida de los humanos se posicionan desenlaces catastróficos en un futuro próximo; en muchas ocasiones, esto corresponde a valoraciones morales descontextualizadas de los avances tecnocientíficos (Zamora, 2021).

El abordaje de las cuestiones expuestas tendrá lugar desde el relativismo moral metaético. Esta posición filosófica expone que los juicios morales a menudo dan directivas prácticamente contradictorias y ninguno de los juicios puede demostrar ser racionalmente superior al otro. La verdad moral no es absoluta, es relativa, por consiguiente, no se deberá entender los juicios morales de modo literal (Tasioulas, 2010; Zamora, 2021; Gowans, 2021).

Algunas interrogantes que conducen el desarrollo del presente capítulo son ¿qué es la IA? ¿qué creemos que es la IA? ¿en qué puede convertirse, de acuerdo con las proyecciones humanas, la IA? ¿las máquinas necesitan tomar decisiones éticas? ¿la ética en el diseño de la inteligencia artificial es fundamental? ¿cómo definir cánones éticos para el desarrollo, implementación y evaluación de programas de IA?

Se destaca que la IA trae muchos beneficios a los diferentes campos del conocimiento (Holmes *et al.*, 2021), así como a la cotidianidad de los seres humanos. Actualmente, el panorama es distinto al de la Revolución Industrial, cuando las máquinas eran un complemento del humano, mientras que ahora lo sustituyen (Brynjolfsson y McAfee, 2014). Esta sustitución da lugar a un sinnúmero de especulaciones relacionadas con el quehacer humano y de las máquinas en un futuro cercano. Un cambio de época trae consigo nuevas narrativas sobre posibilidades soñadas que se cultivan en el lenguaje común y se posicionan como verdades. Por

ejemplo, la inminente transformación tecnológica actual eran los relatos de la ciencia ficción de los siglos pasados.

Las nuevas realidades, sobre todo las virtuales, traen “consecuencias no deseadas de la tecnología” (Coeckelbergh, 2021, p. 85). Es decir, escapan de la predicción de los sistemas técnicos y su impacto en los diversos escenarios de aplicación. Por lo expuesto, más allá de una IA distópica, general, superinteligente, es decir, un escenario en que se alcanzaría una singularidad tecnológica; la preocupación debería ser centrarse en pensar y construir la ética en la IA, la cual según Coeckelbergh (2021) debería “[...] preocuparse del cambio tecnológico y su impacto en las vidas de los individuos, pero también de las transformaciones en la sociedad y en la economía” (p. 20). En consonancia con lo expuesto en este apartado, se propone como objetivo analizar y describir las implicaciones éticas de la IA y las repercusiones en los contextos sociales.

Inteligencia artificial generativa

Las definiciones de inteligencia artificial (IA) varían considerablemente entre los autores, lo que refleja las diversas perspectivas desde las cuales se aborda el concepto. Chesterman (2020) critica la ambigüedad del término “IA”; señala que su significado se diluye en un espacio de interpretaciones y aplicaciones, lo que puede dificultar su regulación y comprensión pública. Esta ambivalencia se manifiesta en la forma en que distintos sectores, desde la academia hasta la industria y la política, abordan la IA, cada uno con sus propias prioridades y preocupaciones.

Desde un enfoque pragmático, para los propósitos legales y regulatorios es eficaz definir la IA en función de casos de uso y riesgos específicos. De esta manera, los legisladores podrían enfocarse en delimitar la tecnología en contextos particulares donde se presenten riesgos o impactos significativos (Vilaça *et al.*, 2024). Esto como respuesta a la no existencia de una definición universalmente aceptada. Se sugiere que, en lugar de intentar abarcar toda la extensión del concepto de IA, que

es inherentemente vasto y en evolución, es más útil considerar cómo se aplica en situaciones concretas que involucran las decisiones autónomas, los datos o la seguridad.

De acuerdo con lo expuesto, el desafío radica no solo en encontrar un término común que pueda ser aceptado por todas las partes interesadas, sino también en abordar las implicaciones éticas y sociales que emergen en cada contexto, asegurando que las definiciones utilizadas no solo sean técnicamente precisas, sino también responsables y sensibles a las preocupaciones del público. Esto requerirá un esfuerzo colaborativo entre investigadores, reguladores y la sociedad en general para garantizar que la evolución de la IA beneficie a todos de manera equitativa.

Wang (2019) examina las diferentes maneras en que se ha definido la IA a lo largo del tiempo y las compara. Argumenta que se puede entender la IA desde cinco perspectivas: estructura, comportamiento, capacidad, función y principios. Esta clasificación resalta que las definiciones pueden variar según si intentan replicar el funcionamiento cerebral (estructura), su comportamiento observable (comportamiento), su habilidad para resolver problemas (capacidad) o las funciones específicas que desempeña (función). Además, esta diversidad en las definiciones no solo refleja el avance técnico en el campo, sino también la evolución de la comprensión humana sobre la cognición y la inteligencia, lo que se traduce en debates filosóficos sobre la naturaleza de la inteligencia artificial y su contraste con su contraparte humana.

En un enfoque más amplio, Sheikh *et al.* (2023) conceptualizan la IA como una “tecnología de sistema” que influye en todos los niveles de la sociedad, comparándola con revoluciones tecnológicas pasadas como la electricidad o el motor de combustión. A diferencia de Chesterman (2020) y Wang (2019), estos autores destacan no solo el aspecto técnico de la IA, sino también las profundas implicaciones sociales, económicas y éticas que conlleva. Su argumento sugiere que la integración de los sistemas informáticos autónomos generativos contemporáneos no es un mero fenómeno técnico.

Por ejemplo, Bruderer (2020) explica que la Agencia Europea de Defensa la define como la capacidad de los algoritmos para seleccionar la mejor opción entre un amplio conjunto de posibilidades, mientras que un Centro de Excelencia del Gobierno de EE. UU. la describe como sistemas capaces de realizar tareas de manera autónoma en contextos variables. El desarrollo histórico resalta el viaje de la IA desde un concepto teórico hasta convertirse en una herramienta tangible que permea el área industrial en el diario vivir. A pesar de enfrentar períodos de estancamiento, la IA ha resurgido con avances significativos que han revitalizado el interés en su investigación y aplicación. Programas como Deep Blue, que venció a campeones de ajedrez que respondió preguntas sobre una variedad de temas, simbolizan hitos en este resurgimiento. Las aproximaciones al desarrollo de la IA incluyen herramientas simbólicas, sistemas expertos y razonamiento basado en casos (CBR), que permiten a los sistemas aprender de experiencias previas y adaptarse a nuevas situaciones (Bruderer, 2020). Este enfoque multifacético posibilita que la IA no solo realice tareas específicas, sino que también mejore su rendimiento a medida que interactúa con su entorno, un principio fundamental en su evolución.

Las herramientas no simbólicas comprenden sistemas multiagente, formados por unidades de software que se adaptan y colaboran para abordar problemas complejos. Ha sido fundamental para las aplicaciones actuales de la IA, utilizando arquitecturas y algoritmos que aprenden de los datos a través de redes neuronales artificiales y técnicas de aprendizaje automático, que han evolucionado para permitir a las máquinas aprender y generalizar de manera más eficiente (Ali *et al.*, 2023). Esta evolución ha facilitado el surgimiento de aplicaciones prácticas multidisciplinar hasta la automoción, donde la capacidad de aprender de grandes volúmenes de datos es crucial.

Este avance ha sido impulsado por el acceso a grandes volúmenes de datos y mejoras en las capacidades computacionales, que han permitido entrenar modelos más complejos y precisos. La esencia de la IA generativa radica en algoritmos que aprenden de ejemplos no supervisados, con redes

neuronales como su estructura fundamental. Estas redes, organizadas en capas, ajustan los pesos de las neuronas para minimizar las funciones de costo, demostrando su capacidad de generalizar y aprender a partir de ejemplos (Ali *et al.*, 2023). Esta capacidad de aprendizaje no solo se traduce en una mayor precisión, sino también en una versatilidad que permite a la IA generativa adaptarse a diferentes tipos de contenido y aplicaciones.

La IA generativa comprende métodos y aplicaciones que crean contenido similar al producido por humanos, como texto, imágenes, música o código. Estas tecnologías se entrenan con grandes volúmenes de ejemplos reales, usualmente de manera no supervisada, para aprender características del contenido a generar. Con instrucciones del usuario, generan nuevos contenidos coherentes con su entrenamiento. Programas como *Gemini*, *ChatGPT* y *Copilot* han revolucionado la generación automática de contenido, permitiendo respuestas detalladas ante consultas precisas. Esto ha transformado la interacción con la información y ha abierto nuevas vías para la creatividad e innovación (Casar-Corredera, 2023).

Los enfoques bajo observación son importantes para la inteligencia artificial de tipo generativa. En el modelo supervisado se aprende bajo datos nominados, lo que facilita tareas como clasificación y regresión, aprendiendo a mapear entradas a salidas específicas. En cambio, el aprendizaje no supervisado utiliza datos no etiquetados, permitiendo que los modelos identifiquen patrones sin guía explícita. Esto es esencial para la generación de nuevas muestras en IA generativa, como en *GANs* (*Generative Adversarial Networks*) y *autoencoders* variacionales (Kingma y Welling, 2013). La capacidad de aprender sin supervisión es crucial en contextos donde los datos etiquetados son costosos o difíciles de obtener.

Las aplicaciones de IA generativa trascienden la generación de texto. Herramientas como *DALL-E* y *MidJourney* crean imágenes 'originales' en diversos estilos artísticos, mientras que otras se enfocan en la composición musical y la generación de efectos a largo plazo. La definición de estas herramientas varía según el contexto, reflejando las amplias capacidades de las IA generativas en distintos ámbitos creativos (Gupta *et al.*, 2024).

Esta flexibilidad y adaptabilidad subrayan el potencial no solo como herramientas de producción, sino también como motores de innovación en el arte y el diseño.

En el diseño de productos, por ejemplo, puede generar prototipos innovadores basados en tendencias de mercado y preferencias del consumidor, lo que optimiza el proceso de desarrollo y reduce el tiempo de comercialización (Suphavarophas *et al.*, 2024). En la programación de videojuegos, permite la creación de niveles y escenarios únicos que se adaptan a las decisiones de los jugadores, generando experiencias personalizadas y dinámicas (Hu *et al.*, 2023). Además, en campos como la biotecnología y la medicina, la IA generativa puede simular estructuras moleculares y proponer nuevos fármacos, acelerando el descubrimiento y desarrollo de tratamientos efectivos (Vamathevan *et al.*, 2019). Estos ejemplos ilustran cómo la IA generativa no solo mejora la eficiencia, sino que también impulsa la creatividad y la innovación en disciplinas altamente especializadas.

La adaptación de la arquitectura de los modelos de IA generativa a las características específicas de cada dominio es esencial para maximizar su efectividad. Este es el caso de los diseños de productos, los modelos pueden incorporar algoritmos de optimización para evaluar la viabilidad funcional y estética de los diseños generados (Bongiorno *et al.*, 2022). En el caso de los videojuegos, se pueden utilizar arquitecturas como redes neuronales convolucionales (CNN) para procesar y generar elementos visuales, mientras que redes recurrentes (RNN) pueden ayudar en la creación de narrativas dinámicas (Hu *et al.*, 2023). Esta capacidad de personalizar la arquitectura del modelo según las necesidades del dominio permite que la IA generativa sea una herramienta versátil y poderosa en la innovación y creación de contenido en múltiples campos. Esta personalización también implica un desafío adicional: la idea de un amplio espectro de comprensión características del dominio para que los modelos puedan ser diseñados de manera efectiva considerando diversos lineamientos éticos.

La IA generativa permite crear contenido personalizado que se adapta a las necesidades de cada usuario, utilizando algoritmos que analizan datos sobre comportamientos y elecciones previas. Esta personalización mejora la experiencia del usuario, ya que ofrece recomendaciones y soluciones más relevantes, desde recomendaciones de productos hasta contenidos en plataformas de *streaming* (Ricci *et al.*, 2010). Sin embargo, la personalización propone retos significativos en base de la privacidad y ética, puesto que implica la administración y almacenamiento de muchos datos. Esto puede dar lugar a preocupaciones sobre cómo se utilizan los datos, quién tiene acceso a ellos y cómo se protegen los derechos del usuario (Hagendorff, 2020).

La necesidad de un enfoque ético y transparente en la implementación de la IA generativa es crucial para garantizar que los beneficios de la personalización no se vean eclipsados por violaciones de privacidad y desconfianza en la tecnología (Möllmann *et al.*, 2021). Este aspecto ético es fundamental, dado que una implementación irresponsable puede socavar la confianza del usuario y limitar la adopción de tecnologías basadas en IA generativa.

Sesgo algorítmico, responsabilidad y transparencia

En el mundo contemporáneo, la IA está posicionando una visión de intento de modelado de la sociedad desde formas profundas y omnipresentes. Esta tecnología tiene el poder de transformar campos educativos, económicos y médicos; sin embargo, con grandes capacidades vienen grandes responsabilidades. En particular, existen preocupaciones crecientes sobre el sesgo algorítmico y sobre cómo se garantiza el desarrollo e implementación de los mecanismos de inteligencia artificial. En este epígrafe se explora estos temas, subrayando la importancia de un enfoque ético y responsable en el uso de la IA.

Introducción a las causas y consecuencias del sesgo algorítmico

El sesgo algorítmico es una problemática crítica que surge cuando los modelos de IA reflejan o amplifican prejuicios preexistentes en los datos o en las estructuras de los algoritmos. Estos sesgos no son una simple falla técnica, sino una manifestación de cómo se desarrolla y entrena la tecnología en un mundo que ya está lleno de desigualdades y prejuicios (Aparicio-Gómez y Cortéz, 2024).

El sesgo algorítmico tiene múltiples causas. Un principal desencadenante es la presencia de datos sesgados, los algoritmos de IA aprenden a partir de datos preexistentes y, si estos datos contienen prejuicios, el resultado será un modelo igualmente sesgado. Por ejemplo, esto sucede en el entrenamiento de los sistemas de IA que generan texto; si el modelo de IA fue entrenado con datos que contienen estereotipos relacionados con género, etnia, cultura, religión, es muy probable que el sistema de IA genere textos que refuerzan estos estereotipos, de esta forma se podría ocasionar algún tipo de discriminación, exclusión o desconfianza en la adopción de los programas de IA. Otro ejemplo claro se da en los sistemas de contratación automatizada; si el historial de contratación de una empresa muestra una preferencia por contratar a hombres en lugar de mujeres, un algoritmo entrenado con esos datos tenderá a replicar ese comportamiento, descartando a candidatas calificadas por prejuicio inherente en los datos (Kheiri *et al.*, 2024).

Otra causa importante es el diseño del algoritmo. Los desarrolladores están sujetos a introducir sesgos involuntariamente al elegir cómo se deben ponderar ciertas características o al definir la estructura del modelo. Los algoritmos son productos humanos y, como tales, pueden reflejar las limitaciones y perspectivas del equipo de desarrollo. En los sistemas de reconocimiento facial, por ejemplo, se ha demostrado que los algoritmos tienen mayores tasas de error al identificar personas de color en comparación con personas de tez blanca. Esta situación se debe, en

parte, a que los datos utilizados en el entrenamiento tienden a tener una mayor representación de personas blancas, mientras que la diversidad étnica es insuficiente (Etxeberria, 2024). En el ámbito de la publicidad, los algoritmos de segmentación pueden discriminar de manera indirecta al mostrar ciertos anuncios solo a un segmento específico de la población, limitando las oportunidades para otros. Esto afecta de manera exponencial el acceso equitativo a la información, servicios y productos.

Sobre *cómo mitigar el sesgo algorítmico en la IA*, inicialmente, se sostienen que requiere un enfoque integral que aborde las causas desde diferentes perspectivas (Kaufman, 2024). En primer lugar, es esencial contar con conjuntos de datos que sean representativos y diversos. Un buen punto de partida es realizar una revisión y curación cuidadosa de los datos, para garantizar que no incluyan patrones sesgados de comportamiento o estereotipos sociales (Raghavan *et al.*, 2020). Por ejemplo, en el estudio de la *ciencia de datos* se evidencian esfuerzos por desarrollar técnicas de muestreo para equilibrar el peso de ciertos grupos dentro de los conjuntos de datos de entrenamiento.

Otra estrategia es la auditoría de los algoritmos enfocada en la evaluación constante de los modelos para identificar y mitigar posibles sesgos. Esta práctica puede llevarse a cabo mediante auditorías internas o, preferiblemente, por auditores externos que puedan ofrecer una perspectiva imparcial. Es importante tomar en cuenta que los desarrolladores deben contar con herramientas que posibiliten analizar la equidad de los resultados generados por sus modelos, como indicadores que evalúan la disparidad entre los diferentes grupos demográficos.

A partir de los aportes de Burrell (2016) menciona que la transparencia en los algoritmos es fundamental, ya que no solo permite a los usuarios entender el proceso de toma de decisiones, sino que también les ofrece la oportunidad de identificar y corregir sesgos que podrían contribuir a la perpetuación de desigualdades en las decisiones automatizadas. Finalmente, es esencial promover la diversidad en los equipos de desarrollo. Los desarrolladores y científicos de datos de diversas culturas,

géneros y perspectivas pueden identificar y abordar problemas que por ciertas cuestiones están sujetas a pasar por alto.

La responsabilidad en la IA implica que las empresas y los desarrolladores asuman la obligación ética y moral de los impactos que sus sistemas tienen en la sociedad. La responsabilidad debe ser un principio rector en todas las fases de la vida de un sistema de IA: desde el diseño, pasando por el entrenamiento, la implementación, el uso y la evaluación. Los principios de responsabilidad requieren que las empresas tengan la responsabilidad de las decisiones y acciones de sus algoritmos. Esto significa que no se pueden escudar en la autonomía tecnológica como una forma de evadir la rendición de cuentas.

Una problemática común de este tipo es atribuir la responsabilidad de una decisión errónea a la máquina, ignorando que la programación y el uso del sistema son responsabilidad humana (Mittelstadt, 2019). Para evitar esto, las empresas están en la obligación de establecer mecanismos claros para identificar quién es responsable en cada etapa del desarrollo y uso del sistema. Esto se puede lograr mediante la asignación de roles específicos de revisión ética dentro del equipo de desarrollo, y documentando cada decisión técnica que impacte la equidad y justicia del sistema. La transparencia es un componente vital para lograr la responsabilidad en el trabajo y fomento de la creación de la IA. La transparencia algorítmica implica que los procesos que llevan a la adquisición de la voluntad por parte de la inteligencia artificial los cuales deben ser comprensibles para todos los usuarios. Esto incluye la posibilidad de acceder a información sobre qué datos se utilizaron para entrenar el sistema y cómo estos influyen en sus resultados.

Un ejemplo donde la falta de transparencia ha generado controversia es el caso de los algoritmos de crédito, que determinan la viabilidad crediticia de las personas (Coeckelbergh, 2020a; 2022). Si los usuarios no comprenden por qué se les negó un crédito, se fomenta la desconfianza en el sistema financiero. En el libro *La sociedad del desconocimiento* de Daniel Innerarity (2022), se explica que el problema radica en la 'caja negra' de

los algoritmos, donde ni siquiera los desarrolladores entienden con precisión cómo un sistema de aprendizaje profundo llega a una conclusión (Pasquale, 2015). La transparencia debe incluir no solo la posibilidad de revisar los resultados, sino también hacer públicos los datos y métodos utilizados, de manera que los expertos externos puedan analizarlos y realizar sugerencias.

Conforme con lo expuesto, se plantean algunas ideas que invitan a pensar en estrategias direccionadas a la mitigación del sesgo algorítmico. En la literatura científica se debate la necesidad de implementar estrategias efectivas desde la fase de diseño la IA. Esto incluye la recopilación de datos diversificados y representativos que reflejen la realidad de la población en su conjunto. López (2023) argumenta que la inclusión de diferentes datos de entrenamiento es muy importante para evitar el sesgo en los algoritmos que se usan. Esto puede incluir técnicas de corrección de sesgos que ajusten los resultados algorítmicos para garantizar que no se favorezca a un grupo sobre otro (Coeckelbergh, 2022). Los estudios en esta área están en constante progreso y desarrollan métodos innovadores para abordar el sesgo en tiempo real. Además, es importante fomentar la colaboración interdisciplinaria en la evolución de la IA. Involucrar a expertos en ética, sociología, derecho y otras disciplinas en el proceso de diseño puede ayudar a identificar y abordar preocupaciones relacionadas con el sesgo y la justicia social.

Los desafíos que presenta la IA demanda entender la profunda interrelación entre sesgo algorítmico, responsabilidad y transparencia. La presencia de sesgos en los algoritmos afecta directamente la calidad y equidad de las decisiones, lo cual se traduce en injusticias y desigualdades para aquellos afectados por los resultados del sistema de IA. Por tanto, la responsabilidad implica que los desarrolladores y empresas deben ser conscientes de los sesgos existentes, hacerlos visibles mediante la transparencia y trabajar activamente para reducirlos (Gallifant *et al.*, 2024).

Para enfrentar los problemas derivados del sesgo algorítmico, la transparencia es clave. Sin información clara sobre los datos y métodos

utilizados, se torna casi imposible que los reguladores, académicos y el público en general puedan verificar si un sistema es justo y equitativo. Al promover la transparencia, las empresas tecnológicas están fomentando la confianza en la tecnología y sus nuevos productos. Además, la responsabilidad también implica trabajar activamente para la mejora continua de los sistemas, incorporando mecanismos de auditoría y control que aseguren la equidad, y cuidando que los sistemas de IA no se utilicen para perpetuar o amplificar desigualdades existentes.

Regulación y marcos éticos para la IA

Para este cometido es fundamental el análisis y regulación con el propósito de establecer un marco que garantice el desarrollo ético y responsable de la IA. Dado que los gestores de IA tienen el potencial de impactar a nivel mundial, es importante que las políticas públicas ofrezcan un marco de referencia que defina límites y establezca responsabilidades claras. Algunos países han avanzado en la creación de regulaciones para la IA. La Unión Europea, por ejemplo, ha propuesto la “Ley de Inteligencia Artificial”, que se basa en un enfoque de regulación proporcional al riesgo que representa un sistema (Binns, 2022). Los sistemas de IA que tienen el potencial de afectar derechos fundamentales o causar daños significativos están sujetos a requisitos más estrictos de transparencia y evaluación de riesgos.

Las recomendaciones propuestas por la UNESCO (2021), también ofrecen directrices para el desarrollo responsable de la IA. Estos marcos se centran en garantizar que la IA sea beneficiosa para la sociedad, respetando los derechos humanos y promoviendo el bienestar. Sin embargo, la implementación de estos principios depende de la voluntad política y empresarial de cumplir con los compromisos adquiridos. Es necesario que la ética no sea solo una declaración de intenciones, sino que esté integrada en el proceso de diseño e implementación (Floridi y Cowls, 2019).

Es inminente la dificultad de aplicar un marco ético global que responda a todas las interrogantes que pueden surgir en los diversos

escenarios de aplicación de la IA; esto debido, en primera instancia, a la diversidad de definiciones asignadas al término *ética*. Históricamente se ha evidenciado los acuerdos y desacuerdos en alcanzar definiciones plausibles; sin embargo, lo que se tiene claro es que el uso de este término es relativo a un conjunto de personas que comparten valores, creencias, dogmas, ideologías. Por tal razón, alcanzar un marco ético en el que se refleje el consenso de qué valores y prioridades deben ser atendidas en el desarrollo de la IA, es poco alcanzable. Otra razón que imposibilita alcanzar un marco ético general que exponga ciertos principios es la complejidad de los sistemas de IA, en algunos ámbitos son impredecibles el alcance de estos sistemas. A esto se suma el avance vertiginoso de la tecnología, los cambios continuos harían obsoleto cualquier intento de generalización de lineamientos o principios éticos generales.

A pesar de las diversas dificultades en el contexto ético, puede implementarse medidas útiles para el desarrollo adecuado de la IA. Por ejemplo, a nivel estatal, como ya se evidenció en el inicio de este apartado, es oportuno crear regulaciones y leyes para normar el desarrollo y uso de la IA. Así mismo, es de suma importancia la cooperación internacional destinada a intercambiar avances tecnológicos y marcos éticos que permitan mitigar determinados riesgos. En cuanto a las instituciones que desarrollan sistemas de IA están llamados a transparentar el funcionamiento y proyectar los posibles impactos que pueda llegar a tener; para ello, es fundamental contar con códigos de ética propios de la institución.

De acuerdo con lo expuesto, se plantea un interrogante que invita a pensar el desarrollo, funcionamiento y evaluación de los sistemas de IA: ¿cuál es el rol de la filosofía en el desarrollo tecnológico? Diversas son las razones por las que se debe subrayar la filosofía de la tecnología, sobre todo para constituir un marco ético. Por ejemplo, posibilita un análisis conceptual de términos como 'ética', 'inteligencia', 'responsabilidad', 'conciencia' que se emplean sin concisión ni precisión, lo cual dificulta el debate sobre el desarrollo tecnológico y sus dimensiones. De igual forma, la filosofía proporciona herramientas para el análisis de dilemas

éticos que eclosionan en el ámbito en mención. Estas son dos de varias funciones que podría tener la filosofía de la tecnología, sobre todo para el fortalecimiento de un marco conceptual y metodológico encaminado a dilucidar cuestiones relacionadas con la ética de/en IA.

Privacidad de los datos

Una de las características propias de la tercera década del siglo XXI es la cantidad de información que generan los seres humanos a partir de su interacción con el medio social y natural. Las sofisticadas Tecnologías de la Información y Comunicación (TIC) son medios para crear, difundir, acceder a una vasta información dando paso a una transferencia de datos globalizada, comúnmente se lo ha llamado *big data*. Por lo tanto, la sociedad está inmersa en un mundo *datificado*, preguntarse por la privacidad de los datos es crucial. En el contexto de la IA, los datos son la materia prima para el entrenamiento de los modelos de lenguaje masivos (large language models [LLM]) que están constituidos por redes neuronales artificiales (Wang *et al.*, 2023). En tal sentido, la IA se favorece de los LLM por la capacidad de estos para el procesamiento de datos y, a través de la conjugación de algoritmos, ofrecen respuestas en lenguaje natural a preguntas planteadas por los humanos (Estrada-García y Narváez, 2024).

Es oportuno precisar en el tipo de información del que están constituidos los gigantescos bulos de datos. Como se había mencionado en la parte introductoria del presente capítulo, actualmente, la tecnología es el campo central al que el ser humano se ha mudado, cada vez más son las actividades digitales que alimentan los bulos con información de todo tipo. Por ejemplo, Coeckelbergh (2021) argumenta que “la IA, y en particular las aplicaciones de aprendizaje automático que operan con el *big data* a menudo implican la recogida y el uso de información personal” (p. 85). El entrenamiento de los programas de IA se da con todo tipo de información, sea esta contrastada o no; por ello, la predicción de resultados es cuestionable tanto por el tipo de información utilizada como por la configuración de algoritmos, puesto que pueden crearse a partir de sesgos.

La IA se respalda de campos del conocimiento como la ciencia de datos, específicamente, la estadística se constituye como uno de los aliados más importantes en el entrenamiento de algoritmos. El procesamiento de datos se da a partir de lo conocido como *aprendizaje automático* que, dependiendo de los objetivos, puede ser supervisado o no. Como consecuencia se obtiene la predicción de resultados a partir de la identificación de patrones en una gran cantidad de datos, este proceso en el lenguaje común se lo conoce como minería de datos. El procedimiento mencionado se complementa con el *aprendizaje por refuerzo*, es cual está direccionado para retroalimentar la predicción que el programa de IA ofreció.

Un caso importante que hay que destacar es el relacionado con las redes sociales —cualquiera que sea— con la privacidad de los datos personales que insertan los usuarios. Evidentemente, en algunas ocasiones estas aplicaciones solicitan consentimiento a los usuarios para utilizar los datos personales con el propósito de ‘ofrecerle una mejor experiencia’. Sin embargo, ¿los usuarios conocen lo que estas redes sociales hacen con sus datos? No es posible acceder transparentemente a información sobre cómo se están utilizando los datos que producen los seres humanos diariamente. Esta práctica de las redes sociales se constituye como un mecanismo en que se utiliza al ser humano como mano de obra gratuita, es más se podría decir que se da una explotación, todo esto genera datos que serán analizados y procesados por la IA.

Un uso ético de la IA requiere que los datos sean recogidos, procesados y compartidos de una forma que respete la privacidad de los individuos y su derecho a saber lo que ocurre con ellos, al acceso de los mismos, a objetar su recogida o procesamiento y a saber que se están recogiendo y procesando y (si procede) que están expuestos a la decisión tomada por una IA. (Coeckelbergh, 2021, p. 86)

Los datos generados a partir de las acciones humanas en entornos mediados por la tecnología, se utilizan para entrenar a los programas de IA. Por ello, es trascendente discutir los principios éticos bajos los cuales puede darse la utilización de los datos, configuración de algoritmos,

retroalimentación. La preocupación humana por una ética para/de la IA nace de lo que creemos que es y en lo que pueda convertirse la IA. Por ejemplo, se ha mencionado que la IA tiene el potencial de convertirse en una inteligencia superior a la que posee el ser humano. Al respecto, la científica Boden (2016) explica que, si bien está de acuerdo en la eminente configuración de una IA general, superinteligente, sostiene que esto no ocurrirá en la práctica; por lo tanto, la IA, comparada con la inteligencia humana, es menos prometedora de lo que mucha gente piensa y aún no conseguirá superarla, al menos no sucederá en las próximas décadas.

Manipulación, vigilancia y totalitarismo

La obtención, procesamiento y análisis de grandes cantidades de información insta a discutir sobre la manipulación, vigilancia y totalitarismo en diversos escenarios. La información personal de los usuarios ha sido utilizada por programas de IA para definir preferencias relacionadas con productos y servicios ofertados por empresas. Las redes sociales gracias a sus características propias acaparan datos personales de los usuarios que son usados, en principio, para una supuesta personalización de la información que estas redes sociales ofrecen al público. Sin embargo, está la posibilidad de que los datos personales sean usados para fines políticos, económicos o de otra índole.

Indudablemente, tecnologías como la internet y los *smartphones* han cambiado totalmente la realidad histórica del ser humano y no siempre estas tecnologías aportan solo cuestiones positivas, sino también negativas (Tollon, 2024). De hecho, las nuevas tecnologías instauran formas de manipulación, vigilancia y totalitarismo diferentes a las tradicionales, formas que quizá sean poco evidentes para las masas, no obstante, efectivas para determinados fines. Los *smartphones*, acompañantes diarios de los seres humanos, se constituyen a partir de aplicaciones que generan datos continuamente, a través de estas se da una suerte de vigilancia indirecta.

En los escenarios de la política, la IA en equipo con las redes sociales puede convertirse en una herramienta utilizada para la manipulación de

las sociedades. Por ejemplo, mediante la utilización de los datos generados por las redes sociales se puede influir en campañas políticas. Por las cuestiones expuestas y otras más, se habla del uso ético de las tecnologías ante la preocupación por las decisiones personales o de grupos de poder que no están prescritas por la ley (Etzioni y Etzioni, 2017).

De acuerdo con la visión de Coeckelbergh (2011) “[u]na forma de analizar y evaluar lo que las tecnologías de la información hacen y podrían hacer a las personas y a la sociedad, es utilizar el enfoque de capacidad como un marco normativo-ético” (p. 81). El enfoque referido hace hincapié en la comprensión de una ética de la tecnología centrada en las capacidades humanas. Un enfoque funcional que propone revisar el proceso de cómo la tecnología aporta a lo que las personas son capaces de hacer de acuerdo con un contexto social, cultural y tecnológico particular.

De conformidad con lo expuesto, la preocupación por la ética, específicamente, en el escenario de la IA recae en dos cuestiones. La primera, consiste en lo que la IA puede hacer, y segundo, en lo que el ser humano está en la capacidad de hacer con la IA. Pensar la primera cuestión lleva a la aceptación de una IA autónoma, si bien esta idea sea, al menos por unos años, utópica. Mientras que la segunda cuestión es la que demanda de la atención prioritaria con el propósito de plantear la integración de la ética en la configuración técnica de los programas de la IA, una revisión y análisis de las implicaciones éticas durante el diseño y códigos de conducta y estándares éticos para programadores (Morley *et al.*, 2021; 2023).

Shanahan (2015) explica que la IA puede llevar a los humanos a ser “menos capaces de pensar o de decidir por sí mismos lo que hacer” (p. 170). Esta cuestión desencadena un sinnúmero de preocupaciones como la infantilización de los usuarios ante la posibilidad de que la IA asuma tareas propias del ser humano, como las cognitivas. De ser este el caso, los usuarios consumen combinaciones de algoritmos y la actividad cognitiva se verá menos forzada. Lo presentado son escenarios en los que la manipulación está presente, de igual forma, la explotación de datos personales (Coeckelbergh, 2019; 2020b).

Una cuestión que se debe discutir es si ¿las máquinas equipadas con inteligencia artificial necesitan ser capaces de tomar decisiones morales?, de ser este el caso ¿existe responsabilidad moral en las decisiones que toman las máquinas? Al respecto, Etzioni y Etzioni (2017) explica que las máquinas con IA no necesitan tomar decisiones éticas, siempre y cuando, estén en entornos gobernados por la ley, puesto que estas regulan y guían el comportamiento adecuado, al igual que establecerían los límites de acción de las máquinas y de los humanos que configuran a estas.

Por último, en un mundo interconectado es evidente la susceptibilidad de los programas y dispositivos tecnológicos a sufrir hackeos o cualquier tipo de ataque malintencionado. De igual forma está la vigilancia y el control de los gobiernos a las sociedades a través de la IA. Como respuesta, es fundamental promover una ética positiva para la privacidad, la transparencia y la rendición de cuentas. Todos son, una u otra forma, vulnerables a la tecnología, desde los problemas ligados a la seguridad hasta la creciente dependencia humana a las máquinas con IA; es decir, la creciente designación de tareas a las máquinas.

Conclusión

La vida de los seres humanos cambia constantemente por la introducción de programas, dispositivos tecnológicos cada vez más sofisticados, lo que demanda de una reflexión crítica a fin de discutir y exponer implicaciones positivas o negativas que podrían originarse de la introducción de programas de IA en la cotidianidad humana. La IA ha dotado a las máquinas funciones que hacen posible la interacción mediante lenguaje natural con el ser humano, posibilidad que hace que muchas tareas históricamente complejas, se hagan más sencillas.

Los humanos nos hemos convertido en seres tecnológicos, a diario se utiliza un sin número de programas, dispositivos tecnológicos para la consecución de objetivos de la vida diaria, profesional, entre otras cuestiones dependiendo de los intereses. Cabe destacar que no todos los avances tecnológicos son positivos, también se dan aspectos que pueden

ser catalogados como negativos. Por ejemplo, la privacidad de los datos que los usuarios generen y son utilizados para el entrenamiento de programas de IA a partir del aprendizaje automático. De la misma manera, la transparencia en la configuración de algoritmos y el aprendizaje por refuerzo que retroalimenta al programa de IA para mejorar las respuestas.

Es importante mencionar que el sesgo algorítmico, la responsabilidad y la transparencia son elementos fundamentales en el desarrollo de sistemas de IA que respeten los principios éticos y promuevan el bienestar social. El sesgo algorítmico puede tener graves repercusiones en distintos ámbitos —desde la justicia hasta el acceso a servicios financieros— y su mitigación requiere un enfoque sistémico que abarque desde la calidad de los datos hasta la estructura de los algoritmos y la diversidad en los equipos de desarrollo. Dependerá de la capacidad del ser humano para enfrentar estos desafíos éticos con honestidad y dedicación. Si se logran implementar de manera adecuada prácticas de transparencia y responsabilidad y se trabaja activamente para mitigar los sesgos, se podrá crear sistemas de IA justos, inclusivos y capaces de mejorar la vida de las personas, sin importar su origen o condición.

La responsabilidad y la transparencia son esenciales para asegurar que los desarrolladores y empresas asuman las consecuencias de sus creaciones. Estos principios no solo promueven la confianza del público en los sistemas de IA, sino que también permiten la rendición de cuentas y la implementación de mejoras para garantizar la equidad. El desafío de reducir los sesgos, implementar la transparencia y garantizar la responsabilidad recae en los desarrolladores, las empresas y, en última instancia, los reguladores y la sociedad en general.

Una conjugación entre la IA, acceso a internet y las redes sociales ofertará más oportunidades para la manipulación, aparte que diversos mecanismos complejizan este problema. Por ejemplo, investigadores de la Universidad de Washington (BBC, 2017) crearon con IA un discurso de Barack Obama. Acciones como estas pueden desencadenarse con mayor frecuencia, la creación y difusión de noticias falsas elaboradas por IA,

utilizando la identidad de personajes influyentes en la sociedad puede ocasionar impactos catastróficos en diversos campos.

La mezcla entre hechos y ficción cada vez es más frecuente, el problema de la verdad y posverdad entra en cuestión; saber distinguir será uno de los desafíos en de la era de la IA. La manipulación, vigilancia y totalitarismo disfrazados de libertad acecha a una sociedad tecnológica, gobernada por máquina y programas que cada vez escapan del “control” humano.

Referencias bibliográficas

- Ali, R., Hussain, A., Nazir, S., Khan, S. y Khan, H. U. (2023). Intelligent Decision Support Systems - An Analysis of Machine Learning and Multicriteria Decision-Making Methods. *Applied Sciences*, 13(22), 12426. <https://doi.org/10.3390/app132212426>
- Aparicio-Gómez, O. y Cortés Gallego, M. (2024). Desafíos éticos de la inteligencia artificial en la personalización del aprendizaje. *Revista Interamericana de Investigación Educación y Pedagogía* 17(2), 377-392. <https://doi.org/10.15332/25005421.10000>
- BBC (17 de julio de 2017). El falso Barack Obama creado con inteligencia artificial capaz de hablar como si fuera el original. <https://bit.ly/46OPLWO>
- Binns, R. (2022). Human Judgment in algorithmic loops: Individual justice and automated decision-making. *Regulation & Governance*, 16(1), 197-211. <https://doi.org/10.1111/rego.12358>
- Boden, M. (2016). *AI: Its Nature and Future*. Oxford University Press.
- Bongiorno, A., De La Torre, L. y Caffaro, F. (2022). Generative design in the digital era: A review. *Journal of Computing and Information Science in Engineering*, 22(1), 1-20.
- Bruderer, H. (2020). *Milestones in analog and digital computing* (3rd ed.). Springer.
- Brynjolfsson, E. y McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & company.
- Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 2-12. <https://doi.org/10.1177/2053951715622512>
- Casar-Corredera, J. (2023). Inteligencia artificial generativa. *Anales de la Real Academia de Doctores*, 8(3), 475-489.

- Chesterman, S. (2020). Artificial intelligence and the limits of legal personality. *International & Comparative Law Quarterly*, 69(4), 819-844. <https://doi.org/10.1017/S0020589320000366>
- Coeckelbergh, M. (2011). Human development or human enhancement? A methodological reflection on capabilities and the evaluation of information technologies. *Ethics and Information Technology*, 13, 81-92. <https://doi.org/10.1007/s10676-010-9231-9>
- Coeckelbergh, M. (2019). Artificial intelligence: some ethical issues and regulatory challenges. *Technology and regulation*, 2019, 31-34. <https://doi.org/10.26116/techreg.2019.003>
- Coeckelbergh, M. (2022). *Robot ethics*. MIT Press.
- Coeckelbergh, M. (2020a). *AI ethics*. MIT Press.
- Coeckelbergh, M. (2020b). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and engineering ethics*, 26(4), 2051-2068. <https://doi.org/10.1007/s11948-019-00146-8>
- Estrada-García, A. y Narváez, J. P. (2024). ChatGPT y la superficialidad del conocimiento: implicaciones académicas y éticas en el siglo XXI. *Yachana Revista Científica*, 13(2), 19-36. <https://doi.org/10.62325/10.62325/yachana.v13.n2.2024.911>
- Etxeberría Guridi, J. F. (2024). The use of artificial intelligence (AI) systems for remote biometric identification in publicly accessible spaces in the European AI law. *Actualidad Jurídica Iberoamericana*, 21, 528-565.
- Etzioni, A. y Etzioni, O. (2017). Incorporating ethics into artificial intelligence. *The Journal of Ethics*, 21, 403-418. <https://doi.org/10.1007/s10892-017-9252-2>
- Floridi, L. y Cows, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1-18. <https://doi.org/10.1162/99608f92.8cd550d1>
- Gallifant, J., Bitterman, D., Celi, L., Gichoya, J., Matos, J., McCoy, L. y Pierce, R. (2024). Ethical debates amidst flawed healthcare artificial intelligence metrics. *npj Digital Medicine*, 7(243). <https://doi.org/10.1038/s41746-024-01242-1>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S. y Courville, A. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27.
- Gowans, C. (2021). *Moral Relativism*. The Stanford Encyclopedia of Philosophy. Edited by Edward N. Zalta. <https://bit.ly/4gYRdKY>
- Gupta, P., Ding, B., Guan, C. y Ding, D. (2024). Generative AI: A systematic review using topic modelling techniques. *Data and Information Management*, 8(2), 100066. <https://doi.org/10.1016/j.dim.2024.100066>

- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and machines*, 30(1), 99-120. <https://doi.org/10.1007/s11023-020-09517-8>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B. y Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 2, 504-526. <https://doi.org/10.1007/s40593-021-00239-1>
- Hu, Z., Ding, Y., Wu, R., Li, L., Zhang, R., Hu, Y. y Fan, C. (2023). Deep learning applications in games: a survey from a data perspective. *Applied Intelligence*, 53(24), 31129-31164. <https://doi.org/10.1007/s10489-023-05094-2>
- Innerarity, D. (2022). *La sociedad del desconocimiento*. Galaxia Gutenberg.
- Kaufman, D. (2024). Logic and foundations of artificial intelligence and society's reactions to maximize benefits and mitigate harm. *Unisinos Journal of Philosophy*, 25(1), 1-13. <https://doi.org/10.4013/fsu.2024.251.10>
- Kheiri, M. A., Qashta, N., & Aljaradat, D. I. (2024). Authenticity of using artificial intelligence systems in proving electronic evidence. *Pacific Journal of Legal & Social Sciences*, 23(1), 37-52. <https://doi.org/10.57239/PJLSS-2025-23.1.004>
- Kingma, D. P. y Welling, M. (2013). Auto-encoding variational Bayes. *arXiv preprint arXiv:1312.6114*.
- López, D. M. (2023). Retos de la inteligencia artificial y sus posibles soluciones desde la perspectiva de un editorialista humano. *Biomédica*, 43(3), 309-314.
- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501-507. <https://doi.org/10.1038/s42256-019-0114-4>
- Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mökander, J. y Floridi, L. (2021). Ethics as a service: a pragmatic operationalisation of AI ethics. *Minds and Machines*, 31(2), 239-256. <https://doi.org/10.1007/s11023-021-09563-w>
- Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M. y Floridi, L. (2023). Operationalising AI ethics: barriers, enablers and next steps. *AI & Society*, 38, 411-423. <https://doi.org/10.1007/s00146-021-01308-8>
- Möllmann, N. R., Mirbabaie, M. y Stieglitz, S. (2021). Is it alright to use artificial intelligence in digital health? A systematic literature review on ethical considerations. *Health Informatics Journal*, 27(4), 14604582211052391. <https://doi.org/10.1177/14604582211052391>
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Raghavan, M., Barocas, S., Kleinberg, J. y Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In Proceedings of

- the 2020 Conference on Fairness, Accountability, and Transparency (pp. 469-481).
- Ricci, F., Rokach, L. y Shapira, B. (2010). Introduction to recommender systems handbook. In Ricci, F., Rokach, L., Shapira, B., Kantor, P. (eds). *Recommender systems handbook* (pp. 1-35). Springer. https://doi.org/10.1007/978-0-387-85820-3_1
- Shanahan, M. (2015). *The technological singularity*. The MIT Press.
- Sheikh, H., Prins, C., Schrijvers, E. (2023). AI as a System Technology. In: Mission AI. Research for Policy (pp. 84-134). Springer. https://doi.org/10.1007/978-3-031-21448-6_4
- Suphavarophas, P., Wongmahasiri, R., Keonil, N. y Bunyarittikit, S. (2024). A Systematic Review of Applications of Generative Design Methods for Energy Efficiency in Buildings. *Buildings*, 14(5), 1311.
- Tasioulas, J. (2010). Relativism, realism, and reflection. *Inquiry*, 41(4), 377-410. <https://doi.org/10.1080/002017498321706>
- Tollon, F. (2024). Technology and the situationist challenge to virtue ethics. *Science and Engineering Ethics*, 30(10), 1-17. <https://doi.org/10.1007/s11948-024-00474-4>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura – UNESCO (2021). *Recomendación sobre la ética de la inteligencia artificial*. Conferencia General de la UNESCO, Paris del 9 al 24 de noviembre de 2021. <https://bit.ly/3KsXvpW>
- Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G. y Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug discovery*, 18(6), 463-477. <https://doi.org/10.1038/s41573-019-0024-5>
- Vilaça, M., Lopes, I. y Ferro, M. (2024). AI beyond a new academic hype: an interdisciplinary theoretical analytical experiment (computational, linguistic and ethical) of an AI tool. *Unisinos Journal of Philosophy*, 25(1), 1-14. <https://doi.org/10.4013/fsu.2024.251.12>
- Wang, F.-Y., Miao, Q., Li, X., Wang, X. y Lin, Y. (2023). What does ChatGPT say: The DAO from algorithmic intelligence to linguistic intelligence. *IEEE/CAA Journal of Automatica Sinica*, 10(3), 575-579. <https://doi.org/10.1109/JAS.2023.123486>
- Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, 10(2), 1-37. <https://doi.org/10.2478/jagi-2019-0002>
- Zamora Bonilla, J. (2021). *Contra apocalípticos: Ecologismo, Animalismo, Posthumanismo*. Shackleton Books.

<https://doi.org/10.17163/abyaups.141.4>

Hacia una perspectiva ética e integral del uso de la inteligencia artificial aplicada a la educación: apuntes para la reflexión

Joe Llerena Izquierdo

Universidad Politécnica Salesiana

jllerena@ups.edu.ec

<https://orcid.org/0000-0001-9907-7048>

Raquel Ayala Carabajo

Universidad Politécnica Salesiana

rayala@ups.edu.ec

<https://orcid.org/0000-0002-5173-099X>

Introducción

Comparar el desempeño de una máquina con el del ser humano en el manejo de información es complejo. La primera afirmación que surge es que, si dicha máquina cuenta con el conocimiento acumulado de varios o muchos seres humanos, tendrá ventajas al procesar información en grandes volúmenes de la misma (Ivcevic y Grandinetti, 2024). Esta mayor capacidad para almacenar conocimiento supondría una ventaja significativa al proporcionar respuestas más rápidas (Malinka *et al.*, 2023). Sin embargo (y como educadores), ¿es posible superar esta capacidad de las máquinas al reflexionar y actuar el conocimiento desde perspectivas humanísticas, altruistas, científicas y/o filantrópicas (*cf.* Loos *et al.*, 2023)?

Más allá de esto, ¿de qué forma podemos integrar, como educadores, estas capacidades en beneficio de los estudiantes, de forma ética, integral y significativa?

El objetivo de este trabajo es explorar, de forma reflexiva y crítica, las implicaciones, fortalezas, riesgos, dificultades y desafíos de esta integración de la IA a la educación, a partir de los hallazgos, conceptualizaciones y reflexiones de autores (implicados en la práctica educativa) en el marco de la revisión de literatura. Este trabajo propone —de forma teórica y general, sin aplicación a nivel/contexto educativo alguno— contenidos originales de reflexión y, en consecuencia, de investigación y de acción docente. En este sentido, se propone analizar cómo el uso de estas tecnologías enriquecen y potencian el proceso educativo, en lugar de simplemente sustituir métodos previos (Zhu *et al.*, 2024). Es decir, se apuesta por una visión donde la tecnología no se erige como una entidad separada, sea tangencial o central, sino como un componente integrado que apoya y amplifica la eficacia y personalización de las estrategias pedagógicas.

Antes de introducirnos en el análisis de estas aportaciones, cabe destacar que consideramos necesario que la discusión actual —que ocurre en los diversos contextos educativos (entre directivos, profesores, estudiantes, padres, etc.)— sobre la interacción entre las máquinas y los seres humanos, debe ir más allá del problema de sustitución de tareas del profesor en contextos específicos de enseñanza-aprendizaje. En efecto, no se trata de dilucidar simplistamente la dicotomía de a quién y cómo reemplaza la IA, sino considerar cómo estos dos agentes pueden colaborar de manera sinérgica para alcanzar logros de aprendizaje significativo en el estudiante y una mejora evidente de la calidad de los procesos educativos (personalización, retención, pertinencia, etc., por solo nombrar algunas de las preocupaciones acuciantes que atraviesan los sistemas educativos hoy (Segrera, 2023).

Una primera reflexión que surge es que, en los contextos educativos particulares y concretos, podemos observar hoy en día la singular interacción entre tecnologías emergentes y prácticas pedagógicas tradicionales; lo

que exige una reconfiguración del conocimiento y práctica de los docentes sobre el aprendizaje y la enseñanza (Ayala-Carabayo y Llerena-Izquierdo, 2024). En este sentido, en cualquier nivel y área, hoy en día los profesores dedican tiempo para aprender y/o usar herramientas tecnológicas para su actividad docente y, también, para desempeñar roles como gestores, tutores, guías y evaluadores, desarrollando nuevas competencias, especialmente si estas herramientas están basadas en la inteligencia artificial (Ng *et al.*, 2023). Ya en el terreno de las habilidades específicas, la incorporación de esta tecnología va desde la adaptación curricular hasta la mejora en la escritura y la comprensión de preguntas complicadas.

En línea con el cuestionamiento del primer párrafo, a diferencia de las máquinas, un profesor puede responder a sus estudiantes (cuando, por ejemplo, obtienen notas deficientes) más allá de la automatización y de la artificialidad, empleando tonos, gestos y expresiones cuidadosamente escogidas para evitar el desánimo; intencionalmente buscando un cambio de actitud y planificando/implementando estrategias y recursos para la mejora de su rendimiento académico. En este sentido, aunque los algoritmos de las máquinas predicen o ‘pronostican’ lo que un estudiante será capaz de hacer o su desempeño futuro según su patrón de comportamiento o sesgos que discriminan cursos de acción (Burton, 2023), no pueden influir positiva ni directamente en la actitud de un ser humano para motivarlo y enriquecerlo en el camino del aprendizaje (Li, 2023).

Consideramos, en ese sentido, que una forma de concebir esta integración es el de una *convergencia* de capacidades y potencialidades que van más allá de la funcionalidad técnica y operativa. Para Halaweh (2023), por ejemplo, y refiriéndose concretamente a esta interacción con la IA a través de ChatGPT, la convergencia radica en la manera en que humanos y máquinas pueden resolver trabajos de alta demanda de tiempo y de búsqueda de información para mejorar la eficiencia. Así también, según Vorobieva (2024), esta colaboración no se limita a la optimización de necesidades tecnológicas, sino que también contempla la capacidad de las máquinas para contribuir a la resolución de problemas complejos

introduciendo un entendimiento ético y contextualizado. La efectividad, desde este punto de vista, se redefiniría no solo en términos de rendimiento, sino valorando la calidad de esta integración y la proyección del impacto positivo en el entorno.

Referentes teóricos

Numerosos estudios abordan las ventajas y usos de la IA en la práctica educativa. En ellos se han planteado enfoques generales acerca de la relación educación-IA y también estudios específicos en contextos de aplicación. En el primer caso, por ejemplo, hay estudios que presentan a la IA como una oportunidad para acelerar la consecución del objetivo del desarrollo sostenible 4 (ODS4) de las Naciones Unidas (Flores-Vivar y García-Peñalvo, 2023; López-Chila *et al.*, 2024; UNESCO, 2019), esto es, desempeñando un rol fundamental y potencial para garantizar una oferta educativa pertinente, accesible, equitativa y sostenible en el tiempo. En el mismo ámbito de la práctica docente, los estudios analizados determinan aspectos clave para tener en cuenta en la incorporación de las IA. Por ejemplo, la temprana alfabetización en su uso está mostrando utilidad como herramientas de apoyo en la gestión académica y a la docencia (Kong *et al.*, 2023; Mohamed, 2024).

En el campo de la educación se han generado propuestas de integración de las herramientas inteligentes como una palanca para la potenciación y/o automatización del proceso de enseñanza y aprendizaje. Sin embargo, persiste la preocupación por un desarrollo de las IA sin 'empatía' o sin la incorporación del agente humano que, con madurez académica, científica, pedagógica y tecnológica incorpore a los algoritmos y a los grandes datos una garantía humanista sin discriminación ni sesgos (Ruíz Aldasoro y Baradat Mari, 2024; Tapia *et al.*, 2023; UNESCO, 2021). Así también, y ya en el contexto aula, para muchos educadores una educación mediada por tecnología les enfrenta a prejuicios y resistencia (propias y ajenas) acerca de su propósito y formas de empleo (Lepri *et al.*, 2021; Osborne *et al.*, 2023). Además, esta conflictividad también se extiende a la

posibilidad de incorporar elementos robóticos o de inteligencia artificial en sus actividades de enseñanza (Corbett-Davies *et al.*, 2024). Algunos autores, consideran como un incremento de carga cognitiva para el docente el esfuerzo de adquirir competencias en herramientas TIC inteligentes (Botwe *et al.*, 2021; Osborne *et al.*, 2023; Reza *et al.*, 2024; Slater, 2024). Sin embargo, puede pensarse también lo contrario: en qué medida puede facilitarse y aligerarse la carga docente en tareas específicas (repetitivas o no) que pueden ser perfectamente transferidas a una herramienta IA. También hay que destacar la relevancia de la discusión sobre la integridad académica y la actitud de los estudiantes hacia el conocimiento con sus patrones de comportamiento al usar IA en sus actividades y tareas de aprendizaje.

Avanzando en la discusión, pueden definirse varias respuestas ante esta posibilidad de integración. En primer lugar, la de aquellos que adoptan las tecnologías con el fin de facilitar o desarrollar (curricular, didáctica, cognitivamente, etc.) el aprendizaje de las ciencias, las matemáticas y las ingenierías en sus programas de estudio o asignaturas y, por tanto, conciben la ayuda de herramientas inteligentes como prometedora (Schiavo *et al.*, 2024; Xie y Wang, 2024). Una segunda postura, es la de aquellos docentes que perciben a las herramientas inteligentes como fuente de información instantánea, de fácil acceso/uso, que pueden llegar a convertir el aprendizaje en un proceso de aprendizaje informativo, pero no reflexivo ni crítico (Alfredo *et al.*, 2024; Belkin, 2024). En tercer lugar, una postura que se puede considerar transversal a las anteriores, es la creación de recursos abiertos generados con las IA, desde una conciencia ética y que sirvan al bien común (García-Peñalvo *et al.*, 2023; Kortemeyer, 2023), adoptando formas colaborativas de trabajo con las IA (Botwe *et al.*, 2021).

Una forma clara de integración —objeto de numerosas investigaciones— es el uso de la IA para identificar características y/o patrones en el desempeño académico o su trayectoria educativa de los estudiantes; validando un clasificador no discriminante para la detección y predicción de situaciones de riesgo, así como, sus componentes cognitivos y emotivos

de actividad en entornos digitales y/o virtuales de aprendizaje, estrategias y tecnologías basadas en la inteligencia artificial. En otro campo de investigación prolífico, se aspira a determinar el efecto de las tecnologías en el desarrollo académico, desempeño y actitud de los estudiantes en la medida y en consonancia con su forma de usar, incorporar y asumir intelectualmente los resultados que requieren/obtienen (por ejemplo, las búsquedas en ChatGPT) de la Inteligencia Artificial. En este punto se cuestionan y/o proponen consideraciones éticas para el aprendizaje de los estudiantes; en efecto, se discuten las características de una experiencia formativa que lleve al alumnado a un aprendizaje autónomo en las actividades de aprendizaje desarrolladas con IA.

Como se adelantó en párrafos anteriores, un aspecto crítico de esta integración es la necesidad y el reto de desarrollar competencias digitales en IA. En el caso de los profesores, la aspiración es la de fortalecer sus habilidades técnicas y críticas para el desarrollo, orientación y retroalimentación en la enseñanza con apoyo de la IA. Junto con estas competencias, un sentido crítico respecto de las implicaciones éticas de la incorporación de la IA, ya que ésta genera preocupación, expectativa y también dilemas sobre su influencia en el aprendizaje de los estudiantes, crea patrones de comportamientos que deben considerarse en las nuevas modalidades de aprendizaje en los contextos educativos (Flogie y Aberšek, 2022). Acercarse a las aplicaciones basadas en IA como un simple ‘consumismo’ de la información y/o del conocimiento, ha sido cuestionado en la medida que podría ralentizar la generación de nuevo conocimiento y reducir las posibilidades del ser humano de tener un protagonismo de creador (Soman y H. G., 2024).

Metodología y resultados

La base de esta propuesta es de una revisión de trabajos relevantes obtenidos de la base de datos Web of Science (WoS). Las búsquedas se generan con la conjugación de los términos clave *artificial intelligence* y *education*. De un total de 63 974 registros se realiza el proceso de selec-

ción a partir de criterios inclusión y exclusión de trabajos relevantes. Los criterios de inclusión se refirieron al tipo de documento (artículo) y de acceso (abierto); trabajos en el idioma inglés y de los tres últimos años. Por los criterios de exclusión se dejaron fuera del conjunto de artículos todos aquellos trabajos que no pertenecen a las categorías de *investigación educativa* y disciplinas científicas de la educación. El resultado fue el de un total de 192 artículos para un análisis de contenido e inclusión por su relevancia. El 85 % de los trabajos provienen de la categoría de investigación educativa y el 15 % “disciplinas científicas de la educación” (según las categorías determinadas de la Web of Science).

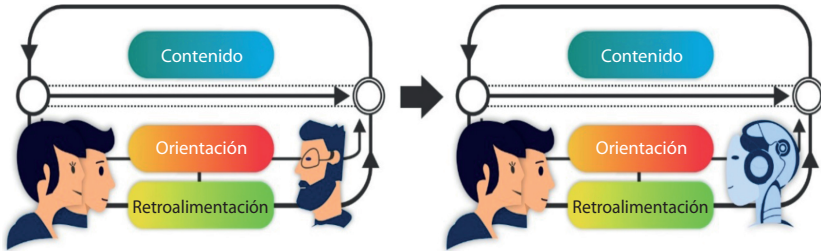
Resultados y hallazgos principales

En primer lugar, cabe destacar el interés generalizado respecto de la transformación de las realidades educativas con la inclusión de robots autónomos y dispositivos con inteligencia artificial (Botwe *et al.*, 2021); así como la necesidad de familiarización y la instrucción teórico-práctica en el ámbito de la tecnología (Nazaretsky *et al.*, 2022). Al mismo tiempo, se estudia cómo el uso educativo de la inteligencia artificial enfrenta barreras de índole diversa (sociales, económicas, tecnológicas, ideológicas, etc.) que limitarían su eficacia (Huh, 2023). Así también, lo tecnológico (indexación, acceso, cobertura, etc.) sigue siendo factor de exclusión/inclusión y, en muchos casos, tiene un trasfondo económico y estaría relacionado con los monopolios y élites (Fallahi *et al.*, 2023).

Al mismo tiempo, se destaca que innovación y mejora educativa no suelen ser resultado de cambios programáticos (Bicknell *et al.*, 2023), sino que está sujeta a las diversas transformaciones tecnológicas y al mismo advenimiento de nuevas tecnologías (Takagi *et al.*, 2023). La integración de la IA en la educación supone una permanente capacitación docente, así como conocer la instrucción personalizada y adaptable a un contexto, el seguimiento continuo al estudiante y la evaluación inteligente de forma masiva con el uso de las IA (Qu *et al.*, 2022) ver Figura 1.

Figura 1

IA mejora la formación docente en contenido, guía y retroalimentación



A pesar de cierto pesimismo o incertidumbre en ciertos contextos educativos, la realidad actual muestra una integración cada vez más extensa de herramientas IA en los contenidos curriculares de la oferta educativa. La IA se presenta como una herramienta que podría transformar el proceso educativo en una experiencia más dinámica y adaptativa, siempre y cuando se utilice con una reflexión crítica y ética. A medida que los educadores se apoyan en la tecnología para mejorar los planes de estudio, proporcionar un seguimiento más personalizado a los estudiantes y ofrecer retroalimentación en plazos cada vez más reducidos (Jia *et al.*, 2024), etc., surge la necesidad de reconsiderar la naturaleza del aprendizaje mismo.

Al mismo tiempo, está viva la discusión sobre el uso ético de la IA. Según Johnston (*et al.*, 2024), existe un riesgo creciente de que los estudiantes, al utilizar IA generativa, lleguen a asimilar los resultados de manera que los consideren propios sin la debida atribución, lo cual podría fomentar un comportamiento deshonesto que resulte imperceptible incluso para la propia conciencia del individuo. Esta situación plantea interrogantes desafiantes sobre cómo la tecnología influye en la ética del aprendizaje y la autenticidad del proceso educativo. Siendo esta la situación concreta más preocupante en contextos educativos, la convergencia entre IA y educación debe ser abordada con una perspectiva ética y filosófica; es preciso contemplar tanto las oportunidades como los riesgos, de manera que se respeten los principios fundamentales del aprendizaje y la integridad.

Otro resultado de este análisis, que recogemos a continuación, es el análisis comparativo entre varias plataformas de inteligencia artificial (IA) aplicadas al ámbito educativo, basado en estudios recientes (año 2023 y 2024) (Tabla 1).

Tabla 1

Comparativas entre plataformas de inteligencia artificial en el ámbito educativo desde una visión de la literatura existente a partir de su objetivo, dificultad de uso, oportunidad que representa y la aceptabilidad del usuario

Año	IA	Objetivo	Dificultades	Oportunidades	Aceptabilidad	Referencia
2023	ChatGPT-4 Plus	Comparar el rendimiento con su predecesora en el ámbito del razonamiento clínico y el conocimiento médico.	La IA requiere del conocimiento, comentarios, participación del ser humano para seguir aprendiendo.	Enfoque de uso más confiable, creativo e idóneo para la comprensión de instrucciones de mayor connotación (hipotetiza).	90 %	(Takagi <i>et al.</i> , 2023) (Mohamed, 2024)
2023	Birdbrain	Aprender del ser humano para personalizar el aprendizaje de otros valorando la mejor experiencia de aprendizaje.	La IA tiene dificultades para mantener la motivación del usuario y el compromiso por el estudio.	Enfoque al aprendizaje de dominio ofreciendo problemas en torno a un concepto similar o componente de conocimiento, requiere de ingeniería inventiva.	90 %	(Bicknell <i>et al.</i> , 2023)
2023	Squirrel AI	Desarrollar rutas de aprendizaje para mejorar una adaptación.	La IA requiere información del estudiante para desarrollar una educación adaptativa inteligente.	Enfoque a un aprendizaje adaptativo en procesos personalizados.	64 %	(Shuai Wang Claire Christensen y Feng, 2023)

2023	Chat-GPT-3.5	Comparar el conocimiento en un temario y la aceptabilidad de las explicaciones con estudiantes de medicina.	La IA no pudo interpretar figuras, gráficos y tablas.	Enfoque escaso a los conocimientos en parasitología y datos epidemiológicos en coreano.	61 %	(Huh, 2023)
2023	Bard (PaLM2) de Google	Ayudar a responder consultas y a entenderlas mediante respuestas de alta calidad, así como explicaciones de situaciones complicadas.	La IA requiere de datos segmentados en tokens como unidades elementales, pero carecen de representaciones.	Enfoque escaso en especificaciones más selectas, que su tendencia en consultas ambiguas se niegue cortésmente a responder.	60 %	(Bal Ram y Pratima Verma, 2023)
2024	Grammarly	Ayudar a la escritura más representativa o convincente de lo que se desea expresar.	La IA aumenta la confianza de los estudiantes en la redacción de ensayos evitando el plagio accidental.	Enfoque de uso de un conjunto extenso de alternativas de respuestas para poner a disposición de otros trabajos del mismo ámbito.	54 %	(Johnston <i>et al.</i> , 2024)
2023	Large Language Model Meta AI (LLaMA)	Comparar con otras IA generativas para establecer la confiabilidad y sesgos existentes.	La IA requiere de datos recientes para mejorar las respuestas en su completitud.	Enfoque a la interpretación de la consulta con un impulso a pedir aclaraciones al usuario de lo que la IA no entiende de la consulta.	50 %	(Soman y H. G., 2024)

En el año 2023, una serie de plataformas de inteligencia artificial se destacaron por sus objetivos, desafíos y las oportunidades que parecían ofrecer en el ámbito educativo. Así, por ejemplo, ChatGPT-4 Plus, con el objetivo de superar a su predecesora en diferentes ámbitos, (como, por

ejemplo en el razonamiento en el área de estudios clínicos y conocimiento médico), enfrenta la dependencia del conocimiento humano para su aprendizaje continuo, pero ofrece oportunidades significativas para la comprensión fiable de instrucciones complejas, destacándose por una aceptabilidad alta del 90 % (Takagi *et al.*, 2023).

Por su parte, Birdbrain aparece en 2023 y se centra en personalizar el aprendizaje basado en la experiencia del usuario (Bicknell *et al.*, 2023). A pesar de los desafíos para mantener la motivación y el compromiso del usuario, Birdbrain destaca por su enfoque en el aprendizaje de dominio con ingeniería inventiva, alcanzando también una alta aceptabilidad del 90%. Así también, Squirrel AI, dedicada a desarrollar rutas de aprendizaje adaptativas basadas en la información detallada que proporciona el estudiante y genera oportunidades mediante la adaptación personalizada de procesos educativos, con una aceptabilidad moderada del 64 % (Shuai Wang Claire Christensen y Feng, 2023). En el mismo año, ChatGPT-3.5 se propuso inicialmente comparar conocimiento y la aceptabilidad entre diferentes usuarios, entre ellos estudiantes y, específicamente, en el área de la medicina (Huh, 2023). Sin embargo, enfrenta limitaciones en la interpretación de figuras y datos específicos, requiriendo mejoras notables en parasitología y epidemiología para elevar su aceptabilidad, la cual es actualmente baja, alcanzando el 61 %.

Por su parte, Bard (PaLM2) de Google, es un modelo que tiene como objetivo responder y entender consultas complejas, pero enfrenta desafíos significativos como la necesidad de datos segmentados y representaciones adecuadas (Bal Ram y Pratima Verma, 2023). A pesar de ello, se destacan oportunidades para mejorar en la respuesta a consultas ambiguas, con una aceptabilidad actual del 60 %. Otro modelo es Large Language Model Meta AI (LLaMA), que se propone comparar la fiabilidad y los sesgos con otras IA generativas. En este sentido, enfrenta desafíos como la necesidad de datos recientes y la clarificación de consultas complejas. No obstante, presenta oportunidades para mejorar la interpretación de dichas consultas, con una aceptabilidad actual baja del 50% (Soman y H.

G., 2024). Lanzado en 2024, Grammarly AI Checker se enfoca en mejorar la calidad y originalidad de la escritura. Aunque depende de un extenso conjunto de respuestas alternativas, ofrece oportunidades para evitar el plagio accidental y mejorar la redacción (Johnston *et al.*, 2024). Su aceptabilidad es moderada, con un 54 %. Estas plataformas representan un espectro diverso de aplicaciones de IA en educación y más allá, cada una con sus propios desafíos y potenciales.

Los desafíos recurrentes incluyen la dependencia de datos humanos para el aprendizaje continuo y la interpretación precisa de consultas complejas. No obstante, se presentan oportunidades claras para optimizar tanto la tecnología como la aceptación de las IA en el ámbito educativo, particularmente en áreas especializadas como la medicina, el aprendizaje de un segundo idioma, la ingeniería y la escritura académica, entre otras. Las inteligencias artificiales, especialmente las generativas, han superado las expectativas de todos (Malinka *et al.*, 2023). Se advierte que para 2025, el 10 % de la información será desarrollada por las IA, especialmente las generativas (Kamran *et al.*, 2020). Aunque se trabaja por determinar los sesgos apropiados para diversidad de problemas reales existentes, sigue siendo un factor predominante alcanzar un adecuado resultado que refleje respuestas convincentes (Thoppilan *et al.*, 2022). La forma compleja y variada del sentir del ser humano es un desafío por alcanzar por las IA. Los resultados evidencian que hay un camino que se está forjando para las IA y ellas mismas están respondiendo. Pero las limitaciones existentes como la complejidad de ser un ser humano, la infraestructura tecnológica en diversos contextos, así como las preocupaciones éticas relativas a la seguridad, la privacidad, los sesgos, tiempos de respuestas o banco de conocimientos disponibles y confiables, entre otros serán los retos que se deberán observar atenta y cautelosamente por la comunidad científica (Ng *et al.*, 2023).

En otro sentido, es fundamental reconocer tanto los aspectos positivos como las limitaciones de la IA en el ámbito educativo. Por un lado, la IA muestra un potencial considerable para transformar y superar diversos aspectos del proceso educativo, ofreciendo herramientas innovadoras que

mejoran la personalización del aprendizaje y la accesibilidad a recursos. Al mismo tiempo, y a pesar de estos avances, la presencialidad aún mantiene ventajas significativas sobre lo digital, que la inteligencia artificial no ha logrado replicar completamente (Gouvea, 2024). En la siguiente tabla, se recogen una síntesis de estos aspectos, según autores varios, en los dos últimos años:

Tabla 2

Fortalezas y ventajas versus riesgos, dificultades y desafíos de la IA en la educación hallados en la literatura relevante

Fortalezas y ventajas	Riesgos, dificultades y desafíos	Referencia
<p>La IA integrada en la educación promete la gestión personalizada y la globalización del aprendizaje. La personalización permite a los estudiantes recibir un aprendizaje adaptado a sus necesidades individuales, mientras que la globalización amplía el acceso a recursos educativos a nivel mundial.</p>	<p>La implementación de la inteligencia artificial en la educación requiere una atención cuidadosa en sus desafíos éticos y usos. Es importante abordar cuestiones como la privacidad, la equidad y la transparencia para garantizar que su uso no comprometa los principios fundamentales de justicia y seguridad en el entorno educativo.</p>	<p>(López-Chila <i>et al.</i>, 2024; Pham <i>et al.</i>, 2023)</p>
<p>Utilizar tecnologías novedosas como la IA y la analítica del aprendizaje (LA) ofrecen nuevas oportunidades para mejorar y personalizar los métodos de evaluación, adaptándolos mejor a las necesidades individuales de los estudiantes. Además, el papel del profesor sigue siendo esencial en el proceso educativo. Al enfatizar que la IA no busca reemplazar al profesor, sino fortalecer su rol. Se subraya la importancia de la intervención humana en la educación y se evita una percepción errónea de que la tecnología sustituya la función crítica del educador.</p>	<p>El tránsito hacia métodos de evaluación más flexibles y tecnológicos enfrenta varios desafíos significativos. La transición complica la implementación de métodos de evaluación tradicionales, dado que estos enfoques convencionales no siempre se alinean bien con las nuevas tecnologías. Además, este cambio requiere una adaptación sustancial tanto a las nuevas expectativas del mercado educativo como a la capacitación de los docentes y a la preparación de los estudiantes. Es importante abordar estos aspectos que aseguren una integración efectiva y equitativa de las nuevas metodologías en el sistema educativo.</p>	<p>(Chen <i>et al.</i>, 2023; Gouvea, 2024; Williams, 2023)</p>
<p>El impacto creciente de la inteligencia artificial (IA) generativa en la educación subraya la importancia de adoptar un enfoque proactivo para comprender y adaptarse a estas nuevas tecnologías. Es necesario desglosar y explicar los conceptos relacionados con la IA.</p>	<p>Actualmente, se necesitan soluciones prácticas y recomendaciones específicas para abordar los desafíos que plantea la inteligencia artificial generativa. Dada la rapidez con la que evoluciona esta tecnología, es esencial desarrollar estrategias basadas en evidencia para</p>	<p>(Lodge <i>et al.</i>, 2023; Nguyen Thanh <i>et al.</i>, 2023)</p>

<p>generativa, especialmente para aquellos que no están familiarizados con el lenguaje técnico. Esta claridad facilita a educadores e investigadores comprender sobre las capacidades y limitaciones de la IA en áreas de la educación, promoviendo así un uso más reflexivo de estas tecnologías avanzadas.</p>	<p>mitigar incertidumbres y adaptar las prácticas existentes. La constante evolución del campo requiere una actualización continua de las directrices y la integración de nuevos hallazgos científicos.</p>	
<p>La adopción efectiva de tecnologías educativas basadas en inteligencia artificial (IA) obedece a la confianza del profesorado en el uso de estas herramientas. Es vital enfocar los esfuerzos en comprender y abordar las actitudes y la confianza de los docentes, ya que estos factores favorecen en la implementación exitosa de la IA en el entorno educativo. Superar las barreras psicológicas y emocionales que dificultan la integración de tecnologías avanzadas en el aula es esencial para maximizar su impacto y eficacia.</p>	<p>La investigación existente sobre la confianza de los profesores en tecnologías educativas basadas en IA es escasa, lo que dificulta la comprensión general del problema y la comparación con estudios previos. Esta falta de antecedentes amplios hace que los hallazgos sean menos contextuales y aplicables a un panorama educativo más amplio.</p>	<p>(Lan & Chen, 2024; Nazaretsky <i>et al.</i>, 2022)</p>
<p>La aplicación innovadora de IA proporciona recursos educativos personalizados y adaptados a las carencias de los estudiantes, fomentando un escenario de aprendizaje basado en el descubrimiento y la experimentación. La integración de IA en la creación de recursos educativos enriquece el proceso de enseñanza y aprendizaje, ofreciendo nuevas oportunidades para el desarrollo profesional de los estudiantes.</p>	<p>La implementación de inteligencia artificial conlleva riesgos significativos de sesgo y abuso, derivados de errores ocultos en los algoritmos cuando se ejecutan en infraestructuras de computación de alto rendimiento. La falta de una discusión sobre las estrategias para abordar y mitigar estos riesgos genera desafíos en la seguridad y la ética al tomar decisiones de integrar esta tecnología.</p> <p>Es esencial desarrollar mecanismos claros para asegurar que la IA se utilice de manera responsable y efectiva en el contexto educativo, garantizando tanto la integridad de los sistemas como su impacto positivo en el aprendizaje.</p>	<p>(Gillani <i>et al.</i>, 2023; Jung, 2024; Kahvazadeh <i>et al.</i>, 2024)</p>
<p>Integrar la IA en la educación influye en la práctica de la evaluación formativa, así como también afecta las habilidades de autogestión y la autoeficacia de los estudiantes. Además, la IA permite identificar diferencias individuales en propio monitoreo de los estudiantes, ofreciendo nuevas perspectivas sobre</p>	<p>Evaluar el impacto de la IA en la práctica educativa es necesaria, pero las herramientas actuales adolecen de acciones significativas como la ausencia de recomendaciones concretas y estrategias específicas para su implementación efectiva en contextos educativos</p>	<p>(Ghaleb y Al-shiha, 2023; Hopfenbeck <i>et al.</i>, 2023; Wang <i>et al.</i>, 2024)</p>

<p>cómo los estudiantes manejan su propio aprendizaje. Este impacto dual subraya la necesidad de explorar cómo las herramientas de IA son utilizadas para mejorar tanto la evaluación formativa como el desarrollo de habilidades asociadas en los estudiantes.</p>	<p>reales. Es esencial contar con una guía detallada sobre cómo integrar y utilizar la IA de manera efectiva en entornos educativos específicos. Sin estas directrices claras, se corre el riesgo de que la implementación de la IA se base en enfoques generales y no en soluciones adaptadas a las necesidades particulares de cada entorno educativo, lo que limita su eficacia y sostenibilidad.</p>	
<p>Integrar la IA en la educación ofrece numerosos beneficios al modernizar métodos de enseñanza obsoletos y facilitar un cambio de paradigma en la educación. Las herramientas innovadoras basadas en IA tienen el potencial de superar desafíos existentes como la falta de recursos y mejorar la personalización del aprendizaje. Los educadores son en general optimistas sobre la capacidad de la IA para ofrecer mejoras a la enseñanza, su calidad y aprendizaje, apreciando las oportunidades que brinda para optimizar la retroalimentación y el análisis de datos educativos.</p>	<p>Las preocupaciones éticas incluyen la dependencia excesiva en la tecnología, la posible falta de originalidad en el trabajo académico y la privacidad de los datos. Además, aunque la percepción del uso de las herramientas de IA es positiva, varios autores no abordan en profundidad cómo superar las limitaciones específicas de las herramientas, como errores en la generación de contenido o sesgos en la información, que afectan la calidad del aprendizaje.</p>	<p>(Das & Madhusudan, 2024; Famaye et al., 2024; Shamsuddinova et al., 2024)</p>
<p>La incorporación de elementos de IA facilita el desarrollo de destrezas como la alfabetización en IA y la ingeniería de prompts (preguntas o instrucciones), esenciales para maximizar el potencial de estas tecnologías. Estos elementos no solo enriquecen las experiencias educativas, sino que también fomentan el pensamiento crítico al permitir a los estudiantes relacionarse en procesos lógicos computacionales con los sistemas de IA. Además, la integración de la IA permite escenarios de adaptabilidad de los currículos y prácticas pedagógicas para incluir estas habilidades transformando significativamente los métodos tradicionales de enseñanza</p>	<p>Implementar la IA en el ámbito educativo contrapone varios escenarios desafiantes, entre ellos la necesidad de una formación exhaustiva para los educadores y la adaptación de los currículos a las nuevas tecnologías. La alfabetización en IA es urgente, pero es difícil de alcanzar sin el apoyo adecuado, lo que podría limitar el impacto positivo de estas herramientas. La ingeniería de <i>prompts</i> (preguntas o instrucciones), aunque valiosa, requiere una comprensión profunda para evitar malentendidos o mal uso, lo cual complica su integración efectiva en el aula. Además, la adaptación de nuevas tecnologías enfrenta resistencia y obstáculos estructurales dentro de los sistemas educativos existentes, lo que se advierte una ralentización en la adopción y limitar sus beneficios potenciales.</p>	<p>(García-Peñalvo et al., 2023; Kelly et al., 2023; Walter, 2024)</p>

<p>Utilizar IA y sus herramientas generativas ofrece el beneficio de enriquecer la enseñanza en la educación superior al facilitar el acceso a información y recursos que complementan la formación de los estudiantes. Sin embargo, esta misma capacidad plantea el riesgo de comprometer la integridad académica, ya que su uso no autorizado resulta en plagio y desvalorización del trabajo académico. Esto enfatiza la urgencia de que las instituciones educativas revisen y actualicen sus políticas de integridad académica para enfrentar estos nuevos desafíos.</p>	<p>El principal peligro radica en la falta de claridad y efectividad de las políticas de integridad académica en muchas instituciones. Las políticas institucionales carecen de elementos esenciales para abordar el uso indebido de herramientas generativas, esto genera incertidumbre tanto en estudiantes como en educadores. Sin actualizaciones periódicas y una comunicación clara sobre políticas enfocadas al uso de las IA, se corre el riesgo de comprometer la integridad académica, afectando negativamente la calidad educativa y la reputación de las instituciones.</p>	<p>(De Maio, 2024; Farrow, 2023; Forero-Corba y Negré Bennisar, 2024)</p>
<p>Utilizar IA e integrar las ciencias computacionales en el ámbito educativo presentan el beneficio significativo de mejorar el rendimiento estudiantil en un aumento en la motivación y la actitud hacia el aprendizaje, especialmente en áreas STEM. La integración de estas tecnologías en el aula enriquece la experiencia educativa, proporcionando herramientas que fomentan un aprendizaje más interactivo y efectivo, lo que se traduce en un mejor desempeño académico en diversas etapas educativas.</p>	<p>Este avance también conlleva peligros importantes, principalmente relacionados con los desafíos educativos y éticos que enfrenta el profesorado en el diseño e implementación de estas tecnologías. La incorporación de IA en los procesos de enseñanza requiere un análisis cuidadoso para asegurar que se utilice de manera efectiva y ética, lo que implica un compromiso significativo por parte de los educadores. Si no se abordan estos desafíos, existe el riesgo de que la implementación de la IA no logre los resultados esperados o, incluso, de que afecte negativamente la calidad de la educación.</p>	<p>(Commission y Directorate-General for Education Sport and Culture, 2022; Martínez <i>et al.</i>, 2023)</p>

Conclusiones

La interacción humana, la capacidad de adaptarse a contextos específicos y las dimensiones emocionales del aprendizaje, siguen siendo áreas donde la presencia física de un educador tiene una eficacia superior. La inteligencia artificial presenta oportunidades valiosas, pero no se debe subestimar el valor insustituible de la presencialidad en el desarrollo integral del estudiante. Se evidencia que el impacto de integrar la IA en el ámbito educativo ofrece una serie de aspectos significativos que enriquecen el proceso de enseñanza y aprendizaje.

Precisamente, entre los aspectos positivos que produce la IA, es la capacidad de generar trayectorias de aprendizaje personalizables, adaptativas a las carencias individuales que tienen los estudiantes, en su desempeño, desde un contexto basado en sus patrones de comportamiento. Esta personalización maximiza el rendimiento académico al proporcionar experiencias educativas ajustadas a cada estudiante, mientras que la globalización del acceso a recursos educativos democratizando el aprendizaje, permitiendo que estudiantes de todo el mundo accedan a materiales y oportunidades educativas previamente inaccesibles.

Además, el uso de tecnologías emergentes como la IA y la analítica del aprendizaje (LA) presenta nuevas oportunidades para la mejora de los métodos de evaluación. Estos enfoques permiten adaptar las evaluaciones a las necesidades específicas de cada estudiante y brindan retroalimentación oportuna. La IA, en lugar de reemplazar a los profesores, tiene el potencial de fortalecer su rol al apoyar y enriquecer la enseñanza. Esta perspectiva recalca que el papel del educador sigue siendo fundamental, evitando malentendidos sobre la sustitución de la intervención humana por la tecnología.

Pese a que, el establecimiento de IA en el ámbito educativo también plantea oportunidades significativas. Uno de los principales desafíos es la necesidad de abordar cuestiones éticas, como la reserva de los datos, la igualdad y la transparencia. La introducción de IA en el entorno educativo debe ser manejada cuidadosamente para garantizar que no se comprometan los principios fundamentales de justicia y seguridad. Además, la transición hacia métodos de evaluación más flexibles y tecnológicos complica la implementación de los métodos tradicionales de evaluación. Esta transición requiere una adaptación significativa en la preparación del profesorado y en la formación del estudiantado, así como una constante actualización de las directrices para asegurar una integración efectiva y equitativa de las nuevas metodologías.

Otro desafío importante es la necesidad de soluciones prácticas y recomendaciones específicas para enfrentar los problemas que surgen con la IA generativa. La rapidez con la que evoluciona esta tecnología exige

estrategias basadas en evidencia para mitigar incertidumbres y adaptar las prácticas educativas existentes. La falta de investigación sobre la confianza de los profesores en tecnologías educativas basadas en IA también limita la comprensión general del problema y la comparación con estudios previos, lo que reduce la aplicabilidad de los hallazgos a contextos educativos más amplios.

Al mismo tiempo, el empleo innovador de la IA en el desarrollo y generación de recursos educativos digitales da la posibilidad de enriquecer el proceso de enseñanza y aprendizaje al entregar recursos personalizados y adaptativos. A pesar de esto, la implementación de estas tecnologías conlleva riesgos de sesgo y abuso derivados de errores en los algoritmos, no patentes. Es esencial desarrollar mecanismos claros para abordar y mitigar estos riesgos, asegurando así que la incorporación de la IA se utilice de manera consciente y efectiva en el contexto educativo, garantizando la integridad de los sistemas y el impacto positivo en el aprendizaje. Esto, a su vez, nos cuestiona sobre el significado de la responsabilidad y la efectividad del trabajo conjunto y/o basado en la IA. Así, por ejemplo, en lugar de considerar a las máquinas como simples herramientas o sustitutos de la capacidad humana, pueden concebirse como extensiones de la mente humana (singular o en red) en la creación de soluciones innovadoras.

En este sentido, la colaboración efectiva entre humanos y máquinas debe ser entendida como una oportunidad para redefinir y mejorar las prácticas educativas, promoviendo un aprendizaje más dinámico y adaptativo que responda a las demandas y desafíos actuales. De modo paralelo, se ha reflexionado sobre el enunciado de una constante amenaza de ‘reemplazo’ del profesor por parte de la IA/tecnologías basadas en IA, que ya es parte del imaginario o de los comentarios fáciles cuando de estos temas se trata en los claustros docentes; en especial si se considera que el antídoto está y estará en la apelación a la imposibilidad e improbabilidad de que llegemos a sustituir las relaciones interpersonales.

Nuestra experiencia cotidiana como usuarios, como aprendices y como docente, en un contexto de rápida integración de la IA en las he-

ramientas y recursos con los que trabajamos, nos invita a una reflexión más profunda sobre cómo la IA redefine no solo el contexto y la práctica educativa; sino, además, la relación entre información, conocimiento y aprendizaje. Como hemos analizado, esta transformación tecnológica basada en el desarrollo de la IA no está exenta de importantes desafíos y problemas. La proliferación exponencial de IA generativa en la actualidad ha suscitado preocupaciones sobre la posibilidad de que los usuarios adopten patrones de conducta que socaven principios fundamentales de honestidad intelectual.

La formación de estudiantes conscientes del impacto de la tecnología en su aprendizaje y comportamiento ético es fundamental para garantizar que la IA sea una fuerza positiva en la educación. Al hacerlo, se podrá aspirar a una evolución del modo de concebir y diseñar entornos de aprendizaje que sean tanto innovadores como éticos y humanistas.

Referencias bibliográficas

- Alfredo, R., Echeverría, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D. y Martínez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence*, 6, 100215. <https://doi.org/10.1016/j.caeai.2024.100215>
- Ayala-Carabajo, R y Llerena-Izquierdo, J. (2024). Bibliometric Review on a Hybrid Learning Model with VLEs from a Higher Education Context. En J. P. Salgado-Guerrero, H. R. Vega-Carrillo, G. García-Fernández y V. Robles-Bykbaev (Eds.), *Systems, Smart Technologies and Innovation for Society* (pp. 148-157). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-52090-7_15
- Belkin, M. (2024). The Necessity of Machine Learning Theory in Mitigating AI Risk. *ACM / IMS J. Data Sci.*, 1(3). <https://doi.org/10.1145/3643694>
- Bicknell, K., Brust, C. y Settles, B. (2023). How Duolingo's AI Learns what you Need to Learn: The language-learning app tries to emulate a great human tutor. *IEEE Spectrum*, 60(3), 28-33. <https://doi.org/10.1109/MSPEC.2023.10061631>

- Botwe, B., Antwi, W., Arkoh, S. y Akudjedu, T. (2021). Radiographers' perspectives on the emerging integration of artificial intelligence into diagnostic imaging: The Ghana study. *Journal of Medical Radiation Sciences*, 68(3), 260-268. <https://doi.org/https://doi.org/10.1002/jmrs.460>
- Burton, J. (2023). Algorithmic extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization and political violence. *Technology in Society*, 75, 102262. <https://doi.org/https://doi.org/10.1016/j.techsoc.2023.102262>
- Chen, X., Cheng, G., Zou, D., Zhong, B. y Xie, H. (2023). Artificial Intelligent Robots for Precision Education. *Educational Technology & Society*, 26(1), 171-186. <https://www.jstor.org/stable/48707975>
- Corbett-Davies, S., Gaebler, J. D., Nilforoshan, H., Shroff, R. y Goel, S. (2024). The measure and mismeasure of fairness. *J. Mach. Learn. Res.*, 24(1).
- Das, S. R. y Madhusudan, J. V. (2024). Perceptions of Higher Education Students towards ChatGPT Usage. *International Journal of Technology in Education*, 7(1), 86-106.
- De Maio, C. (2024). Institutional responses to ChatGPT: Analysing the academic integrity policies of four public and private institutions of higher education in Australia. *Journal of Academic Language and Learning*, 18(1 SE-Responding to the opportunities and challenges of generative AI), T1-T8. <https://journal.aall.org.au/index.php/jall/article/view/917>
- Fallahi, S., Mellquist, A.-C., Mogren, O., Listo Zec, E., Algurén, P. y Hallquist, L. (2023). Financing solutions for circular business models: Exploring the role of business ecosystems and artificial intelligence. *Business Strategy and the Environment*, 32(6), 3233-3248. <https://doi.org/10.1002/bse.3297>
- Famaye, T., Bailey, C. S., Adisa, I. y Irgens, G. A. (2024). What Makes ChatGPT Dangerous Is Also What Makes It Special: High-School Student Perspectives on the Integration or Ban of Artificial Intelligence in Educational Contexts. *International Journal of Technology in Education*, 7(2), 174-199.
- Farrow, R. (2023). The possibilities and limits of XAI in education: a socio-technical perspective. *Learning, Media and Technology*, 48(2), 266-279. <https://doi.org/10.1080/17439884.2023.2185630>
- Flogie, A. y Aberšek, B. (2022). Artificial Intelligence in Education. In *Active Learning Theory and Practice*. IntechOpen. <https://doi.org/10.5772/intechopen.96498>
- Flores-Vivar, J. M. y García-Peñalvo, F. J. (2023). Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality

- education (SDG4). *Comunicar*, 31(74), 37-47. <https://doi.org/10.3916/C74-2023-03>
- Forero-Corba, W. y Negre Bennasar, F. (2024). Técnicas y aplicaciones del Machine Learning e inteligencia artificial en educación: una revisión sistemática. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1 SE-Estudios e investigaciones), 209-253. <https://doi.org/10.5944/ried.27.1.37491>
- García-Peñalvo, F., Llorens Largo, F. y Vidal, J. (2023). *The new reality of education in the face of advances in generative artificial intelligence*. <https://doi.org/10.5944/ried.27.1.37716>
- Ghaleb, M. y Alshiha, A. (2023). Empowering Self-Management: Unveiling the Impact of Artificial Intelligence in Learning on Student Self-Efficacy and Self-Monitoring. *Eurasian Journal of Educational Research*, 2023(107), 68-94.
- Gillani, N., Eynon, R., Chiabaut, C. y Finkel, K. (2023). Unpacking the “Black Box” of AI in Education. *Educational Technology & Society*, 26(1), 99-111. <https://www.jstor.org/stable/48707970>
- Gouvea, J. S. (2024). Ethical Dilemmas in Current Uses of AI in Science Education. *CBE—Life Sciences Education*, 23(1), fe3. <https://doi.org/10.1187/cbe.23-12-0239>
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15(2), ep421. <https://doi.org/10.30935/cedtech/13036>
- Hopfenbeck, T., Zhang, Z., Sun, S., Robertson, P. y McGrane, J. (2023). Challenges and opportunities for classroom-based formative assessment and AI: a perspective article. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1270700>
- Huh, S. (2023). Are ChatGPT’s knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study. *Journal of Educational Evaluation for Health Professions*, 20(1), 1. <https://doi.org/10.3352/jeehp.2023.20.1>
- Ivcevic, Z. y Grandinetti, M. (2024). Artificial intelligence as a tool for creativity. *Journal of Creativity*, 34(2), 100079. <https://doi.org/10.1016/j.yjoc.2024.100079>
- Jia, F., Sun, D. y Looi, C. (2024). Artificial Intelligence in Science Education (2013-2023): Research Trends in Ten Years. *Journal of Science Education and Technology*, 33(1), 94-117. <https://doi.org/10.1007/s10956-023-10077-6>
- Johnston, H., Wells, R., Shanks, E., Boey, T. y Parsons, B. (2024). Student perspectives on the use of generative artificial intelligence technologies in

- higher education. *International Journal for Educational Integrity*, 20(1), 2. <https://doi.org/10.1007/s40979-024-00149-4>
- Jung, I. (2024). Personalized education for all: The future of open universities. *Open Praxis*, 16(1), 24-36. <https://doi.org/10.55982/openpraxis.16.1.612>
- Kahvazadeh, I., Jose, E., Fong, A., Gupta, A., Carr, S., Bhattacharjee, S. y Harnar, M. (2024). Development and evaluation of a modular experiential learning curriculum for promoting AI readiness. *Education and Information Technologies*, 29(3), 3445-3459. <https://doi.org/10.1007/s10639-023-11928-w>
- Kamran, S., Farhat, I., Talha Mahboob, A., Gagandeep Kaur, A., Liton, D., Abdul Ghaffar, K., Rimsha, I., Irum, S. y Afifah, R. (2020). The Impact of Artificial intelligence and Robotics on the Future Employment Opportunities. *Trends in Computer Science and Information Technology*, 050-054. <https://doi.org/10.17352/tcsit.000022>
- Kelly, A., Sullivan, M. y Strampel, K. (2023). Generative artificial intelligence: University student awareness, experience, and confidence in use across disciplines. *Journal of University Teaching and Learning Practice*, 20(6). <https://doi.org/10.53761/1.20.6.12>
- Kong, S.-C., Cheung, W. M.-Y. y Zhang, G. (2023). Evaluating an artificial intelligence literacy programme for developing university students' conceptual understanding, literacy, empowerment and ethical awareness. *Educational Technology & Society*, 26(1), 16-30.
- Kortemeyer, G. (2023). Could an artificial-intelligence agent pass an introductory physics course? *Physical Review Physics Education Research*, 19(1), 10132.
- Lan, Y.-J. y Chen, N.-S. (2024). Teachers' agency in the era of LLM and generative AI. *Educational Technology & Society*, 27(1), I-XVIII. <https://bit.ly/3VHhR19>
- Lepri, B., Oliver, N. y Pentland, A. (2021). Ethical machines: The human-centric use of artificial intelligence. *IScience*, 24(3), 102249. <https://doi.org/10.1016/j.isci.2021.102249>
- Li, H. (2023). Rethinking human excellence in the AI age: The relationship between intellectual humility and attitudes toward ChatGPT. *Personality and Individual Differences*, 215, 112401. <https://doi.org/https://doi.org/10.1016/j.paid.2023.112401>
- Lodge, J. M., Thompson, K. y Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*, 39(1 SE-Editorial), 1-8. <https://doi.org/10.14742/ajet.8695>

- Loos, E., Gröpler, J. y Goudeau, M. (2023). Using ChatGPT in Education: Human Reflection on ChatGPT's Self-Reflection. *Societies*, 13(8). <https://doi.org/10.3390/soc13080196>
- López-Chila, R., Llerena-Izquierdo, J., Sumba-Nacipucha, N. y Cueva-Estrada, J. (2024). Artificial Intelligence in Higher Education: An Analysis of Existing Bibliometrics. *Education Sciences*, 14(1). <https://doi.org/10.3390/educsci14010047>
- Malinka, K., Peresíni, M., Firc, A., Hujnák, O. y Janus, F. (2023). On the Educational Impact of ChatGPT: Is Artificial Intelligence Ready to Obtain a University Degree? *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, 47-53. <https://doi.org/10.1145/3587102.3588827>
- Martínez, I., Batanero, J., Cerero, J. F. y León, S. (2023). Analysing the impact of artificial intelligence and computational sciences on student performance: systematic review and meta-analysis. *NAER: Journal of New Approaches in Educational Research*, 12(1), 171-197.
- Mohamed, A. (2024). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL Faculty Members. *Education and Information Technologies*, 29(3), 3195-3217. <https://doi.org/10.1007/s10639-023-11917-z>
- Nazaretsky, T., Ariely, M., Cukurova, M. y Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, 53(4), 914-931. <https://doi.org/https://doi.org/10.1111/bjet.13232>
- Ng, D., Leung, J., Su, J., Ng, R. y Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research and Development*, 71(1), 137-161. <https://doi.org/10.1007/s11423-023-10203-6>
- Nguyen Thanh, B., Vo, D., Nguyen Nhat, M., Pham, T., Thai Trung, H. y Ha Xuan, S. (2023). Race with the machines: Assessing the capability of generative AI in solving authentic assessments. *Australasian Journal of Educational Technology*, 39 (5 SE-Themed issue 2023-AI in tertiary education: impacts for research and practice), 59-81. <https://doi.org/10.14742/ajet.8902>
- Osborne, M., Omrani, A. y Deghani, M. (2023). The Sins of the Parents Are to Be Laid Upon the Children: Biased Humans, Biased Data, Biased Models. *Perspectives on Psychological Science*, 17456916231180100. <https://doi.org/10.1177/17456916231180099>

- Pham, T., Nguyen, T., Ha, S. y Nguyen Ngoc, N. (2023). Digital transformation in engineering education: Exploring the potential of AI-assisted learning. *Australasian Journal of Educational Technology*, 39(5 SE-Themed issue 2023-AI in tertiary education: impacts for research and practice), 1-19. <https://doi.org/10.14742/ajet.8825>
- Qu, J., Zhao, Y. y Xie, Y. (2022). Artificial intelligence leads the reform of education models. *Systems Research and Behavioral Science*, 39(3), 581-588. <https://doi.org/https://doi.org/10.1002/sres.2864>
- Reza, M., Tasnim, M., Afroz, M. y Ruhi, S. (2024). Exploring nomophobia among university students: Identifying risk factors, correlates, and predictive insights through machine learning. *Journal of Future Sustainability*, 4(4), 243-250. <https://doi.org/10.5267/j.jfs.2024.11.001>
- Ruiz, P. y Baradat, R. (2024). *Metodología para la elaboración de una cartografía de inteligencia artificial responsable*. <https://bit.ly/4pXfXY7>
- Schiavo, G., Businaro, S. y Zancanaro, M. (2024). Comprehension, apprehension, and acceptance: Understanding the influence of literacy and anxiety on acceptance of artificial Intelligence. *Technology in Society*, 77, 102537. <https://doi.org/https://doi.org/10.1016/j.techsoc.2024.102537>
- Segrera, F. L. (2023). Posibles futuros de la educación superior en América Latina y el Caribe: antecedentes, situación actual, escenarios y alternativas. *Revista Educación Superior y Sociedad (ESS)*, 35(1 SE-Artículos Dossier Temático). <https://doi.org/10.54674/ess.v35i1.856>
- Shamsuddinova, S., Heryani, P. y Naval, M. (2024). Evolution to revolution: Critical exploration of educators' perceptions of the impact of Artificial Intelligence (AI) on the teaching and learning process in the GCC region. *International Journal of Educational Research*, 125, 102326. <https://doi.org/https://doi.org/10.1016/j.ijer.2024.102326>
- Wang, S., Christensen, C., Cui, W., Tong, R., Yarnall, L., Shear, L. y Feng, M. (2020). When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction. *Interactive Learning Environments*, 31(2), 793-803. <https://doi.org/10.1080/10494820.2020.1808794>
- Slater, G. B. (2024). Dread and the automation of education: From algorithmic anxiety to a new sensibility. *Review of Education, Pedagogy, and Cultural Studies*, 46(1), 170-182. <https://doi.org/10.1080/10714413.2023.2299521>
- Soman, S. y Ranjani, H. G. (2024). Observations on LLMs for Telecom Domain: Capabilities and Limitations. *Proceedings of the Third International Conference on AI-ML Systems*. <https://doi.org/10.1145/3639856.3639892> <https://bit.ly/48GNBLg>

- Takagi, S., Watari, T., Erabi, A. y Sakaguchi, K. (2023). Performance of GPT-3.5 and GPT-4 on the Japanese Medical Licensing Examination: Comparison Study. *JMIR Medical Education*, 9, e48002. <https://doi.org/10.2196/48002>
- Tapia, J., Avilés, F., Cuesta, D. y García, J. Z. (2023). Digital Transformation in Higher Education: Proposed Initiatives from the Salesian Polytechnic University Assessment. *2023 IEEE Seventh Ecuador Technical Chapters Meeting (ECTM)*, 1-6. <https://doi.org/10.1109/ETCM58927.2023.10309007>
- Thoppilan, R., Freitas, D. De, Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.-T., Jin, A., Bos, T., Baker, L., Du, Y. et al. (2022). *LaMDA: Language Models for Dialog Applications*. <https://bit.ly/4nwWqk>
- UNESCO. (2019). Artificial Intelligence in education: Challenges and opportunities for sustainable development. *Unesco Working Papers on Education Policy*. <https://bit.ly/4nWOeop>
- UNESCO. (2021). *Recomendación sobre la ética de la Inteligencia Artificial*. <https://bit.ly/3WIRJZG>
- Vorobieva, N. (2024). The Mechanisms of Artificial Intelligence in The Process of Teaching English: Risks, Opportunities, Prospects. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4773358>
- Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: the relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, 21(1), 15. <https://doi.org/10.1186/s41239-024-00448-3>
- Wang, D., Tao, Y. y Chen, G. (2024). Artificial intelligence in classroom discourse: A systematic review of the past decade. *International Journal of Educational Research*, 123, 102275. <https://doi.org/https://doi.org/10.1016/j.ijer.2023.102275>
- Williams, P. (2023). AI, Analytics and a New Assessment Model for Universities. *Education Sciences*, 13(10), 1040. <https://doi.org/10.3390/educsci13101040>
- Xie, X. y Wang, T. (2024). Artificial Intelligence: A help or threat to contemporary education. Should students be forced to think and do their tasks independently? *Education and Information Technologies*, 29(3), 3097-3111. <https://doi.org/10.1007/s10639-023-11947-7>
- Zhu, S., Wang, Z., Zhuang, Y., Jiang, Y., Guo, M., Zhang, X. y Gao, Z. (2024). Exploring the impact of ChatGPT on art creation and collaboration: Benefits, challenges and ethical implications. *Telematics and Informatics Reports*, 14, 100138. <https://doi.org/https://doi.org/10.1016/j.teler.2024.100138>

Inteligencia artificial y ética del lenguaje: del atomismo lógico a las redes neuronales artificiales

Sara Eugenia Madera Gómez

Universidad de Milán

sarymadera@hotmail.com

<https://orcid.org/0000-0003-0164-4695>

Introducción

La inteligencia artificial (IA) ha revolucionado numerosos campos, desde la medicina hasta la economía, planteando simultáneamente nuevos y complejos desafíos éticos. En este artículo, defiendo una tesis central: los problemas éticos fundamentales asociados con la IA generativa no son primariamente un problema de algoritmos, sino que derivan directamente de los conjuntos de datos (datasets) utilizados para su entrenamiento. Más aún, sostengo que esto es, en su esencia, un problema del lenguaje.

Mi argumento se basa en una analogía clave: así como el atomismo lógico de Bertrand Russell propone que el lenguaje debe reflejar la estructura de los hechos que describe, los datasets utilizados en el entrenamiento de IA funcionan como un ‘lenguaje’ que intenta representar la realidad. Sin embargo, al igual que el lenguaje puede ser impreciso o sesgado, los datasets pueden contener errores, sesgos y limitaciones que se transmiten directamente a los modelos de IA, generando consecuencias éticas significativas.

Esta perspectiva nos permite reinterpretar los desafíos éticos de la IA no como problemas técnicos aislados, sino como manifestaciones de las limitaciones inherentes a nuestros intentos de representar la realidad a través del lenguaje; en este caso, el lenguaje de los datos. Al reinterpretar los desafíos éticos de la IA a través de esta lente filosófica, este artículo no solo ofrece una nueva perspectiva sobre problemas ya tratados, sino que también proporciona un marco conceptual para abordar estos desafíos de manera más efectiva y contextual.

El presente artículo se estructura de la siguiente manera: primero, estableceré el marco teórico, explorando los principios del atomismo lógico y los fundamentos de la creación de datasets para IA. Luego, desarrollaré la analogía entre el atomismo lógico y la estructura de los datasets, mostrando cómo los datos individuales funcionan como ‘hechos atómicos’ y las etiquetas como ‘proposiciones atómicas’. A continuación, analizaré las implicaciones éticas de esta perspectiva, examinando cómo los sesgos y limitaciones en los datasets se traducen en problemas éticos concretos en las aplicaciones de IA.

Finalmente, propondré algunas direcciones para futura investigación y desarrollo, enfatizando la necesidad de un enfoque interdisciplinario que combine la filosofía del lenguaje, la ética y la ciencia de datos.

Definiciones y conceptos clave

Antes de profundizar en las implicaciones filosóficas y éticas, es crucial establecer una comprensión clara de los términos centrales que se utilizarán en el presente artículo:

- Inteligencia artificial (IA)

Campo de la informática enfocado en crear sistemas capaces de realizar tareas que típicamente requieren inteligencia humana.

- *Machine learning*

Una subcategoría de la IA que implica el desarrollo de algoritmos que permiten a las computadoras aprender de los datos y hacer predicciones o tomar decisiones basadas en esos datos.

- *Dataset*

Grandes conjuntos de datos que se utilizan para entrenar y probar algoritmos y modelos. Estos datasets emplean ‘etiquetas’ (labels) para categorizar la información.

- Red neuronal artificial

Modelos computacionales inspirados en el cerebro humano que se entrenan utilizando datasets. La red procesa estos datos para aprender patrones y hacer predicciones o clasificaciones (un proceso conocido como Deep Learning). La analogía neuronal, aunque poderosa, puede ser engañosa si se toma demasiado literalmente.

Atomismo lógico

El atomismo lógico, propuesto por Bertrand Russell y desarrollado por Ludwig Wittgenstein, representa un intento ambicioso de desentrañar la relación fundamental entre lenguaje y realidad. En esencia, esta teoría propone que existe una correspondencia estructural entre nuestras expresiones lingüísticas y los hechos del mundo. Russell (1914, p.8) argumentaba que “el lenguaje debe tener una estructura que refleje la estructura de los hechos que describe”. Esta afirmación, aparentemente simple, encierra profundas implicaciones filosóficas que nos servirán como punto de partida para nuestra analogía con los datasets en IA. El atomismo lógico se fundamenta en tres principios interrelacionados que merecen un análisis detallado.

- Hechos atómicos y proposiciones atómicas

En el corazón del atomismo lógico yace la idea de que la realidad está compuesta por hechos atómicos, unidades irreducibles que consti-

tuyen los bloques fundamentales del mundo. Correspondiendo a estos hechos atómicos, Russell postula la existencia de proposiciones atómicas en el lenguaje. Como él mismo explica: “Los hechos atómicos son las unidades últimas del mundo externo y las proposiciones atómicas son las unidades últimas de nuestro conocimiento del mundo externo” (Russell, 1918, p. 199).

Esta correspondencia entre hechos atómicos y proposiciones atómicas no es meramente una conveniencia lingüística, sino que representa un intento de alinear nuestra comprensión conceptual con la estructura misma de la realidad. Sin embargo, esta propuesta plantea interrogantes fascinantes: ¿hasta qué punto nuestro lenguaje puede capturar la complejidad del mundo?, ¿existen hechos atómicos que eludan toda representación lingüística?

Descomposición y análisis

El atomismo lógico sostiene que cualquier proposición compleja puede ser descompuesta en proposiciones atómicas. Este proceso de análisis lógico promete una comprensión clara y precisa de las afirmaciones complejas. Wittgenstein (1921/1961, 3.25), en su *Tractatus Logico-Philosophicus*, afirma con característica concisión: “ Toda proposición tiene un análisis completo”.

Este principio de descomposición, aunque atractivo en su simplicidad, plantea desafíos significativos. ¿Es realmente posible descomponer todo enunciado en unidades atómicas sin perder matices o significados emergentes? La idea de que toda complejidad puede reducirse a componentes simples tiene un atractivo intuitivo, pero ¿no corremos el riesgo de caer en un reduccionismo excesivo?

Lenguaje y representación

Finalmente, el atomismo lógico propone que el lenguaje tiene una estructura lógica que refleja la estructura del mundo. Las proposiciones

atómicas son las piezas fundamentales de este lenguaje y su veracidad depende de su correspondencia con los hechos atómicos. Wittgenstein lo expresa de manera poética: “La proposición es una figura de la realidad” (1921/1961, 4.01). Esta metáfora pictórica es poderosa, pero también problemática. ¿Hasta qué punto puede el lenguaje ‘pintar’ una imagen fiel de la realidad?, ¿no está nuestro lenguaje inevitablemente teñido por nuestras percepciones, cultura y limitaciones cognitivas? Estas preguntas adquieren una urgencia renovada en la era de la IA.

La importancia del atomismo lógico para nuestro análisis va más allá de su valor histórico en la filosofía del lenguaje. Nos proporciona un marco para examinar críticamente cómo construimos representaciones de la realidad, ya sea a través del lenguaje natural o de los lenguajes formales de la computación. Al aplicar estos principios a los *datasets* utilizados en IA, podemos obtener perspectivas novedosas sobre cómo estos *datasets* representan (o fallan en representar) la realidad, y cómo estas representaciones pueden llevar a problemas éticos en los sistemas de IA.

Inteligencia artificial y conjuntos de datos

La inteligencia artificial (IA), en su manifestación contemporánea, representa un intento audaz de emular y potencialmente superar capacidades cognitivas humanas a través de sistemas computacionales. Sin embargo, a diferencia de la inteligencia humana, que emerge de procesos biológicos y experiencias acumuladas, la IA se construye sobre una base de datos estructurados y algoritmos de aprendizaje.

Como señala acertadamente Andrew Ng (Marr, 2018), uno de los pioneros en el campo del *Deep Learning*, “[l]os datos son el nuevo petróleo para la IA. Cuantos más datos de alta calidad tengamos, mejores modelos de IA podemos construir”. Esta metáfora del “petróleo de datos” es reveladora en múltiples niveles. Por un lado, enfatiza el valor y la centralidad de los datos en el desarrollo de la IA. Por otro, nos invita a reflexionar sobre las implicaciones éticas y ambientales de esta nueva ‘economía de datos’.

La creación de datasets para IA no es un proceso neutral o meramente técnico. Implica decisiones cruciales sobre qué aspectos de la realidad se capturan, cómo se categorizan, etiquetan y qué se deja fuera. Estas decisiones tienen profundas implicaciones filosóficas y éticas que a menudo se pasan por alto en las discusiones técnicas sobre IA. Esta generación de datasets de alta calidad para IA se rige por varios principios que, aunque presentados como directrices técnicas, tienen profundas raíces filosóficas:

- Relevancia y representatividad

¿Qué constituye una muestra ‘representativa’ de la realidad? Esta pregunta nos lleva al corazón de debates epistemológicos sobre la naturaleza del conocimiento y la posibilidad de una representación objetiva.

- Calidad y precisión

La búsqueda de datos ‘precisos’ presupone una teoría de la verdad y la correspondencia entre nuestras representaciones y la realidad.

- Diversidad y balance

El imperativo de la diversidad en los datasets refleja preocupaciones éticas sobre la equidad y la inclusión, pero también plantea preguntas ontológicas sobre la naturaleza de las categorías y las diferencias.

- Suficiencia de datos

¿Cuántos datos son ‘suficientes’? Esta pregunta tiene resonancias con debates filosóficos sobre la inducción y la justificación del conocimiento.

- Variedad y escalabilidad

La búsqueda de variedad en los datos refleja una comprensión de la complejidad y multidimensionalidad de la realidad que desafía las simplificaciones excesivas.

- Anonimización y privacidad

Las preocupaciones sobre la privacidad de los datos nos llevan a cuestiones éticas fundamentales sobre la autonomía individual y los límites del conocimiento público.

- Consistencia y estabilidad

La búsqueda de consistencia en los datos refleja un compromiso con principios lógicos fundamentales, como la no contradicción.

- Evaluación y validación continua

Este principio reconoce la naturaleza provisional y falible de los datos utilizados.

Al examinar estos principios, se hace evidente que los *datasets* no son meras colecciones de información, sino intentos estructurados de capturar y representar aspectos de la realidad. Cada bit de información en un dataset se convierte, en efecto, en un ‘hecho atómico’ (en el sentido del atomismo lógico) y las etiquetas asociadas funcionan como ‘proposiciones atómicas.’ Gitelman (2013, p. 2) nos recuerda que “los datos brutos son un oxímoron”, enfatizando que incluso los datos aparentemente más básicos ya están imbuidos de decisiones interpretativas y contextuales. Esta observación tiene profundas implicaciones para nuestra comprensión de los *datasets* en IA ya que las etiquetas en los conjuntos de datos de IA, al igual que las proposiciones atómicas en el atomismo lógico, no son meras descripciones neutrales, sino que tienen el poder de dar forma a nuestra comprensión de la realidad.

Kitchin (2014, p. 2) señala que “los datos no existen independientemente de las ideas, instrumentos, prácticas, contextos y conocimientos utilizados para generarlos, procesarlos y analizarlos”. Esta advertencia nos recuerda los peligros de un reduccionismo excesivo, tanto en filosofía como en IA. En última instancia, la creación de datasets para IA nos obliga a confrontar preguntas filosóficas fundamentales sobre la naturaleza de la realidad, el conocimiento y la representación. Preguntas que no son me-

ras abstracciones académicas, sino que tienen consecuencias prácticas y éticas profundas en un mundo cada vez más mediado por sistemas de IA.

Conexión entre atomismo lógico y conjuntos de datos

La aparente disparidad entre el atomismo lógico, una teoría filosófica del siglo XX y los conjuntos de datos modernos utilizados en inteligencia artificial puede, a primera vista, parecer insalvable. Sin embargo, un análisis más profundo revela paralelismos sorprendentes y fructíferos entre ambos conceptos. Esta sección explora estas conexiones, argumentando que existe una analogía significativa entre los principios del atomismo lógico y la estructura de los conjuntos de datos utilizados en IA.

Como ya se señaló previamente, el atomismo lógico postula la existencia de hechos atómicos como las unidades básicas e indivisibles de la realidad. De manera análoga, los datos individuales en un conjunto de datos de IA pueden considerarse como ‘hechos atómicos’ en el sentido de Russell. Cada punto de datos, ya sea un píxel en una imagen, una palabra en un texto, o un valor numérico en una tabla, representa una unidad básica de información. Al igual que los hechos atómicos de Russell, estos datos individuales son las unidades fundamentales sobre las cuales se construye todo el edificio del aprendizaje automático.

Esta analogía no es meramente superficial. Tanto los hechos atómicos del atomismo lógico como los datos individuales en IA comparten una característica crucial: se presentan como elementos objetivos y no interpretados de la realidad. En el atomismo lógico, las proposiciones atómicas son expresiones lingüísticas que corresponden directamente a los hechos atómicos. De manera similar, las etiquetas (*labels*) en los conjuntos de datos de IA funcionan como ‘proposiciones atómicas’, proporcionando una interpretación lingüística o categórica de los datos brutos.

Por ejemplo, en un conjunto de datos de imágenes, la etiqueta ‘gato’ asociada a una imagen particular funciona de manera análoga a una proposición atómica que afirma ‘esto es un gato’. Esta etiqueta, al igual

que una proposición atómica, hace una afirmación sobre la realidad (en este caso, el contenido de la imagen) que puede ser verdadera o falsa.

Sin embargo, esta analogía también revela las limitaciones y peligros potenciales de nuestros sistemas de etiquetado. Como advierte Bowker (2014, p. 65), “las formas en que categorizamos y clasificamos dan forma al mundo en que vivimos”. Las etiquetas en los conjuntos de datos de IA, al igual que las proposiciones atómicas en el atomismo lógico, no son meras descripciones neutrales, sino que tienen el poder de dar forma a nuestra comprensión de la realidad.

El atomismo lógico sostiene que cualquier proposición compleja puede ser descompuesta en proposiciones atómicas. De manera similar, el proceso de creación de conjuntos de datos para IA implica la descomposición de fenómenos complejos en unidades discretas y etiquetadas de información. Este proceso de descomposición y análisis es fundamental tanto para el atomismo lógico como para la creación de conjuntos de datos de IA. Sin embargo, como advierte Kitchin (2014, p. 2), “los datos no existen independientemente de las ideas, instrumentos, prácticas, contextos y conocimientos utilizados para generarlos, procesarlos y analizarlos”. Esta advertencia nos recuerda los peligros de un reduccionismo excesivo, tanto en filosofía como en IA. La analogía entre el atomismo lógico y los conjuntos de datos de IA tiene implicaciones prácticas y éticas significativas:

- Limitaciones de la representación

Al igual que el atomismo lógico ha sido criticado por su incapacidad para capturar ciertos aspectos de la realidad (Quine, 1951), los conjuntos de datos de IA pueden fallar en representar adecuadamente la complejidad del mundo real.

- Sesgos implícitos

Así como las proposiciones atómicas del atomismo lógico pueden contener sesgos lingüísticos y culturales, las etiquetas en los conjuntos de datos de IA pueden perpetuar y amplificar sesgos sociales existentes.

- Poder ontológico

Tanto el atomismo lógico como los conjuntos de datos de IA tienen el poder de dar forma a nuestra comprensión de la realidad, influyendo qué consideramos como ‘hechos’ y cómo los categorizamos.

- Necesidad de contextualización

Al igual que las proposiciones atómicas necesitan ser entendidas en un contexto más amplio, los datos individuales en IA requieren una contextualización cuidadosa para ser interpretados correctamente.

En conclusión, la analogía entre el atomismo lógico y los conjuntos de datos de IA nos proporciona una nueva lente a través de la cual podemos examinar críticamente nuestras prácticas de creación y uso de datos en IA. Nos recuerda que, al igual que el lenguaje en la filosofía, los datos en IA no son un reflejo neutral de la realidad, sino una construcción que incorpora decisiones filosóficas profundas sobre la naturaleza de la realidad y nuestro conocimiento de ella.

Implicaciones éticas

Al considerar los conjuntos de datos como representaciones de la realidad, similares a las proposiciones atómicas en el atomismo lógico, nos enfrentamos a cuestiones fundamentales sobre la objetividad, la representación y la justicia en los sistemas de IA. Un problema ético central que emerge de esta analogía es el del sesgo de representación. Al igual que el lenguaje en el atomismo lógico puede fallar en capturar toda la complejidad de la realidad, los conjuntos de datos en IA a menudo presentan una visión parcial o sesgada del mundo. Como señala O’Neil (2016), estos sesgos pueden tener consecuencias devastadoras cuando se incorporan en sistemas de IA que toman decisiones que afectan la vida de las personas. Por ejemplo, un conjunto de datos de imágenes médicas que no representa adecuadamente la diversidad de tonos de piel podría

llevar a diagnósticos erróneos para ciertos grupos étnicos, perpetuando y exacerbando las desigualdades existentes en la atención médica.

La falta de diversidad cultural en los conjuntos de datos presenta otro desafío ético significativo. Al igual que el atomismo lógico de Russell y Wittgenstein surgió de un contexto filosófico occidental específico, los conjuntos de datos de IA a menudo reflejan los sesgos culturales de sus creadores. Esta falta de diversidad puede llevar a sistemas de IA que malinterpreten o subestimen las expresiones lingüísticas y las necesidades de diferentes grupos culturales. Esto no solo plantea problemas de equidad, sino que también limita la eficacia y aplicabilidad global de los sistemas de IA.

El problema del sesgo de selección en la creación de conjuntos de datos refleja las limitaciones del atomismo lógico en su intento de representar la realidad a través de proposiciones atómicas. Así como la elección de qué proposiciones considerar como ‘atómicas’ puede ser arbitraria o culturalmente sesgada, la selección de datos para incluir en un conjunto de datos de IA puede introducir sesgos significativos. Mittelstadt (*et al.*, 2016) señalan que estos sesgos pueden ser particularmente perniciosos cuando se trata de decisiones automatizadas en áreas sensibles como la justicia penal o la evaluación crediticia.

El impacto de los conjuntos de datos sesgados en las decisiones automatizadas es quizás la implicación ética más urgente de esta analogía. Al igual que las proposiciones atómicas en el atomismo lógico se consideraban los bloques de construcción de afirmaciones más complejas sobre la realidad, los datos individuales en los conjuntos de datos de IA se convierten en la base sobre la cual se toman decisiones automatizadas. Barocas y Selbst (2016, p. 671) argumentan que esto puede llevar a una “discriminación por diseño”, donde los sesgos históricos y sociales se codifican y amplifican a través de sistemas de IA aparentemente objetivos.

El reforzamiento de estereotipos a través de conjuntos de datos sesgados es otra consecuencia ética preocupante. De la misma manera que el lenguaje puede perpetuar estereotipos y prejuicios, los con-

juntos de datos de IA pueden reforzar y amplificar las desigualdades existentes. Noble (2018) demuestra cómo los algoritmos de búsqueda, entrenados en conjuntos de datos sesgados, pueden perpetuar estereotipos raciales y de género, influyendo en la percepción pública y las oportunidades individuales.

Estas implicaciones éticas nos obligan a reconsiderar fundamentalmente nuestro enfoque de la creación y uso de conjuntos de datos en IA. Al igual que el atomismo lógico nos invita a examinar críticamente la relación entre lenguaje y realidad, debemos escrutinar cuidadosamente cómo nuestros conjuntos de datos representan y dan forma al mundo.

Conclusiones

El análisis de la relación entre el atomismo lógico y los conjuntos de datos de IA nos permite comprender mejor los desafíos éticos que enfrentamos en el desarrollo y aplicación de la inteligencia artificial. Este enfoque filosófico-lingüístico nos ofrece una nueva perspectiva para abordar y mitigar los problemas éticos en IA, enfatizando la importancia de la diversidad, representatividad y equidad en la creación de conjuntos de datos.

En última instancia, la analogía entre el atomismo lógico y los conjuntos de datos de IA nos recuerda que la creación de conocimiento, ya sea a través del lenguaje filosófico o de los datos computacionales, nunca es un proceso neutral. Nos enfrentamos al desafío ético de crear sistemas de IA que no solo sean eficientes y precisos, sino también justos, inclusivos y respetuosos de la diversidad humana. Este desafío requiere una colaboración continua entre filósofos, éticos, científicos de datos y desarrolladores de IA para garantizar que nuestros sistemas tecnológicos reflejen nuestros valores éticos más elevados.

Propongo que futuros desarrollos en IA deberían considerar explícitamente esta conexión entre lenguaje, representación y ética, para crear sistemas más justos y equitativos. Esto podría implicar nuevas metodologías

para la creación de conjuntos de datos, así como un escrutinio más riguroso de los principios filosóficos que subyacen a nuestras prácticas de IA.

En conclusión, al reinterpretar los desafíos éticos de la IA a través de la lente del atomismo lógico y la filosofía del lenguaje, podemos obtener nuevas perspectivas sobre cómo crear sistemas de IA más éticos y equitativos. Este enfoque nos recuerda que la IA, al igual que el lenguaje, es una herramienta poderosa para dar forma a nuestra comprensión del mundo, y como tal, debe ser desarrollada y utilizada con un profundo sentido de responsabilidad ética.

Referencias bibliográficas

- Barocas, S. y Selbst, A. (2016). Big Data's Disparate Impact. *California Law Review*, 104(3), 671-732.
- Bowker, G. C. (2014). The Theory/Data Thing. *International Journal of Communication*, 8, 1795-1799.
- Gitelman, L. (Ed.). (2013). "Raw Data" Is an Oxymoron. MIT Press.
- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. SAGE Publications.
- Marr, B. (2018, May 21). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. *Forbes*. <https://bit.ly/41EgGTA>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S. y Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1-21.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishers.
- Quine, W. V. O. (1951). Two Dogmas of Empiricism. *The Philosophical Review*, 60(1), 20-43.
- Russell, B. (1914). *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. Open Court Publishing Company.
- Russell, B. (1918). The Philosophy of Logical Atomism. *The Monist*, 28(4), 495-527.
- Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus* (D. F. Pears & B. F. McGuinness, Trans.). Routledge & Kegan Paul.

Transhumanismo y la inteligencia artificial: un análisis de sus implicaciones éticas para la humanidad

Johan Manuel Méndez Reyes
Universidad Politécnica Salesiana
jmendezr@ups.edu.ec
<https://orcid.org/0000-0002-9349-223X>

Introducción

La sociedad actual, inmersa en la lógica del libre mercado y la globalización, está en la encrucijada impuesta por el paradigma de la racionalidad instrumental moderna. Esta perspectiva ve el progreso científico como la clave para alcanzar el máximo potencial humano, guiado por un determinismo tecnológico. La ciencia se presenta como una actividad neutral que observa la naturaleza para comprender sus leyes y estructuras, con la promesa de un conocimiento tan acelerado que nos permitirá manipularla a nuestro libre albedrío. La tecnología, actuando como un sistema autónomo que impulsa el cambio social y cultural, devendría la piedra angular del progreso y la prosperidad de la humanidad.

La hegemonía de la racionalidad moderna, caracterizada por su énfasis en el individualismo desmedido y el consumismo desenfrenado, ha propiciado el surgimiento de una neocolonialidad o colonialidad epistémica que permea tanto el ámbito científico como el político. La tecnociencia se erige como un elemento indispensable, generando una

creciente dependencia que reproduce y exagera las desigualdades estructurales. En este contexto, surge la propuesta del transhumanismo que busca superar las limitaciones inherentes a la condición humana mediante la manipulación artificial de la evolución. Ella aspira a impulsar modificaciones en el cuerpo humano con el objetivo de alcanzar la inmortalidad, la perfección y la eventual sustitución de la memoria humana por inteligencia artificial. Este panorama suscita un debate ético sobre el porvenir de la humanidad y las profundas implicaciones que conlleva la tecnociencia en un mundo globalizado.

Tanto el transhumanismo como el posthumanismo procura superar las limitaciones inherentes a la condición humana mediante avances científicos y tecnológicos. Estas posturas suscitan un debate sobre los desafíos éticos que plantean, especialmente en lo concerniente a la validez científica y la aplicabilidad universal de sus propuestas. Aunque es factible y realista considerar avances como la prolongación de la vida, la prevención de enfermedades mortales, mejoras en la calidad de vida en etapas avanzadas y el desarrollo de inteligencias ampliadas, resulta imperativo cuestionar el grado en que estas aspiraciones están respaldadas por la ciencia actual (Cortina, 2022).

La revolución de la industria 4.0 ha puesto de manifiesto el potencial transformador de la tecnología en términos de mejorar nuestras capacidades y nuestro bienestar. No obstante, es importante reconocer los límites claros en cuanto a lo que la ciencia puede lograr. Si bien el aumento de la longevidad es un objetivo alcanzable, la aspiración a la inmortalidad, la salud perpetua o el desarrollo de superinteligencias y una inteligencia artificial general aún no gozan de una aceptación global.

A medida que la tecnología evoluciona y se perfecciona, la Inteligencia Artificial (IA) ha logrado una rápida expansión y dado lugar a múltiples cambios en diversos sectores de la vida. La capacidad que posee la tecnología se ha transformado en un elemento fundamental para la evolución de la sociedad actual. No obstante, es importante estudiar sus

alcances, efectos y los peligros que su utilización en el ámbito humano puede implicar.

El transhumanismo tecnocientífico impulsa una carrera hacia la creación y fabricación de máquinas que se asemejen cada vez más a los seres humanos. Inicialmente, se adopta una estructura conductista que se caracteriza por el énfasis en la acumulación de información, la memorización repetitiva y la limitación en ciertos ámbitos. Sin embargo, con el avance y la proliferación de la inteligencia artificial, se busca desarrollar una especie de máquina con mayor autonomía. Esta máquina no solo sería independiente del ser humano, sino que también podría fusionarse de alguna manera con el cuerpo humano, trascendiendo así incluso los límites de la ciencia ficción (Kurzweil, 2005).

Uno de los principales retos éticos consiste en distinguir entre lo que es científicamente factible y lo que podría representar un salto hacia un nuevo tipo de ser humano o post humano. Es importante mantener una perspectiva crítica y reflexiva sobre estas propuestas utópicas para evitar caer en excesos o distorsiones de lo que la ciencia puede ofrecer. La búsqueda de la mejora humana y el avance tecnológico deben ir acompañados de una evaluación cuidadosa de sus implicaciones éticas, sociales y filosóficas. Es fundamental asegurar que los beneficios se equilibren con los posibles riesgos y desafíos que estas transformaciones puedan conllevar (Cortina, 2022).

Efectivamente, la ciencia y la tecnología, herramientas originalmente concebidas para el progreso, pueden convertirse en instrumentos de dominio en manos de grupos de poder. Así lo advirtió Habermas (1984), hoy en día lo vemos reflejado en la forma en que empresas como la industria farmacéutica y las grandes plataformas digitales utilizan la opinión pública y las redes sociales para promover sus agendas e intereses. Esta instrumentalización de la ciencia y la técnica se enmarca en una estructura ideológica que legitima la dominación, desviando su potencial de transformación social hacia el control y la manipulación (Cortina, 2022). Apel (1985) y Habermas (1984) coinciden en la importancia de distinguir

entre los intereses que impulsan el conocimiento. El enfoque científico-técnico busca el dominio, mientras que el enfoque práctico se centra en la comprensión mutua. El diálogo entre individuos es fundamental para determinar la utilidad y el impacto de la tecnociencia en pro del progreso y la libertad, que permita superar las sociedades ideologizadas.

Si la ciencia y la tecnología, ahora convertidas en tecnociencias, se erigen como instrumentos ideológicos que dominan el espacio público con promesas inalcanzables, no solo no contribuyen al progreso, sino que refuerzan la dominación al adueñarse de este espacio. Aunque los defensores del transhumanismo y el posthumanismo promuevan debates públicos, si la información que recibe el público es engañosa, el discurso transhumanista se convierte en una herramienta para legitimar una ideología distorsionada, como bien señala Cortina (2022).

El análisis crítico desde una perspectiva ideológica nos plantea un reto ético adicional en el contexto del transhumanismo y el posthumanismo. ¿Acaso estas corrientes utilizan promesas inalcanzables para seducir las aspiraciones humanas, impulsadas por quienes controlan la economía y la ciencia a nivel global? Es vital dilucidar entre una comunicación que busca la dominación y aquella que busca una comprensión genuina en lugar de una adhesión superficial. La intersubjetividad humana debe guiar el uso de las tecnologías, incluidas las tecnocientíficas, a través de un diálogo informado y reflexivo (Cortina, 2022).

Es importante destacar que este trabajo no busca generar temor irracional hacia la tecnología, sino más bien promover una reflexión crítica sobre el impacto del transhumanismo tecnocientífico y el futuro de las inteligencias no biológicas. Se invita a considerar cuidadosamente los posibles conflictos éticos que podrían surgir en su aplicación en el ámbito humano.

Este ensayo tiene como objetivo analizar las implicaciones éticas que tiene para la humanidad el transhumanismo y la inteligencia artificial. Se basa en una investigación cualitativa, utilizando principalmente la hermenéutica y la exégesis como herramientas de análisis. El estudio

invita a una reflexión interdisciplinaria y transdisciplinaria sobre el transhumanismo tecnocientífico, la inteligencia artificial y las implicaciones éticas de este pensamiento. En efecto, este análisis adopta un enfoque cualitativo basado en los juicios de quien recopila los datos, lo que le permitirá examinarlos e interpretarlos según el contexto de la investigación. Se recurrió a múltiples fuentes bibliográficas para investigar diversas teorías de referencia. Asimismo, se aplicó el método hermenéutico con el propósito de interpretar los datos relacionados con la realidad objeto de estudio. Durante este proceso, se implementaron técnicas propias del diseño de investigación documental, tales como la revisión de material digital e impreso relevante para identificar, describir y entender la temática abordada.

La metodología documental no se limita a la revisión de textos, sino que busca revelar tanto los aspectos teóricos como empíricos de los argumentos presentes en los datos estudiados y en la perspectiva de quien desarrolla el estudio, lo que permite un análisis crítico que enriquece el trabajo (Sabino, 2007). Este método, de carácter cualitativo y hermenéutico, explora conceptos teóricos relacionados con el transhumanismo y la inteligencia artificial, así como sus implicaciones éticas para la humanidad. Con el fin de examinar esta realidad desde un enfoque cualitativo, se eligieron cuidadosamente los textos que facilitaron el estudio del transhumanismo y la inteligencia artificial y su impacto en la humanidad. Asimismo, del método hermenéutico, se aplicó el análisis de contenido, que Hurtado (2005) describe como un proceso sistemático y organizado de estudio de documentos.

Transhumanismo: una aproximación general a su concepto

El transhumanismo —también conocido como H— promueve el uso creciente de la tecnología para potenciar las capacidades humanas, integrando progresivamente estos avances en las personas con el propósito de perfeccionar su vida, dando lugar a individuos dotados de mayor inteligencia, inmunidad a enfermedades, belleza y fuerza física. En este

sentido, el transhumanismo sostiene que los humanos superarán su estado natural actual, al que se refiere como estado intermedio, superando sus limitaciones mediante el uso de medios tecnológicos. Así, se plantea una evolución autodirigida hacia formas de vida más avanzadas (Alba y Calero, 2023).

Las primeras ideas sobre el transhumanismo se remontan a la obra *Daedalus; or, Science and the Future* del bioquímico británico Haldane en 1923, donde se vislumbra la posibilidad de utilizar la tecnología para mejorar las capacidades humanas y se anticipa un futuro de prosperidad y salud para la sociedad. Sin embargo, se considera a Julian Huxley como uno de sus principales pioneros, quien introduce, en 1946, por primera vez el término transhumanismo como sinónimo de humanismo evolucionista, argumentando que los seres humanos pueden potenciarse a sí mismos mediante el uso de la tecnología, alcanzando así un nuevo nivel de existencia (Garcés Castellote y Jiménez Rodríguez, 2016).

En su obra *La perspectiva de la inmortalidad*, Ettinger planteó la posibilidad de una alternativa a la muerte, dando origen al tema de la criogenización en 1962. Esta técnica plantea la posibilidad de preservar a las personas después de la muerte con la esperanza de resucitarlas en el futuro. Esta obra es, también, considerada como un estudio precursor del transhumanismo (Garcés Castellote y Jiménez Rodríguez, 2016). En la década de 1980, el transhumanismo experimentó un renovado interés. Destacados autores como Bostrom, de Grey, More, Vita-More, Kurzweil, Pearce, entre otros, han profundizado en esta temática. Comparten la perspectiva de que el transhumanismo brinda la oportunidad de potenciar las facultades del ser humano a través de la incorporación de los avances tecnológicos. Igualmente, diversas organizaciones e instituciones han dedicado esfuerzos significativos al estudio del transhumanismo. Entre ellas se encuentran la Asociación Transhumanista Mundial, el IIEET (Institute for Ethics and Emerging Technologies), la Universidad de la Singularidad, la Alcor Life Extension Foundation, entre otras instituciones de renombre. Todas ellas promueven el transhumanismo y realizan investigaciones

sobre las tecnologías que podrían hacer realidad sus principios (Garcés Castellote y Jiménez Rodríguez, 2016).

El transhumanismo ha sido interpretado de diversas maneras. Por un lado, se le ha visto como una expresión científica futurista con aspiraciones escatológicas, es decir, que la tecnociencia posee el potencial de trascender los límites biológicos de la humanidad. Por otro lado, se percibe como una corriente materialista que busca mejorar las capacidades humanas mediante los avances tecnológicos. Sin embargo, también ha sido objeto de críticas, ya que se argumenta que sus contribuciones pueden desatender las implicaciones éticas y sociales inherentes a la humanidad. Su impacto podría intensificar la desigualdad y la discriminación existentes, dado que solo aquellos con recursos económicos podrían acceder a las tecnologías que mejoran la vida humana.

A la vez, puede interpretarse como una corriente ideológica que sostiene la obligación moral de optimizar las capacidades corporales e intelectuales de la humanidad, utilizando los avances tecnológicos para eliminar aspectos indeseables y prevenibles de la imperfección del ser humano, como la angustia, el dolor, la decadencia e incluso la propia muerte. Este propósito apunta a la creación de una nueva especie posthumana. Si bien este mensaje resulta atractivo para aquellos que forman parte de una élite económica, científica e intelectual, en realidad no constituye una propuesta accesible para todos.

Serra (2022) explora los conceptos de posthumano, transhumano y transhumanismo, enfocándose en el uso de tecnologías emergentes para el mejoramiento humano y sus implicaciones científicas y sociales. Es importante valorar las técnicas de edición genética, sus beneficios y riesgos, adoptando una perspectiva de responsabilidad personal y colectiva. Al mismo tiempo, se debe fomentar un enfoque inclusivo que respete los valores de la persona, sus costumbres e incluso su propia esencia, independientemente de sus capacidades. La ausencia del marco ético global sobre la edición genética y la necesidad de alcanzar un consenso ético son aspectos fundamentales, alertando contra una sociedad dividida por

el mejoramiento humano. Según Pearce (2004), los objetivos generales del transhumanismo abarcan la consecución de la superinteligencia, la superlongevidad y el superbienestar para toda la humanidad. Este autor sostiene que el progreso continuo en tecnología, impulsado por el desarrollo de la IA y la correlación de las NBIC¹ eventualmente conducirá a una singularidad tecnológica en un futuro cercano.

Resulta cada vez más evidente que las tecnologías NBIC están destinadas a transformar el rumbo de la humanidad. La nanotecnología promete resolver una amplia gama de problemas globales y abrir nuevas posibilidades en la creación de materiales. Por su parte, los avances en biotecnología sentarán las bases para el desarrollo de otros fármacos y tratamientos. Aunado, al desarrollo de la neuromedicina que ampliarán la comprensión del funcionamiento cerebral, mientras que el desarrollo de las TIC ofrecerá una base común que permitirá integrar los aportes del lenguaje algorítmico y de la informática a diversas disciplinas.

Kurzweil (2012) sostiene que la singularidad tecnológica representa un punto en el tiempo en el que el progreso tecnológico se acelerará a una velocidad tal que tendrá un impacto irreversible en la vida humana. Este avance podría conducir a la humanidad hacia una forma superior de existencia, conocida como posthumanos o posthumanidad. Si bien la singularidad tecnológica se percibe como una oportunidad para forjar un futuro mejor, también plantea desafíos éticos que deben abordarse con cautela. Estamos adentrándonos en una era posthumana en la cual la evolución ya no es exclusivamente natural, sino que estará dirigida por la tecnología. La manipulación de la línea germinal, que permite modificar el ADN de las células reproductivas mediante el uso de la tecnociencia, podría eventualmente extenderse a una gran parte de la población humana (Mehlman, 2009). Esto implica que la humanidad podría tomar las riendas de su propio proceso evolutivo, lo que tendría un impacto significativo en la naturaleza humana.

1 Nanotechnology, Biotechnology, Information technology, and Cognitive science.

El transhumanismo parte de la premisa de que el progreso científico y tecnológico confiere a los seres humanos la capacidad de emanciparse de las leyes de la evolución y la naturaleza. Se convierten en agentes que controlan su propio destino y el de los seres vivos. Se les atribuye la facultad de generar vida y corregir las imperfecciones biológicas inherentes a la existencia terrenal. Este cambio de paradigma sugiere una visión en la que la autonomía y la capacidad de manipular la realidad biológica definen la relación de la humanidad con la evolución y la naturaleza (Méndez Reyes y Padrón Medina, 2024).

Otros autores, como Abbate (2022), exploran el paralelismo entre la gnosis y el transhumanismo, destacando su relación con el posthumanismo. Ambos comparten un enfoque dualista que percibe la realidad material, especialmente el cuerpo, como una limitación que debe ser superada. En la gnosis, el descubrimiento del *pneuma*² simboliza el inicio de un camino redentor, mientras que el transhumanismo busca la salvación a través de la técnica aplicada al cuerpo. Sin embargo, este esfuerzo por recrear la realidad, basado en una visión del mundo inhóspito e injusto, no resuelve la angustia humana, sino que la intensifica y conduce a un mayor desarraigo y soledad en la posmodernidad.

Dobre y Pavón (2024) proponen un enfoque crítico desde la filosofía de Søren Kierkegaard hacia el transhumanismo y su concepto de singularidad. En primer lugar, examinan cómo los transhumanistas interpretan la singularidad y sus limitaciones, para luego argumentar, basándose en el pensamiento kierkegaardiano, que la verdadera singularidad reside en la persona humana como una relación fundamental, caracterizada por la angustia, la libertad, el pecado, el amor y el espíritu.

2 Se entiende como el principio espiritual o esencia divina que otorga al ser humano la capacidad de vincularse con la dimensión trascendente. Esta concepción hunde sus raíces en la filosofía griega, particularmente en Platón y en la escuela estoica, donde fue concebido como sopro vital y racional. No obstante, en los sistemas gnósticos de los siglos II y III (d.C) adquirió un matiz redentor, al ser considerado el componente que faculta al individuo a trascender la materia y dirigirse hacia lo divino.

Por lo que, ninguna creación artificial puede reproducir esta relación esencial, subrayando la exclusividad de la condición humana frente a las promesas del transhumanismo.

En su estudio, González Ramírez (2024) explora el vínculo conceptual entre el transhumanismo y la filosofía de García Bacca, destacando la transfinitud como base teórica para comprender la inclinación humana hacia la trascendencia. Por un lado, el transhumanismo pretende minimizar las imperfecciones del ser humano a través de la tecnociencia; por otro, la transfinitud representa la fuerza inherente del hombre para negarse a aceptar límites infranqueables. Ambas corrientes comparten un profundo apego a la tecnociencia, considerada casi como una religión capaz de resolver todos los problemas humanos. La transfinitud, entendida como clave para alcanzar la inmortalidad, posiciona a la filosofía garciabaquiana como precursora del transhumanismo.

Los transhumanistas buscan diferenciarse de proyectos previos de modificación humana de corte totalitario, enarbolando la bandera del progresismo y vinculando sus ideas al pluralismo cultural y sistemas democráticos globales. Ejemplo de ello es James Hughes (2004), quien propone un “transhumanismo democrático” que integra la biopolítica transhumanista con la política socialdemócrata, económica y liberal cultural. Hughes (2004) argumenta que el mejor futuro posthumano se alcanzará asegurando la seguridad y accesibilidad de las tecnologías para todos, respetando los derechos individuales sobre el control de sus propios cuerpos; sin embargo, para garantizar que se cumplan estos planteamientos en la sociedad, es importante que los gobiernos centrales hagan cumplir, por medio de sus leyes, estos beneficios que provienen de la tecnociencia. Esto nos lleva a cuestionarnos si estas afirmaciones son simplemente retórica políticamente correcta o si representan propuestas éticamente sólidas.

El transhumanismo, para Sánchez Salazar (2024), plantea la mejora humana como un medio para lograr el máximo desarrollo de las facultades físicas y cognitivas. Esta idea ha originado el llamado derecho al autodiseño, el cual alude a la capacidad de modificar el propio cuerpo

según la libertad personal, con el fin de potenciar las habilidades humanas. Esta perspectiva redefine los derechos humanos al incorporar el concepto del derecho a la auto-mejora, proponiendo que este derecho al autodiseño sea protegido y esté exento de intervenciones por parte de terceros o del Estado.

López Fernández (2024) hace referencia a la trilogía *Coup de sang* (2009-2015) de Enki Bilal para explorar el transhumanismo y sus implicaciones en un mundo transformado por el Antropoceno. En este contexto postapocalíptico, los sobrevivientes, obligados por un planeta dañado por el impacto humano, adoptan formas híbridas de existencia, fusionando características humanas y animales. Se plantea cuestiones éticas y filosóficas sobre los avances tecnológicos propios del transhumanismo, advirtiendo sobre los peligros de cruzar límites biológicos y morales. Lo que es importante reflexionar sobre los desafíos del posthumanismo y la sostenibilidad de la vida en un entorno radicalmente alterado.

Por otro lado, Gaviria (2024) argumenta que el transhumanismo no es incompatible con la educación y que puede existir una versión humana de este enfoque. Aunque los bioconservadores critican las intervenciones biotecnológicas por considerar que atentan contra la dignidad humana, Gaviria señala que estas pueden ser éticamente válidas si aumentan la autonomía de los individuos transformados. El transhumanismo comparte con la educación la intención de optimización, diferenciándose solo en los métodos empleados: biotécnicos y simbólicos, respectivamente, es por eso, que plantea criterios para evaluar cuándo estas modificaciones son aceptables, destacando la importancia de reflexionar sobre los límites que enfrenta la educación.

Castillo (2023) examina la viabilidad de un nuevo ideal educativo de racionalidad frente a dos corrientes opuestas: el transhumanismo y el pensamiento crítico. Sostiene que el transhumanismo, al intentar superar el humanismo, genera desconfianza hacia la razón. Por otro lado, el pensamiento crítico, fundamentado en la racionalidad de Scheffler, revela las falencias del imperialismo científico que oculta los aspectos

ético-rationales de la educación. En consecuencia, no sería necesario un nuevo ideal, sino más bien preservar el equilibrio entre lo moral y lo racional, reforzando el rol del docente en la enseñanza crítica sin ceder al reduccionismo tecnocientífico.

Pavón y Macip (2023) critican la visión tecnocrática del transhumanismo en la educación, que favorece un enfoque reduccionista centrado en el desarrollo funcional y exitoso, más que en el crecimiento integral de los estudiantes. Señalan que el avance humano, cuando depende de factores externos como la tecnología o el diseño genético, puede desmotivar el aprendizaje y la formación del carácter. En contraste, abogan por una educación humanista que valore la libertad, la motivación y el sentido del aprendizaje, lo que promueve la autonomía individual y la participación comunitaria. Por lo que, defienden un modelo educativo inclusivo que favorezca la justicia curricular y el desarrollo de capacidades colectivas.

Transhumanismo tecnocientífico o cibernético

El transhumanismo tecnocientífico o cibernético, impulsado por la inteligencia artificial (IA), el software y la robótica, como lo visionaron Minsky (1960), Moravec (1998), Kurzweil (1999), Bostrom (2002) y Sandberg (2001), busca fusionar humanos y máquinas mediante IA y robótica. En lugar de la biología, se pretende llegar a una era posthumana desde la tecnológica, ofreciendo un nuevo ser dotado físicamente con una inteligencia superior a la del ser humano. Esta visión representa un tipo de posthumanismo en el que la fusión hombre-máquina supera la inteligencia humana con IA, con el potencial de eventualmente reemplazar a los humanos (Cortina, 2022).

La corriente del transhumanismo tecnocientífico se basa en la convicción de que la capacidad humana es capaz de generar una identidad completamente nueva, redefinir su posición en el mundo, transformar su relación con otros seres en el universo y reconsiderar su concepción de la trascendencia humana. Esta corriente parte de la premisa fundamental

de que los individuos poseen la innata habilidad de moldear su propio destino, así como de crear nuevas creencias y narrativas (Méndez Reyes y Padrón Medina, 2024).

La noción de que la tecnología va más allá de ser solo una herramienta y está intrínsecamente vinculada con la esencia misma de la humanidad ha sido una narrativa arraigada en nuestra historia. Algunos planteamientos, en especial aquellos que abordan la mejora de la especie mediante avances tecnológicos, la consideran como un medio para alcanzar la libertad y la superación. El transhumanismo se basa en este criterio al abogar por una evolución artificial dirigida. Sus seguidores, como Bostrom (2005), sostienen la creencia de que la tecnología tiene el potencial de convertirnos en *posthumanos* con capacidades extraordinarias: una longevidad infinita, un intelecto superior y un control total sobre nuestras emociones.

A diferencia de la evolución biológica, la evolución dirigida implica una orientación deliberada y fundamentada en la razón, con miras a la creación de una especie *sucesora*. De esta premisa emerge el término ‘transhumano’, que representa un ser en transición entre lo humano y lo posthumano (FM-2030, 1989). La visión de Kurzweil (2005), que contempla la fusión de humanos con nanobots e inteligencias artificiales, es una de las versiones más conocidas de esta narrativa, aunque existen otras igualmente plausibles, como el uso de la ingeniería genética con el mismo propósito (Vaccari, 2016).

La aspiración central del proyecto transhumanista reside en la realización de los anhelos fundamentales de la humanidad: alcanzar una autonomía total, eludir la muerte y perseguir la perfección. Esto implicaría otorgar un propósito esencial a la existencia y establecer los cimientos para una sociedad innovadora regida por la organización y la racionalidad tecnológica y digital. En este contexto, la información se almacenaría en dispositivos que encapsulan todo el conocimiento de estos individuos posthumanos, marcando así un cambio radical en la comprensión y ex-

perencia de la vida en esta nueva etapa de la humanidad (Méndez Reyes y Padrón Medina, 2024).

Díaz Alzate *et al.* (2023) reflexionan sobre cómo los avances tecnocientíficos han transformado las condiciones humanas y destacan su influencia en la sociedad contemporánea y el capitalismo. Estos promueven, desde la globalización, las nuevas formas en que se determinan las relaciones humanas y dan paso a nuevas subjetividades y a la transformación del concepto del ser humano. El humanismo pierde vigencia y es reemplazado por el posthumanismo, un ser que trasciende sus condiciones naturales mediante la tecnología, buscando la inmortalidad y cuestionando los límites de la vida y la muerte. Este cambio de paradigma desafía los límites humanos, mejorando sus capacidades físicas, intelectuales y emocionales.

Según Camargo Brito y Ried (2024), el transhumanismo plantea desafíos significativos a la bioética contemporánea, abordados desde una perspectiva genealógica que cuestiona las condiciones de inteligibilidad de este fenómeno. Este enfoque permite analizar cómo el transhumanismo impacta los criterios con los que se manipula la vida. En este escenario, donde se trastoca las bases humanistas, el transhumanismo procura reconfigurar la discusión sobre la condición humana, formulando nuevamente la pregunta sobre la posibilidad y la necesidad de trascender lo humano, reavivando el debate crítico sobre los dilemas éticos derivados de los avances tecnocientíficos.

El transhumanismo está ganando relevancia en la cultura de masas debido al creciente interés de los científicos en las posibilidades que ofrece para mejorar la condición humana mediante tecnologías avanzadas. Es fundamental analizar los dilemas éticos que enfrenta la bioética en relación con el transhumanismo, destacando tanto las convergencias como las divergencias en sus perspectivas. Es necesario considerar el impacto en diversos ámbitos (social, legal, tecnológico, moral y filosófico) y cómo el transhumanismo ha generado nuevos desafíos. Es esencial desarrollar principios éticos que guíen estos avances tecnológicos de manera responsable y moralmente alineada (Hernández *et al.*, 2023).

Inteligencia humana e inteligencia artificial (IA): una mirada ética

Abordar los complejos dilemas éticos del transhumanismo cibernético requiere una comprensión de lo que se ha entendido como inteligencia, junto con un análisis de, al menos, tres tipos de inteligencia artificial (IA) (Cortina, 2022). Si bien la inteligencia suele considerarse una característica exclusiva de los seres racionales, su alcance no se limita al ámbito humano, sino que se manifiesta en diversas formas en el mundo natural, incluyendo animales, plantas e incluso el planeta Tierra, tal como lo propone la hipótesis Gaia de James Lovelock (1983). Esta teoría postula que la vida en la Tierra funciona como un sistema autorregulado que mantiene condiciones idóneas para la vida. Todos los seres vivos interactúan y contribuyen al control de los atributos del planeta, demostrando la unidad de la biosfera. Los humanos, como parte integral de esta biota, pueden actuar como un sistema de alerta para Gaia. Comprender estos mecanismos nos permite entender la habitabilidad terrestre y diseñar hábitats sostenibles en el espacio (Sagan y Margulis, 1984).

Etimológicamente, la palabra inteligencia deriva del latín *inter* —que significa entre— y *eligere* —que significa escoger—. Desde este contexto, se puede definir la inteligencia como la capacidad cerebral para comprender y seleccionar el curso de acción más adecuado. Desde tiempos remotos, la humanidad ha sentido una profunda curiosidad por comprender el origen y los mecanismos de la inteligencia, impulsada por el deseo de dominarla. La conducta inteligente se caracteriza por la capacidad de enfrentar y resolver con éxito desafíos tanto internos como externos (Ponce y Sanmartín, 2010).

La inteligencia, en términos psicológicos, es entendida como la habilidad para asimilar y utilizar el aprendizaje dado por la experiencia a situaciones nuevas. Las definiciones más comunes enfatizan la capacidad de pensamiento abstracto, aprendizaje, adaptación y resolución de problemas. Binet y Simon (1916) la consideraban como la aptitud para aprender, mientras que Piaget (1985) la veía como un proceso evolutivo

de adaptación al medio. La inteligencia, lejos de ser una capacidad única y monolítica, se manifiesta en una rica diversidad de formas, lo que exige un enfoque educativo inclusivo y en constante evolución. Según Gardner (1994), la inteligencia va más allá de la mera resolución de problemas; implica generar acciones que beneficien tanto al individuo como a la comunidad.

La propuesta de Gardner (1999) sobre las inteligencias múltiples cuestiona la idea convencional de la inteligencia limitada por la capacidad cognitiva. Esta teoría reconoce la importancia de las emociones, la personalidad y el entorno sociocultural en el desarrollo de la inteligencia. Este autor aboga por crear entornos educativos que fomenten el desarrollo de todas las habilidades humanas, partiendo de dos premisas fundamentales: todos los individuos nacen con habilidades innatas que se desarrollan a lo largo de su vida y el aprendizaje continuo es esencial para el desarrollo personal, aunque no garantiza la obtención de una verdad absoluta (Méndez Reyes y Padrón Medina, 2022). La comprensión de la inteligencia es fundamental para el desarrollo humano y la solución de problemas complejos (O’Conor, 1999). En este sentido, la inteligencia también se asocia a la capacidad de establecer objetivos, planificar, anticipar consecuencias y emplear herramientas para alcanzarlos, siendo la capacidad de resolver problemas con herramientas un componente esencial (Cortina, 2022).

Por otra parte, el término de inteligencia artificial (IA) surge en 1955 durante un congreso en Los Ángeles sobre máquinas con capacidad de aprendizaje. En 1956, John McCarthy introduce el concepto inteligencia artificial, refiriéndose a la creación de máquinas que pueden considerarse inteligentes debido a su interacción con humanos, al punto de que una persona no puede distinguir si está interactuando con una máquina o con otro ser humano (test de Turing³) (McCarthy *et al.*, 2006 citado en Cortina, 2022).

3 Alan Turing (1912-1954) es reconocido como el pionero de la informática debido a su invención de la máquina en 1936, que representa una conceptualización inicial

La (IA) se define como la rama de la informática que busca crear máquinas capaces de realizar tareas que, generalmente, requieren inteligencia humana (Iberdrola, 2019). Esta tecnología se basa en el desarrollo de algoritmos complejos que posibilitan que las máquinas adquieran conocimientos, piensen y decidan por sí mismas. El Manifiesto de Dartmouth, publicado en 1956, predijo que la IA alcanzaría la inteligencia general en un futuro cercano. Sin embargo, esta ambiciosa meta aún no se ha logrado y el alcance de la IA sigue siendo un tema de debate. Algunos autores sostienen que la IA se centra en imitar la forma de actuar o pensar de los humanos, mientras que otros enfatizan la creación de sistemas que razonan y actúan como ellos (Marín García, 2019).

Una definición más técnica de la IA, según el Grupo de Expertos de Alto Nivel en Inteligencia Artificial (2019), es la de sistemas que perciben su espacio, procesan datos y toman decisiones para cumplir objetivos complejos, por medio del análisis de acciones anteriores. Asimismo, pueden emplear reglas simbólicas o aprender modelos numéricos, y pueden ajustar su comportamiento desde entornos similares (Cortina, 2022). Según la clasificación de Russell y Norvig (2016) citada por Marín García (2019), las definiciones de inteligencia artificial se pueden agrupar en cuatro categorías según los objetivos que se persiguen:

- a. Sistemas que imitan el comportamiento humano: inspirados en el Test de Turing (1950), estos sistemas tienen la capacidad de procesar lenguaje natural, representar conocimiento, razonar de forma automática y aprender para adaptarse a nuevas situaciones (Kurzweil, 1999).
- b. Sistemas que replican el pensamiento humano: estos sistemas automatizan operaciones mentales como la toma de decisiones, la resolución de problemas y el aprendizaje (Bellman, 1978).

de la Inteligencia Artificial. En 1950, en su artículo en la revista *Mind*, argumenta que los ordenadores tienen la capacidad de exhibir comportamientos inteligentes.

- c. Sistemas que piensan racionalmente: emulan el pensamiento lógico racional y alcanzan conclusiones siguiendo leyes universales del pensamiento definidas por la lógica (Winston, 1992).
- d. Sistemas que actúan racionalmente: amplían la racionalidad más allá de la lógica, incluyendo elementos como la incertidumbre, la autonomía y el cambio (Poole, Mackworth y Goebel, 1998).

Dentro de este marco de ideas, la IA se concibe como una herramienta complementaria de las actividades que hace el ser humano y no como su sustituto. Bajo este criterio, la estructura de la IA es instrumental, enfocándose en expandir las habilidades humanas en dominios donde las máquinas pueden desempeñarse mejor. Este enfoque ha impulsado investigaciones sobre cómo la IA puede colaborar con otros actores, como personas, grupos sociales y computadoras (Malone, 2018, citado en Marín García, 2019).

Según Palma y Marín (2008), la IA se divide en dos áreas principales: la ciencia de la IA y la ingeniería del conocimiento. La ciencia de la IA busca comprender los principios fundamentales de la inteligencia y desarrollar modelos teóricos que expliquen cómo funciona la cognición humana. Su objetivo final es crear una teoría computable del conocimiento humano, similar a las ecuaciones de Maxwell en el electromagnetismo, que permita predecir y comprender el comportamiento inteligente. Para lograr este objetivo, se enfoca en analizar la estructura del conocimiento humano, descomponiendo los procesos cognitivos en sus componentes más básicos, hasta llegar a las inferencias primitivas que los sustentan.

La ingeniería del conocimiento, por otro lado, se centra en el desarrollo de sistemas y aplicaciones prácticas de IA. Se basa en los modelos teóricos desarrollados por la ciencia de la IA para crear herramientas y soluciones que puedan resolver problemas del mundo real. Para ello, reescribe formalmente las inferencias identificadas por la ciencia de la IA en lenguajes comprensibles para las computadoras y luego programas operadores y sistemas que puedan ejecutar esas inferencias (Palma y Marín, 2008).

Ambas áreas de la IA se complementan entre sí y son esenciales para el avance del campo. La ciencia de la IA proporciona la base teórica necesaria para el desarrollo de sistemas de IA efectivos, mientras que la ingeniería del conocimiento pone en práctica sus resultados para crear soluciones tangibles. Juntas, ambas áreas contribuyen al objetivo final de la IA: comprender y replicar la inteligencia humana (Palma y Marín, 2008).

El avance de la tecnología y la ciencia desde esta perspectiva ha dado lugar a un tipo de transhumanismo tecnocientífico o cibernético, que se nutre de los aportes de la (IA), el software y la robótica. Que busca la compatibilidad entre el cuerpo humano y un sistema tecnológico complejo que abarca el aprendizaje automático (incluyendo el *deep learning* y el *reinforcement learning*), el razonamiento artificial (con áreas como la planificación, la programación, la representación del conocimiento, el razonamiento, la búsqueda y la optimización), la robótica (que integra control, percepción, sensores y actuadores), y la fusión de estas técnicas en sistemas ciberfísicos (Cortina, 2022).

Planteamientos que nos invita a una reflexión más profunda sobre la inteligencia en sus diversas formas y la naturaleza de la IA, lo que es fundamental para abordar los complejos desafíos éticos que conlleva la implantación del transhumanismo y posthumanismo. En este sentido, las IA, sintetizadas en al menos tres tipos, suscitan ciertos problemas éticos que necesitan ser abordados. El primer enfoque trata de una *inteligencia superior o superinteligencia*, que podría reemplazar a la humana y da lugar a conceptos transhumanistas y posthumanistas centrados en la *singularidad* (Cortina, 2022). Pioneros como John von Neumann (1966) vislumbraron este potencial, y autores como Vernor Vinge (1994) y Raymond Kurzweil (1999) han desarrollado visiones sobre la superación de la mente humana por máquinas inteligentes. La idea de la Singularidad, respaldada por avances tecnológicos, plantea la posibilidad de que los humanos transfieran su inteligencia a máquinas, generando una nueva especie.

Existen debates sobre la viabilidad científica de estas visiones. Algunos proyectan la creación de superinteligencias artificiales para este siglo; sin embargo, otros cuestionan esta perspectiva debido a la limitada comprensión del cerebro humano. Independientemente, estas ideas abren un vasto campo de interrogantes éticas. Los defensores del transhumanismo argumentan que es un deber moral buscar la mejora humana a través de la tecnología. Sin embargo, surge la pregunta sobre si trascender la humanidad hacia una supuesta superioridad implicaría una división en la inteligencia racional, potencialmente creando una jerarquía de dominación.

¿Qué ética guiaría a estas superinteligencias? Algunos proponen integrar valores en ellas, pero si estas máquinas llegaran a aprender y desarrollarse independientemente, podrían establecer sus propios valores, lo que plantea desafíos éticos fundamentales. La inversión masiva en la creación de estas entidades también suscita interrogantes sobre la priorización de recursos frente a la urgencia de abordar problemas sociales como la guerra, la pobreza y la injusticia. ¿Es ético canalizar recursos hacia la construcción de seres presuntamente perfeccionados mientras persisten estas crisis? (Cortina, 2022).

Otro enfoque dentro del campo de la IA se centra en la *inteligencia general*, la cual se caracteriza por la habilidad para solucionar problemas diversos. Esta inteligencia, por excelencia humana, es la base de la investigación en IA, donde las mentes más preclaras se esfuerzan por imbuir a las máquinas con esta inteligencia de amplio espectro, similar a la nuestra. La meta principal de la IA como esfuerzo científico es dotar a las máquinas de esta inteligencia general de nivel humano (Cortina, 2022).

La distinción entre IA fuerte y débil, introducida por John Searle en 1980, marcó un hito importante en esta temática, para él, la IA débil también conocida como IA limitada, se caracteriza por su capacidad de hacer ciertas actividades concretas, pero carece de un pensar autónomo. Se pueden mencionar algunos sistemas de IA limitada: como las computadoras con sus sistemas operativos, navegadores web y software

de oficina emplean IA débil para automatizar tareas como la búsqueda, organización de archivos y corrección de errores ortográficos. Los sistemas de los celulares que vienen ya con la tecnología basada en la IA que, entre otras cosas, permiten reconocer voz, responder preguntas y realizar tareas como programar alarmas o enviar mensajes de texto. Asistentes de Amazon: Alexa y otros asistentes de voz recurren a IA débil para controlar dispositivos inteligentes, reproducir música y efectuar compras. La IA débil es la tecnología más utilizada en la actualidad y se aplica en una amplia gama de sectores. Sin embargo, es importante tener en cuenta que la IA débil no puede pensar ni actuar de manera independiente y siempre requiere la intervención humana para su funcionamiento

La IA fuerte propone la posibilidad de que una máquina posea una mente y un pensamiento similar al humano. Sin embargo, Searle (1980) refuta esta idea argumentando que las máquinas carecen de la intencionalidad necesaria para otorgar significado a los símbolos. Aunado a la ausencia de un cuerpo físico que les impide adquirir el conocimiento de sentido común que nosotros obtenemos a través de nuestras experiencias corporales.

La pregunta clave es si es posible dotar a las máquinas de sentido común y conciencia moral, incluso sin un cuerpo humano. Aunque la financiación para este fin es considerable, hasta ahora no se ha logrado. Qué pasaría si las máquinas pudieran desarrollar una inteligencia general similar a la humana, surgirían interrogantes sobre su autonomía, dignidad, derechos y responsabilidades. Sin embargo, es improbable que se logre esta construcción, ya que las máquinas solo podrían simular emociones y valores, pero no sentir realmente, dado que la experiencia emocional está ligada al cuerpo. Por lo tanto, el riesgo de otorgarles autonomía plena en decisiones políticas y sociales parece poco probable (Cortina, 2022).

Un tercer enfoque dentro del ámbito de la IA conocido como *inteligencia especial* se caracteriza por centrarse en tareas específicas. Estos sistemas inteligentes exhiben una capacidad notable para superar ampliamente las habilidades humanas en tareas delimitadas, gracias a su manejo de grandes volúmenes de datos y algoritmos sofisticados. Estos

sistemas inteligentes se aplican en diversos campos, como la salud, la predicción climatológica, la eficiencia empresarial, la comunicación y el asesoramiento agrícola, utilizando técnicas como la minería de datos para encontrar patrones y predecir comportamientos futuros. Sin embargo, a pesar de su poder, sigue siendo un aspecto fundamental la dirección humana para aprovechar su potencial y garantizar el respeto a los derechos y la autonomía de las personas (Cortina, 2022).

Por otro lado, Rueda (2023) analiza la relación entre la inteligencia artificial (IA) y la ética, reflexionando sobre si la IA puede ayudarnos a tomar decisiones más éticas o mejorar nuestras capacidades morales. A partir de la discusión de varios capítulos del libro *Más (que) humanos. Biotecnología, inteligencia artificial y ética de la mejora* de Lara y Savulescu (2021) se plantea que limitar el potencial de la IA a la mejora moral individual es insuficiente. La IA tiene un gran potencial para apoyar la toma de decisiones éticas en contextos profesionales e institucionales, como en el ámbito de la salud o la formulación de políticas públicas. Al procesar grandes volúmenes de datos y considerar preferencias morales, la IA podría contribuir a diseñar políticas más éticas. Sin embargo, la decisión final siempre corresponderá a la persona.

Los retos éticos derivados de la IA actual exigen un compromiso activo para garantizar un desarrollo responsable y beneficioso para la sociedad. Entre ellos se encuentran la protección de los derechos laborales en un mercado en constante transformación, la preservación de la autonomía individual frente a sistemas de IA no autónomos y la organización de una sociedad donde los algoritmos desempeñan un papel cada vez más significativo. El imperativo ético fundamental en relación con la IA es garantizar que esta revolución tecnológica beneficie a toda la humanidad.

Ética de la IA: principios para un futuro responsable

Marín García (2019) nos presenta algunos principios éticos tanto para el diseño como el desarrollo de la IA, los cuales promueven la importancia de un uso y aplicación más consciente de esta tecnología. No obstante, conviene subrayar que estos principios no poseen la precisión de un marco normativo de comportamiento. En cambio, constituyen una serie de directrices que resaltan la importancia de respetar y fomentar los bienes humanos en cualquier contexto y acción vinculada a la IA.

Las Leyes de la Robótica de Asimov (1988) han sido fundamentales para establecer principios éticos en el ámbito de la IA, a pesar de ser una creación de la narrativa de ciencia ficción. Estas leyes establecen una guía ética para la interacción entre robots y seres humanos. La Primera Ley estipula que un robot no debe perjudicar a una persona en ningún caso. La Segunda Ley señala que un robot debe acatar las disposiciones impartidas por las personas, salvo cuando dichas disposiciones contradigan la Primera Ley. Por último, la Tercera Ley indica que un robot debe proteger su propia existencia, siempre y cuando esta protección no entre en conflicto con la Primera o la Segunda Ley (Marín García, 2019). Estas leyes proporcionan un punto de partida para reflexionar sobre la ética de la IA, pero es fundamental reconocer que son solo el comienzo. Abordar los complejos desafíos éticos que plantea la IA en el mundo real requiere un análisis más profundo y detallado. Siguiendo a Marín García (2019), es importante resaltar algunos criterios éticos en relación a la utilidad y el manejo de la IA, tomando como base las propuestas de diversos organismos internacionales y declaraciones emblemáticas.

- Respeto a la autonomía humana: la inteligencia artificial debe desarrollarse considerando sus límites y reconociendo la autonomía y los derechos fundamentales de los seres humanos, sin discriminación alguna. Esto implica resguardar el cuerpo, la mente y las capacidades intelectuales de la persona. Las decisiones tomadas por estos sistemas no deben afectar negativamente la autonomía individual ni limitar las posibilidades de desarrollo personal. Es fundamental

que las personas tengan control sobre sus datos personales y cómo se utilizan en sistemas de IA.

- **Transparencia:** se refiere a la capacidad de explicar y comprender su funcionamiento, incluyendo el razonamiento detrás de las decisiones que toman y las consecuencias que generan, esta información debe estar presentada en un lenguaje básico y natural. La transparencia es esencial para generar confianza en los sistemas de IA y permitir que las personas tomen decisiones informadas sobre su uso.
- **Confianza:** la confiabilidad recae en la importancia de crear algoritmos infalibles para funcionar con precisión. Esto implica desarrollar sistemas que consideren la posibilidad de ciberataques, fallas técnicas y errores humanos. La seguridad de los sistemas de IA es fundamental para proteger a las personas y evitar daños.
- **Equidad:** garantizar la justicia y evitar la discriminación requiere que todos los grupos afectados por la IA participen activamente en el desarrollo de sus sistemas. Esto ayudará a garantizar que los sistemas se diseñen de manera justa, equitativa y que no discriminen a ningún grupo en particular. Los datos utilizados para entrenar sistemas de IA deben ser cuidadosamente seleccionados y procesados para evitar sesgos y discriminación. Es importante monitorear el uso de sistemas de IA para detectar y abordar cualquier problema de discriminación o injusticia. La justicia y la no discriminación son principios éticos fundamentales que deben guiar el desarrollo y uso de la IA para garantizar que todos los individuos y grupos sociales se beneficien de sus avances (Marín García, 2019).

Incorporar algunos de estos principios éticos representa un desafío significativo para una sociedad que busca ir más allá de las premisas propuestas por el transhumanismo y el posthumanismo. Aunque la adopción de estos criterios éticos es importante, no garantiza por sí sola un uso seguro y responsable de la IA ni la viabilidad del proyecto transhumanista tecnocientífico. Por lo tanto, entran en juego fuerzas sociales, políticas y

económicas en el debate sobre el impacto que estos principios tendrán en nuestras vidas futuras. Por lo que, se hace necesario reflexionar sobre cómo estas fuerzas moldearán nuestro porvenir y trabajar para asegurar que la implementación de la IA esté alineada con valores éticos compartidos y el bienestar de toda la humanidad.

Hacia una ética interdisciplinaria e intersubjetiva del transhumanismo tecnocientífico e inteligencia artificial

Uno de los dilemas éticos centrales que nos plantea el transhumanismo y el posthumanismo radica en la distinción entre propuestas respaldadas científicamente por experiencias actuales o posibles y aquellas que carecen de este fundamento. Se necesita de una ética que pueda dar cuenta de los desafíos que lleva implícito los avances tecnocientíficos y dar respuesta responsable a esa realidad.

Lo que demanda el reconocimiento abierto de la falta de legitimidad científica al pasar de un tipo de propuesta al otro, como si todas tuvieran igual validez científica. Esto requiere el ejercicio de la crítica racional, que es el camino ilustrado hacia un conocimiento confiable. Es esencial considerar que la confianza es un pilar fundamental para el funcionamiento de las sociedades, ya sea en ámbitos como las tecnociencias, la política, la economía o cualquier otra esfera de la vida en comunidad (Cortina, 2022).

Un segundo dilema ético surge al considerar si la presentación a la sociedad global de ciertos proyectos posthumanistas y transhumanistas, responde más a un uso ideológico de las tecnociencias (dada la influencia económica y política que estas conllevan) que a un verdadero aporte al bienestar social. En este contexto, se justifica aplicar una crítica ideológica para contrarrestar las pretensiones de los grandes centros de poder que pretenden usar, para sus propios beneficios, los aportes de la tecnociencia en este aspecto. Sin embargo, en términos éticos son las personas quienes deben guiar el futuro de la humanidad mediante un diálogo abierto basado en información veraz y colocar límites normados

a la praxis tecnocientífica. El interés genuino en el entendimiento mutuo y el consenso debe orientar las tecnociencias hacia la creación de una sociedad emancipada, liberada de ideologías perversas.

Sin duda alguna, la revolución 4.0 forma parte de nuestro mundo actual y nos brinda herramientas invaluable para construir una sociedad que, a través de estas nuevas posibilidades, se pueda mejorar la calidad de vida de todos. Es imperativo, sin embargo, el desarrollo de una ética interdisciplinaria fundamentada en la intersubjetividad e interculturalidad que enfatice la importancia de la convivencia comunitaria, promoviendo el bien común, reconociendo que somos diferentes culturalmente y que es necesario preservar las identidades de los pueblos y nacionalidades. Esta ética debe ser capaz de valorar los avances científicos y tecnológicos sin que estos perjudiquen a la naturaleza ni a los seres vivos, ni que busquen aniquilar a la humanidad en pos de un mundo dominado exclusivamente por robots y/o homo cyber dirigidos por un tipo de inteligencia artificial bajo una visión poshumanista excluyente.

Esta ética, que se fundamenta en la interdisciplinaria y la intersubjetividad, se basa en el respeto a la autonomía individual, la del otro como ser social parte de una comunidad y los aportes de las diversas disciplinas a la sociedad del conocimiento. Se procura reconocer la autonomía que tienen las personas y las comunidades para decidir sobre la adopción o el rechazo de tecnologías que puedan afectar sus territorios, recursos naturales o culturas, empoderándolas para proteger y promover la biodiversidad desde criterios éticos y responsables.

Conclusiones

El transhumanismo y la IA representan avances tecnocientíficos que plantean profundas implicaciones éticas para la humanidad. Este análisis ha destacado la necesidad de abordar estos desarrollos desde una perspectiva crítica, consciente y responsable, reconociendo que el conocimiento científico institucionalizado y los avances tecnológicos no

son imparciales y que su progreso está influenciado por intereses específicos. Si bien estas innovaciones han generado beneficios significativos en diversos sectores, su implementación debe orientarse hacia la equidad, la sostenibilidad y el respeto por los valores humanos fundamentales.

La reflexión ética sobre el transhumanismo tecnocientífico y la inteligencia artificial exige analizar sus efectos en la esencia y naturaleza humanas, cuestionando las promesas de superinteligencia y la búsqueda de la singularidad. Es imprescindible establecer límites morales claros que preserven principios esenciales como la autonomía, la justicia, la seguridad y la transparencia, asegurando que estas tecnologías no comprometan la dignidad humana ni exacerbem desigualdades sociales.

Por ende, la responsabilidad por el diseño y el impacto de estos sistemas recae en quienes los desarrollan y promueven. Se requiere un enfoque interdisciplinario que articule la ética, la tecnología y las necesidades sociales que garantice que las aspiraciones de un mundo posthumanista estén guiadas por valores que prioricen el bienestar humano y la defensa de la vida en el planeta.

Referencias bibliográficas

- Abbate, S. (2022). Transhumanismo y gnosis: un paralelismo. *Scientia et Fides*, <https://doi.org/10.12775/SetF.2022.010>
- Alba, J. y Calero, M. (2023). Transhumanismo: Una perspectiva desde las diferencias biológicas basadas en la desigualdad de riqueza. *Revista de Filosofía*, 40(105), 119-130. <https://doi.org/10.5281/zenodo.7857567>
- Apel, K. O. (1985). *La transformación de la filosofía*. Taurus I.
- Asimov, I. (1988). *Yo, robot*. Edhasa.
- Bellman, R. E. (1978). *An Introduction to Artificial Intelligence: Can Computers Think?* Boyd & Fraser Publishing Company.
- Binet, A. y Simon, T. (1916). *The development of intelligence in children (the Binet-Simon Scale)*. The Training School.
- Bostrom, N. (2002). Riesgos existenciales: Análisis de escenarios de extinción humana y peligros relacionados. *Journal of Evolution and Technology*.
- Bostrom, N. (2005). *Valores transhumanistas*. Universidad de Oxford.

- Camargo Brito, R. y Ried, N. (2024). Los desafíos del transhumanismo a la bioética. *Acta Bioethica*, 30(1), 9-17. <https://doi.org/10.4067/S1726-569X2024000100009>
- Castillo, P. (2023). Does there exist a need for a 'new' educational ideal of rationality? The crossroads between transhumanism and Israel Scheffler's conception of critical thinking. *Encyclopaideia*, 27(66), 49-61. <https://doi.org/10.6092/issn.1825-8670/16490>
- Comisión Europea, Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial. (2019). Directrices éticas para una IA fiable: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60423
- Cortina, A. (2022). Los desafíos éticos del transhumanismo. *Pensamiento. Revista De Investigación E Información Filosófica*, 78 (298 S. Esp), 471-483. <https://doi.org/10.14422/pen.v78.i298.y2022.009>
- Díaz Alzate, M. V., González-Cortes, J. H. y Gómez Tabares, A. S. (2022). Papel de las tecnologías digitales en la consolidación del posthumanismo como nueva forma de humanismo. *Revista de Filosofía*, 39(102), 346-362. <https://doi.org/10.5281/zenodo.7045669>
- Diéguez, A. (2017). *Transhumanismo. La búsqueda tecnológica del mejoramiento humano*. Herder
- Dobre, C. E. y Pavón, R. G. (2024). Singularidad individual versus "singularity". Una crítica al transhumanismo desde el pensamiento de Søren Kierkegaard. *Tópicos*, 69, 389-420. <https://doi.org/10.21555/TOP.V690.2503>
- Ferry, L. (2017). *La revolución transhumanista. Cómo la tecnología y la uberización del mundo van a transformar nuestras vidas*. Alianza Editorial
- FM-2030 (1989). *Are you a transhuman? Monitoring and stimulating your personal rate of growth in a rapidly changing world*. Warner Books.
- Garcés Castellote, E. y Jiménez Rodríguez, M. (2016). Transhumanismo: cómo el mejoramiento humano cambiará el cuidado: un análisis desde la teoría general del déficit de autocuidado. *Ene*, 10 (3). <https://bit.ly/4ngPVwv>
- Gardner, H. (1994). *Estructuras de la mente: La Teoría de las inteligencias múltiples*. FCE.
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. Basic Books.
- Gaviria, J. L. (2024). ¿Transhumanismo 'contra' educación? *Teoría de la Educación*, 36(2), 1-23. <https://doi.org/10.14201/teri.31762>
- González Ramírez, Y. J. (2024). Transhumanismo: una versión de la «Transfinitud» garcíabaquiana. *Bajo Palabra*, 37, 395-412. <https://doi.org/10.15366/bp2024.37.034>

- Habermas, J. (1984). Ciencia y técnica como “ideología”. Tecnos.
- Hernández, R. M., Fernández, A. V., Vilca, L. A. P., Mallqui, N. C. y Fuster-Guillén, D. (2023). Between bioethics and transhumanism: Remarkable perspectives and acute challenges in transit to a possible dystopian future. *Revista Relações Internacionais do Mundo Atual*, 6(39), e06297. <https://doi.org/10.21902/Revrima.v6i39.6297>
- Hugues, J. (2004). *Citizen Cyborg: why democratic societies must respond to the redesigned human of the future*. Westview Press
- Hurtado, J. (2005). *Metodología de la Investigación Holística*. Editorial Magisterio.
- Iberdrola (2019). *¿Somos conscientes de los retos y principales aplicaciones de la inteligencia artificial?* Innovación.
- Kurzweil, R. (1999). *La era de las máquinas espirituales*. Planeta.
- Kurzweil, R. (2005). *The Singularity is near: When humans transcend biology*. Viking Press
- Kurzweil, R. (2012). *La singularidad está cerca*. Lola Books.
- Lara, F. y Savulescu, J. (Eds.). (2021). *Más (que) humanos: Biotecnología, inteligencia artificial y ética de la mejora*. Tecnos.
- López Fernández, D. (2024). Lo posthumano en la era del Antropoceno. Transhumanismo, hibridación y agencia no-humana en la trilogía Arrebató, de Enki Bilal. *Revista de Filología de la Universidad de La Laguna*, 48, 81-106. <https://doi.org/10.25145/j.refull.2024.48.04>
- Lovelock, J. (1983). *Gaia, una nueva visión de la vida sobre la tierra*. Hermann Blume.
- Malone, T. (2018). *Superminds: The Surprising Power of People and Computers Thinking Together*. Little, Brown & Company.
- Marín García, S. (2019). Ética e inteligencia artificial. *Cuadernos de la Cátedra CaixaBank de Responsabilidad Social Corporativa*, (42). <https://dx.doi.org/10.15581/018.ST-522>
- McCarthy, J., Minsky, ML, Rochester, N. y Shannon, CE. (2006). Una propuesta para el proyecto de investigación de verano de Dartmouth sobre inteligencia artificial, 31 de agosto de 1955. *Revista AI*, 27(4), 12. <https://doi.org/10.1609/aimag.v27i4.1904>
- Mehlman, M. (2009). Will direct evolution destroy humanity, and if so, what can we do about it? Saint Louis University. *Journal of Health Law & Policy*, 3(1). <https://bit.ly/3VKFBkP>
- Méndez Reyes, J. y Padrón Medina, A. (2024). El transhumanismo: una mirada crítica desde la filosofía y la teología. [Manuscrito no publicado]. *Encuentros. Revista de Ciencias Humanas, Teoría Social y Pensamiento Crítico*.

- Méndez Reyes, J. y Padrón Medina, M. (2022). *Consciencia psicológica e inteligencias múltiples en el contexto universitario: Una mirada fenomenológica*. Universidad Nacional Experimental “Rafael María Baralt”.
- Minsky, M. (1960). *Pasos hacia la inteligencia artificial*. <https://bit.ly/42jgxFj>
- Moravec, H. (1998). ¿Cuándo el hardware de la computadora coincidirá con el cerebro humano? *Revista de Evolución y Tecnología*, 1(1), 1-12. <https://bit.ly/3IJvAl5>
- Neumann, J. von (1966). *Theory of Self-Reproducing Automata*. University of Illinois Press.
- O’Conor, A. (1999). *Desarrollo de la Inteligencia*. S.A de C.V y Alfaomega Grupo Edito.
- Palma, J. y Marín, R. (2008). *Inteligencia artificial*. McGraw-Hill.
- Pavón, S. A. y Macip, C. G. (2023). Futuro y transhumanismo: ¿culminación del proyecto moderno de control racional o fin de la educación? *Perspectiva abierta*, 14(32), 1-14. <https://doi.org/10.23924/OI.V14I32.563>
- Pearce, D. (2004). The hedonistic imperative. <https://bit.ly/475EmmI>
- Piaget, J. (1985). *La equilibración de las estructuras cognitivas*. Siglo XXI Editores.
- Ponce, V. y Sanmartín, E. (2010). *Las inteligencias múltiples y su relación con el aprendizaje en niños de educación básica*. Universidad de Cuenca.
- Poole, D., Mackworth, A. K. y Goebel, R. (1998). *Computational Intelligence: A Logical Approach*. Oxford University Press.
- Rueda, J. (2023). ¿Automatizando la mejora moral humana? La inteligencia artificial para la ética. *Daimon*, 89, 199-209. <https://doi.org/10.6018/daimon.508771>
- Russell, S. y Norvig, P. (2016). *Artificial Intelligence: A Modern Approach* (3.ª edición). Pearson Education Limited.
- Sabino, C. (2004). *Metodología de la Investigación*. Editorial Logos.
- Sagan, D. y Margulis, L. (1984). *Gaia y filosofía*. University of Notre Dame Press.
- Sánchez Salazar, D. V. (2024). El derecho al autodiseño en el transhumanismo: ¿un derecho humano? *Revista de la Facultad de Derecho y Ciencias Políticas*, 54(141).
- Sandberg, A. (2001). Morphological freedom - Why we not just want it, but need it. En More, M. y Vita-More, N. (Eds.), *The transhumanist reader: Classical and contemporary essays on the science, technology, and philosophy of the human future*. Wiley-Blackwell.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424. <https://doi.org/10.1017/S0140525X00005756>

- Serra, M. Á. (2022). Mejora humana y diversidad funcional: Preocupaciones éticas de las tecnologías emergentes y el transhumanismo. *Método*, 12, 20676. <https://doi.org/10.7203/método.12.20676>
- Vaccari, A. (2016). Aporías transhumanistas: Ideologías de la tecnología en el proyecto del auto-diseño humano. *Quadranti. Rivista Internazionale di Filosofia Contemporanea*, 4(1-2), 1-18.
- Vernor, V. (1994). La singularidad tecnológica. *Revista electrónica AXXO*.
- Winston, P. H. (1992). *Artificial Intelligence*. Addison-Wesley.

La bioética y el cambio climático: uso humano de plantas medicinales contaminadas¹

Fredi Portilla Farfán
Universidad Politécnica Salesiana
fportilla@ups.edu.ec
<https://orcid.org/0000-0002-1234-9758>

Prefacio

El clamor de la Tierra

Desde la nada surgí, cuando aún no existía forma ni esencia, cuando el vacío flotaba sin rumbo en la inmensidad de la nada. Un origen remontado a los albores del tiempo, cuando reinaba el caos y el orden aún no hallaba su lugar. En aquel vastísimo universo, compuesto por diminutas partículas, fui solo una más, en constante cambio. Mi existencia quedó ligada a la órbita de una estrella, cuyos destellos me han brindado luz y calor ininterrumpidamente. Atada a una red invisible, conocida por los sabios como la red cuántica, giro sin descanso, marcando con mi movimiento el ciclo del día y la noche. En mi esencia, tres cuartas partes son agua, y el resto, la tierra firme que da nombre a mi ser y refugio a innu-

1 La elaboración de este artículo está basada en las tesis de pregrado de los ahora exalumnos ambientales: Darwin Abad, Bryam Poluche, Diego Nieves y Hernán Solano. A quienes se les agradece por su cooperación.

merables seres diminutos y gigantes. Desde siempre fui un hogar común, un manantial inagotable de vida. Sobre mi suelo crecieron plantas que ofrecieron alimento al ser humano, mientras que las semillas, esparcidas por el viento y transportadas por los insectos, dieron origen a un mundo de diversidad. Árboles majestuosos alimentaron a los animales, praderas cubrieron mis llanuras y flores silvestres adornaron mis acantilados. Mis jugosos frutos, mis hierbas aromáticas y mis dorados granos fueron sustento para los seres humanos. Hoy me estremezco con la furia de los volcanes y lloro con la fuerza de tormentas y huracanes. Mis montañas tiemblan, mis suelos se agrietan, y el dolor se expande en cada rincón de mi ser. Me desangro con cada herida causada por la codicia humana, que año tras año arrasa mis bosques, envenena mis ríos y hurga en mis entrañas en busca de riquezas. En nombre del desarrollo, han convertido mis campos fértiles en terrenos áridos y han contaminado mis aguas. ¡Basta ya de agresiones constantes! Mi paciencia se desvanece como el agua de un río que se seca lentamente. La infertilidad que el hombre ha sembrado en mis tierras se traducirá en hambre y escasez. La contaminación que se ha propagado sin medida será una deuda que las demás especies le cobrarán sin tregua. Y cuando el daño sea irreversible, cuando todo lo que ignoró finalmente le alcance, entonces lamentará su descuido. Pero para entonces, su grito de todo lo que ignoró finalmente le alcance, entonces lamentará su descuido.

Introducción

El cambio climático, impulsado tanto por procesos naturales como por la acción humana a lo largo de los siglos, se ha acelerado de manera alarmante en el último siglo. Aunque algunos aún niegan su existencia, la mayoría de la población lo percibe y lo vive en su día a día. Los ciclos evolutivos de las especies se ven afectadas y las temperaturas han alcanzado niveles extremos nunca registrados. Al mismo tiempo, los humedales y los bosques desaparecen a un ritmo preocupante, mientras que en áreas densamente pobladas el agua escasea inexorablemente. A esto se suma la contaminación de las pocas fuentes de agua restantes por parte de diversas

industrias. El aire se vuelve irrespirable, gravado por nubes amenazadoras de lluvia ácida que se ciernen sobre los campos y ciudades. El bosque de cemento reemplaza la cobertura natural del suelo, privando de la luz natural a las casas pequeñas de habitación tradicional.

La prensa habla de violencia generalizada promoviendo el miedo colectivo a caminar, a saludar al vecino, a ser solidario con el desconocido. Cada ciudadano, a medida de su bolsillo, irrumpe en la construcción de muros altos para sus casas, barrotes para sus puertas y ventanas, sistemas electrónicos de video de seguimiento continuo, aunque éste atente a la intimidad de las personas. Los supermercados venden más papel y plástico de frente a una necesidad creada de sanidad y estatus social, mientras que se olvidan de la necesidad real de la alimentación humana y animal. Niños, ancianos, madres y padres, buscan sobras en los basureros mientras que, perros callejeros esperan su oportunidad para hacer lo mismo. La seguridad alimentaria es un concepto esencial que debería beneficiar a todas las naciones, pero, lamentablemente, parece estar al alcance principalmente de los países ricos.

En los países del cuarto y quinto mundo, las altas tasas de desnutrición y las enfermedades relacionadas continúan siendo un grave problema. Mientras que la seguridad alimentaria brinda ventajas significativas, como la reducción del hambre y la mejora de la salud pública, los peligros asociados a su inadecuada gestión son alarmantes. La llegada de alimentos híbridos a través de tratados de libre comercio ha favorecido a quienes ya tienen acceso a una variedad de recursos, a la vez que ha socavado la soberanía alimentaria de pueblos que tradicionalmente han practicado el uso sostenible de sus frutos y semillas nativas. Esta situación no solo compromete la diversidad agrícola, sino que también afecta a la población local, que se ve forzada a depender de productos menos nutritivos y más costosos, exacerbando las disparidades en salud y alimentación.

La salud pública es solo un cliché de onerosas campañas políticas, mientras que en los escasos hospitales campea la muerte por doquier, a pesar de que los sanitarios, hacen lo posible por evitarla. La educación se

ha convertido en un privilegio reservado para unos pocos, mientras que la masificada educación de baja calidad, presente en muchas comunidades, actúa como una forma de control para las masas. Esta desigualdad en el acceso a la educación resalta los privilegios de quienes tienen más recursos, quienes pueden acceder a instituciones que ofrecen una formación real y de excelencia. En contraste, los menos favorecidos se enfrentan a una educación disfrazada de calidad, que rara vez les brinda las oportunidades necesarias para mejorar su situación. Así, la falta de acceso a una educación auténtica perpetúa ciclos de pobreza y limita el desarrollo de potenciales talentos en comunidades marginadas.

Los sistemas imperantes imprimen su lógica productiva en desmedro de la ética del desarrollo humano, a quien consideran un tornillo más de su enorme fábrica. Algunos inhumanos, encapsulados en sus religiones, conspiran con los poderes de turno para consolar a los necesitados, sobre invisibles prados floridos de futuros mundos distantes del sufrimiento actual. Se olvidan de los mensajes centrales de sus fundadores de paz, fraternidad, equidad, amor y compromiso con los menos favorecidos. Si aparecieran en la actualidad: ¿qué diría el Cristo que predicó sobre el Dios del amor universal?, ¿qué diría Mahoma que habló del Dios de la misericordia?, ¿qué diría Buda el buscador de la felicidad plena?, ¿qué diría Abraham el padre de la fe? ¿qué diría Gandhi el incansable luchador por la no violencia? y así muchos hombres y mujeres que fueron ejemplo de lucha, hicieron historia y dieron su vida por los derechos humanos. El padre del nihilismo, Friedrich Nietzsche, con su frase “Dios ha muerto”, invita al atrevimiento de decir, que pocos poderosos humanos, cada día, matan al Dios verdadero, para vender a otro dios que es falso pues representa la inequidad, el consumismo, la violencia, la envidia, etc., que mancilla a la mayoría de una población que ha caído en la desesperanza y desilusión. Todo esto suma al deterioro de las relaciones humanas, que se expresa en hechos de violencia, deshonestidad, corrupción y daño a la naturaleza y por ende suman fuerzas para que actúe el cambio climático, por lo cual es menester responsablemente proponer alternativas para luchar en algunos frentes y ser resilientes en otros. Uno de ellos es la

agricultura urbana como fuente de la seguridad y soberanía alimentaria conjugada con la educación ambiental participativa.

La utilización de plantas medicinales para curar distintas enfermedades es considerada una práctica ancestral que, durante gran tiempo, fue el único recurso empleado para calmar enfermedades (Aguirre, 2014). Se estima que cerca de 35 000 especies de plantas presentan potencial medicina (Bermúdez *et al.*, 2005). Las plantas de la región Sierra son las más conocidas (Aguirre, 2014). Los contaminantes presentes en el medio ambiente se han incrementado drásticamente debido a actividades antropogénicas principalmente, representando una amenaza para los organismos que la habitan, incluyendo las plantas medicinales (Martín, 2017). Esta contaminación ha provocado cambios en la cobertura vegetal y la atmósfera terrestre y producido una interacción de los metales pesados y plantas. La exposición a estos contaminantes puede retener estos metales y con esto ingresar a la cadena alimenticia (Carrera, 2016).

Las plantas utilizadas para elaborar medicamentos son muy importantes para la medicina tradicional y moderna (Akerele, 1994). En Ecuador, el uso de estas plantas es común debido a la gran diversidad de plantas que existe (De la Torre *et al.*, 2008), pueden conservarse por mucho tiempo y emplearse en la elaboración de jarabes, jabones, té, pomadas, entre otros (Perez, 2008). Después de un tiempo se desarrollaron otro tipo de tecnologías que formaron una agricultura avanzada y permitió que la región evolucione culturalmente (Almeida, 2000). La flora medicinal ya sea como compuesto puro o como prototipos, se ha convertido en una fuente importante de fármacos (Maregesi y Mtewa, 2007). Se estima que un 30 % de los fármacos comercializados se consideran como fito medicamentos (Oliveira *et al.*, 2012). La fitoterapia se ha convertido en una importante y eficiente solución para enfermedades tanto crónicas como moderadas (Abe y Ohtani, 2013). Una base fundamental para el éxito de la medicina tradicional es que el conocimiento se trasmite de generación en generación (Pabón *et al.*, 2017). Muchas de estas prácticas han contribuido al uso sostenible de los recursos naturales y desencadenado

una diversidad biocultural (Gómez-Baggethun *et al.*, 2010). La medicina tradicional siempre está abierta a nuevas perspectivas (Onofre, 2013), pero a pesar de que la medicina tradicional puede ser muy efectiva cuando se conoce la dosis y las plantas a utilizar, a veces, personas que desconocen de estos factores pueden aumentar el riesgo de toxicidad (García, 2004). En Colombia, por ejemplo, en el caso de la diabetes se utilizaba plantas para regular el azúcar, en el caso de la anemia se utilizaba el consumo del romero, la yerbabuena, entre otros (Cardona y Rivera, 2012). Otro ejemplo es tratar infecciones de las vías respiratorias en niños y trastornos de sueño (Yanchaguano y Pérez, 2019) y para el beneficio ocular. La medicina tradicional es considerada una práctica profesional por la OMS por su aporte a la salud (Acosta y Salcedo, 2020). Por ejemplo, frente a la resistencia que genera el cuerpo humano a los antibióticos, se han investigado otros tratamientos que emplean extractos de plantas y metabolitos secundarios con eficiencia (Lara y Sánchez, 2018).

La bioética y la idea del cambio

En una trivial entonación de un coro sin integrantes, la voz bioética de los que no tienen voz nunca fue escuchada por quienes en su momento debieron hacerlo sin ambigüedades. Nadie supo decir en su entorno cuán difícil fue enrumbar el mundo nacido del hambre y la miseria, del desencuentro entre la necesidad y la opulencia. Fueron las brisas de los atardeceres del otoño del milenio pasado que soplaron la comarca de las verdes colinas donde nacieron los niños que pronto fueron hombres luchadores de la causa por los más pobres en medio de un cambio climático. Ese tenue viento que no abriga ni enfría, bordeó la mente de los pensadores, y navegó por lo sideral de la idea del cambio sin límite de ideología ni religiones, y aparcó en la levedad de lo inhóspito de lo lejano del ser.

Aquí tuvo nacimiento la idea con base en la necesidad colectiva y en la incertidumbre singular del ser humano que trasciende la causa sin hallar el verdadero origen de la circunstancia. No fluye el pensamiento en medio de un caudaloso río cargado de ilusiones que se dirige a un bioterio

de desigualdad en las reparticiones. El pensamiento se anquilosa cuando aún no ha nacido el lazo de la perplejidad uniendo a la certeza que mucho le falta para ser considerada verdadera. Como un tronco sin hojas que permanece vivo en medio de un desierto donde el agua se evapora sin haber llovido, así se esfuma el pensamiento cuando la mente no trabaja en función de la razón razonada sino aquella de la especulación consumada.

Cuando el bioético pensador se aproxima a lo incierto del camino sinuoso de la incertidumbre tras la pantalla del descubrimiento, se conmueve entonces ante lo acontecido y exclama para su interior en forma compungida, diciendo si ahí estuvo y nunca lo vi en su destino. La idea sin origen y sin fin habrá de permanecer en aquello que considera su nido, el cerebro humano compuesta de masa gris y neuronas que hacen brillo al tocar sus terminaciones que son el inicio de otras. La idea del cambio climático ni se mueve ni se traslada en el espacio sin tiempo si el cerebro no lo permite, ahí se queda detenida como gélida lápida de cementerio que hace guardianía de algo que se descompone con el pasar del tiempo.

La idea nacida del ser que lo promueve es parte de él y al mismo tiempo no es, porque una vez producida se independiza, aunque territorio común estrene como morada del pensamiento que es abstracto sin dimensión y sin dirección. La libertad aparece entonces desde la idea concebida en un ente material llamado cerebro, que es parte de un ser de dimensiones conocidas, más la esencia que provoca la idea es libre para vagar por el universo de las mil y una noches, esa es la libertad en sí, aquella que se vuelve transhumante, victoriosa de sus victorias y de sus sueños soñadora. Libérrima es la libertad al llevar la idea de un lado para el otro, donde no hay semáforo que la detenga ni policía alguna que la amoneste. No hay materia alguna que la detenga ni voz taciturna que la increpe. Solo acatará la voz de quienes no tienen voz, porque de cierto es que no existe, porque aquellos tampoco lo están, pues yacen dormidos en el campo florido de la demagogia, en la esperanza de un más allá que cada vez está más lejano del más acá.

Cuando iniciado el sendero como investigador, jamás volteé la vista atrás cuando avanzaba con premura. Nunca lo hice porque mi mirada permanecía fija en el horizonte, allí donde la realidad se presentaba, el cambio climático, impreso como una pintura de Goya, un manantial nocturno de conocimiento que emergía sobre la duda de aquellos que, en la penumbra, se inclinan como marionetas temerosas de espectros inexistentes, cuyas siluetas se alojan en frágiles mentes humanas. La esencia del razonamiento siempre fue originada desde el porvenir de un ambiente sano para convertirse en pasado, donde lo inmaterial, convertido en palabra, se instala en el santuario de las joyas intocables de la naturaleza. Aquella idea, sumisa antes de ser pronunciada, deja de ser pensamiento aislado al convertirse en frase, donde las letras no alcanzan a plasmar con precisión lo que la mente concibió en su ética natural. Un resplandor lunar surge desde el intelecto, que no encuentra espacio dentro del cerebro y, sin embargo, alberga allí ideas que no son incubadas, sino que se desvanecen como gases nobles en la vastedad del cosmos, un cambio llegado, un clima alterado. La fluidez de lo intangible se transforma en la solidez de lo perpetuo, en un creador de ilusiones, en una corriente de emociones, en una melodía interior del pensador que solo él comprende, aunque quienes escuchan su bioética voz lo aclamen sin haber captado su esencia.

Un viajero prisionero de sus sentimientos que no se detiene a mirar atrás, porque su mente bulle con el sueño de lo alcanzable y con la interrogante de cómo lograrlo sin haber recorrido lo suficiente. Su obsesión se instala en el pensamiento, atrapado en su delirio de grandeza, repitiéndose que la verdad absoluta no existe y que solo la acumulación de certezas lo acerca a su meta, aunque también las incertidumbres, que enseñan a través del error, lo conducen al mismo destino. El caminante del poeta Antonio Machado, traza su camino al andar, nunca se detiene a mirar hacia atrás para no quedar atrapado en el ayer, contemplando lo que fue y lamentando lo que nunca sucedió. No es un viajero errante, sino un visionario, alguien que, en la práctica de los valores, ve lo malo y lo bueno que el ser humano hace en la naturaleza y provoca un cambio. Aquel que, paso a paso, valora su experiencia y deja que sus ideas fluyan

en el incesante devenir, las recicla, las reutiliza y las transforma en algo nuevo, con la convicción de que la utopía deja de serlo en el instante en que pisa tierra firme, nunca sobre el polvo que el viento disipa, sino sobre la roca que graba sus huellas.

La palabra cobra vida en el instante mismo en que es pronunciada, pues transforma la idea incandescente en prosa o verso abundante que, al final, expresa lo que el viajero desea. Ahora la palabra se vuelve tangible, legible en signos y transmisible en distintos lenguajes. Lo concebido se materializa, mientras lo inexplorado guarda en el infinito de los universos paralelos. No mirar atrás se convierte, así, en una necesidad que brota del conocimiento, para prevenir errores venideros o, simplemente, para abstraerse del pasado y justificar un presente injusto. Esa misma actitud ha llevado a la humanidad, salvo raras excepciones, a olvidar que la bioética surge del principio del derecho del respeto mutuo, rigiéndose por normas universales que fomentan el diálogo continuo entre las sociedades, y un solo cambio en ese camino, altera el clima con todas las consecuencias vividas y previstas.

Recordando los principios bioéticos: beneficencia, no-maleficencia, autonomía y justicia, cuyos conceptos es menester recordarlos:

- Beneficencia

Principio que exige la realización de acciones que produzcan beneficios. Es el más evidente, por ejemplo, el profesional se forma y educa con el fin de hacer el bien, no solo para sí, individualmente, sino para la sociedad en su conjunto.

- No-maleficencia

Inspirado en el adagio hipocrático *Primum non nocere* o ‘antes que nada, no causar daño’, este principio establece que ningún ser humano debe emplear sus conocimientos o posición para ocasionar perjuicio a otro ser humano.

- Autonomía

Este principio se relaciona con la capacidad del ser humano para tomar decisiones de acuerdo con su voluntad.

- Justicia

Corresponda a la noción de que toda persona, simplemente por existir, tiene la misma dignidad sin importar sus circunstancias, tiene derecho a una justicia aplicada con respeto.

Vistos estos principios bioéticos y para enrumbar la presente ponencia desde la bioética y el cambio climático, he tomado como ejemplo la experimentación llevada por Abad y Poluche (2018), que plantea la búsqueda de Pb (plomo) y CO₂ (dióxido de carbono) en cuatro especies de plantas manzanilla (*Chamaemelum nobile*), toronjil (*Melissa officinalis*), romero (*Rosmarinus officinalis*) y ortiga (*Urtica dioica*), cultivadas en la ciudad de Cuenca y una comparación con plantas medicinales adquiridas en tres mercados de la misma ciudad.

Experimentación, metodología y recursos

Diseño experimental

La siembra de las cuatro especies medicinales: manzanilla, toronjil, romero y ortiga, se llevó a cabo en tres zonas de la ciudad de Cuenca: manzanilla (*Chamaemelum nobile*), toronjil (*Melissa officinalis*), romero (*Rosmarinus officinalis*) y ortiga (*Urtica dioica*), se llevó a cabo en tres zonas de la ciudad de Cuenca - Ecuador (norte, centro y sur). Estadísticamente se utilizó el diseño de bloques al azar. En cada zona se ubicaron doce macetas, distribuidas en tres tratamientos y en cuatro bloques; en cada bloque se cultivaba una especie de planta medicinal. Los tratamientos estuvieron separados por un metro. Los tratamientos estuvieron separados por 1 metro. Se utilizó humus de lombriz como sustrato. Para realizar las

comparaciones, se tomaron muestras de las mismas especies de plantas medicinales en los tres mercados de la ciudad de Cuenca.

Concentración de carbono y plomo

Para evaluar el carbono secuestrado por las plantas medicinales en los tres sectores de la ciudad de Cuenca y en los tres mercados de las zonas Norte, Sur y Centro, se aplicó la técnica planteada por Osinaga (2014). Esta metodología detalla el proceso para medir la biomasa de vegetación no arbórea, como hierbas, arbustos y gramíneas, en diversos tipos de uso del suelo, incluyendo forestal, agrícola y pastizales. Asimismo, se establece la relación entre la biomasa y el carbono: 1 tonelada de biomasa equivale a 0.5 toneladas de carbono, y 1 tonelada de carbono representa 3.67 toneladas de CO₂. Para implementar esta guía, se recolectaron todos los vegetales (incluidas las raíces) que fueron pesados en su estado fresco. Posteriormente, se tomó una submuestra de este peso fresco que se guardó en bolsas plásticas debidamente etiquetadas antes de ser trasladadas al laboratorio de la Universidad Politécnica Salesiana. Allí, las submuestras se pesaron y se secaron en un horno a 60 °C hasta alcanzar un peso constante. Este proceso permitió establecer la relación entre la masa seca y húmeda, así como la cantidad de carbono presente. Para analizar la concentración de plomo en las plantas medicinales, se recolectaron muestras tanto de las plantas cultivadas en los tres sectores de Cuenca como de aquellas adquiridas en los mercados de las zonas Norte, Sur y Centro. En total, se recogieron 12 muestras de cada sector, sumando 72 en total, que se conservaron en bolsas herméticas. El procedimiento operado con base en estudios realizados por Ali y Al-Qahtani (2012) y con el aporte del informe realizado por el laboratorio de Ciencias de la vida de la Universidad Técnica Particular de Loja (Abad y Poluche, 2018)

Resultados

Los resultados obtenidos indican que las plantas medicinales, específicamente manzanilla, toronjil, romero y ortiga, cultivadas en la Univer-

sidad Politécnica Salesiana de la Zona Norte, fueron las que almacenaron la mayor cantidad de carbono, alcanzando un total de $0.005951 \text{ TnCO}_2/\text{m}^2$. De estas, el toronjil destacó como la especie con la mayor capacidad para capturar carbono, con un total de $0.0022075 \text{ TnCO}_2/\text{m}^2$. Al analizar los resultados por sectores, se observó que las plantas medicinales provenientes del mercado 10 de agosto, ubicado en la zona centro de Cuenca, lograron una notable retención de carbono, con un total de $0.008753 \text{ TnCO}_2/\text{m}^2$. La manzanilla también mostró un desempeño considerable, capturando $0.0030204 \text{ TnCO}_2/\text{m}^2$. En contraste, las especies sembradas en la zona de María Auxiliadora mostraron una mayor concentración de plomo en comparación con las otras ubicaciones. Además, los vegetales adquiridos en los mercados de Cuenca demostraron una mayor capacidad de captura de carbono en comparación con aquellos que fueron cultivados. En particular, las especies del mercado 10 de agosto con una retención de carbono de $0.008753 \text{ TnCO}_2/\text{m}^2$ destacaron, mientras que las plantas sembradas en la ciudadela 11 de octubre registraron la menor captura de carbono, con un total de $0.004663 \text{ TnCO}_2/\text{m}^2$. Por último, las plantas cultivadas en María Auxiliadora (centro histórico de Cuenca) destacaron por revelar la mayor concentración de plomo, alcanzando los $210,57 \text{ mg/kg}$. Abad y Poluche (2018).

Discusión científica

La etnobotánica juega un papel fundamental dentro de la mitigación del cambio climático (Zhiminaicela *et al.*, 2020), ya que la población podría adaptarse en medida que retome el conocimiento cultural de sus generaciones pasadas (Castellanos, 2011). Las plantas medicinales mayoritariamente se refugian en los bosques naturales, así los estudios etnobotánicos reflejan la realidad acerca del uso del bosque, el cual debe ser considerado para el control de intervenciones, ya que se pierde la biodiversidad y conocimiento generacional (Jijón, 2015). Lo anterior implica pérdida del saber ancestral, uso insostenible de los recursos naturales, cambios de los hábitos alimenticios, medicinales y de la cosmovisión tradicional (Aguinda, 2015). Se suma a esto que las comunidades aledañas

(a los bosques), además de dedicarse a la actividad agrícola, también se dedican a la crianza de animales mayores como el ganado bovino y ovino, las cuales son actividades que están avanzando hacia el bosque (Taípe, 2016), devienen en una lata tasa deforestación. La flora medicinal juega un papel fundamental en el diario vivir de las comunidades cercanas al bosque. Nieves y Solano, 2021).

Entre todas las plantas medicinales cultivadas en las tres zonas de Cuenca, se registró una captura total de $0,015389 \text{ TnCO}_2/\text{m}^2$. Por otro lado, las plantas compradas en los tres mercados alcanzaron una captura total de $0,023046 \text{ TnCO}_2/\text{m}^2$, lo cual es similar a lo informado por Galarza, P. (2017). En, el caso de las hortalizas, estas lograron capturar un total de $0,1028 \text{ TnCO}_2/\text{m}^2$. Abad y Poluche (2018).

Además, estudios realizados por Fernández (2016), Calle y Zhindon (2019) quienes capturaron 2,70 tn y 0,38192 tn respectivamente en hortalizas corroboran lo publicado por Cuéllar (*et al.*, 2015) que menciona que las plantas son uno de los principales sumideros de GEI planeta desempeñando un papel fundamental en la mitigación del cambio climático debido a su capacidad de fijar carbono atmosférico. Los límites permisibles establecidos en la Norma Técnica Ecuatoriana NTE INEN 2392:2007 de requisitos de hierbas aromáticas, nos indica que la máxima concentración de plomo es de 10 mg/kg (INEN, 2016).

Con respecto a las plantas adquiridas en los mercados los niveles más altos de plomo se encuentran en la zona sur-Mercado el Arenal (197.73 mg/kg) y las más baja en la zona centro-Mercado 10 de Agosto (56 mg/kg) superando los límites permisibles establecidos por la Norma Técnica Ecuatoriana NTE INEN 2392:2007 valores superiores a los encontrados por (Carrera, 2016) quien realizó un estudio en el que se analizó la concentración plomo en plantas de toronjil y taraxaco en treinta mercados del distrito metropolitano de Quito encontrado un promedio de 0,267 mg/kg de plomo, valores que están dentro del límite aceptado por la Norma Técnica Ecuatoriana NTE INEN 2392:2007.

Las concentraciones de Pb que se obtuvo en las plantas sembradas en las tres zonas de Cuenca presentan un valor promedio de 35,93 mg/kg siendo similares a las reportadas por (Królak, 2003) 37,3 mg/kg en el área metropolitana de Silesia en Polonia, la semejanza se debe a que se encuentran en un ambiente severamente contaminado por actividades industriales.

Conclusiones, análisis y relevancias

En el experimento llevado a cabo por Abad y Poluche (2018), la concentración de plomo medida en las plantas medicinales cultivadas fue de 144.01 mg/kg, mientras que en las compradas en mercados se registró 129.63 mg/kg, ambos valores exceden significativamente el límite establecido por la Norma Técnica Ecuatoriana NTE INEN 2392:2007 (10 mg/kg). En lo que respecta a la captura de CO₂, las especies sembradas acumularon un total de 0.015389 Tn/m², frente a las adquiridas en el mercado, que alcanzaron 0.022986 Tn/m². Entre las especies, la manzanilla cultivada fue la que más carbono absorbió, con 0.0048977 TnCO₂/m², mientras que el toronjil comprado destacó al capturar 0.0072557 TnCO₂/m². Esta habilidad para absorber CO₂ ayuda a mitigar los efectos del cambio climático, evitando la dispersión de gases en la atmósfera. El análisis comparativo reveló que las plantas cultivadas en las áreas de estudio mostraron un nivel total de plomo de 431.50 mg/kg, superando a las plantas compradas en los mercados, que presentaron 388.9 mg/kg. En cuanto a la captura de carbono, las plantas del mercado demostraron una capacidad mayor, totalizando 0.028986 TnCO₂/m², mientras que las sembradas alcanzaron 0.015389 TnCO₂/m².

Las plantas medicinales contribuyen a la lucha contra el cambio climático, pues los resultados habidos muestran que todas las especies (cultivadas y las adquiridas en los mercados) capturan plomo, un metal nocivo para la salud humana y animal, así como carbono, lo que inhibe el desplazamiento de gases a la atmósfera disminuyendo el efecto invernadero. Sin embargo, desde la bioética no deberían ser comercializados

sin las debidas precauciones. Este estudio marca un límite entre lo bueno que significan las especies para la captura del carbono y para diversificar estéticamente los parques y jardines de los ciudadanos, mas no significa esto que son amigables para el consumo humano. Los datos expresados en la experimentación llevada a cabo por Abad y Poluche (2018), expresan el peligro de envenenamiento crónico o afecciones diversas a la salud humana de los consumidores. Que sirva esta información para que las entidades que tienen poder de decisión y control tomen cartas en el asunto, y no permitan la comercialización de estos vegetales provenientes de los cultivos en la ciudad y los que están a la venta en los mercados.

Los gobiernos de turno han respondido a intereses particulares, de poderes supranacionales, cuya práctica del extractivismo de los recursos naturales, entre ellos las plantas medicinales, implica también el dominio de las masas populares a través de una soterrada democracia donde el voto es obligatorio, los impuestos son coercitivos y la justicia es inequitativa y antiética. Supuestamente el poder radica en el pueblo, pero quienes representan a éste en el poder legislativo, son marionetas alza manos y ahora aplasta botones, que le ponen precio a sus votos a favor o en contra de leyes que aprietan el yugo de los pobladores. Súmese a esta criminalización la actitud sordomuda de las religiones imperantes y de las otras confesiones emergentes, que llevan a su grey al prado florido de la esperanza en un más allá desconocido donde la leche y miel mana como el agua para quienes en lo terrenal hayan sufrido, mientras que los pastores o líderes religiosos gozan de la opulencia en la presente vida.

La confesión del pecado de palabra, obra y omisión no aplica para los pastores, solo para las ovejas cuyo balido retumba en las quebradas, en la verde montaña, en el campo florido, o en el atestado barrio popular suburbano donde los indígenas han hallado transitorio refugio porque de sus tierras por las mineras, petroleras y madereras desplazados han sido. También son alcahuetes de la criminalización algunos medios de comunicación, que priorizan la noticia hacia la sangre, los golpes y el vandalismo, dejando de lado el análisis del porqué de la lucha indígena por

la conservación de la naturaleza, el agua, la medicina ancestral y sus derechos propios de seres humanos, sus causas y consecuencias y las posibles soluciones en el marco del diálogo corresponsable, abierto y transparente.

Cosa curiosa es que quienes criminalizan desde el Estado son hijos e hijas de indígenas también, pues son policías y militares, que claro está ejecutan una orden sin opción de conciencia. Aunque históricamente el desarrollo y crecimiento de nuestro país se ha basado en el extractivismo antiético que es una de las causas del cambio climático, la implementación y beneficios de este sistema no han sido administrados de forma justa por parte del Estado para los diferentes sectores de la sociedad ecuatoriana. Los ojos de la educación primaria, secundaria y universitaria deben mirar hacia la bioética de la:

- Conexión con la madre naturaleza y el inmenso universo
- Importancia de la comunidad formada por personas y la reciprocidad como principio.
- Balance entre el aspecto espiritual y el material en el ser humano.
- Atención en la armonía y el respeto hacia todas las formas de vida, es decir, el biocentrismo.

Reflexión / invitación / convocatoria

“*La verdad os hará libres*” (Jn, 8, 31-32). La investigación de la naturaleza debe ser la búsqueda de la verdad y en libérrima libertad, uno de los pilares de la bioética para la construcción de una sociedad equitativa y fraterna, en medio de una Tierra cambiante, de un clima variante, de una adaptación al cambio climático. Se hace menester un diálogo fraterno entre la fe y la razón, para alcanzar mejores tecnologías limpias aplicables por la ciencia experimental para el cuidado de la casa común que es el planeta Tierra.

Finalmente, es menester sumarse a una oración que algún agricultor algún día lo hiciera en la cuenca del río Paute, un río caudaloso de la zona sur del Ecuador, cuando miraba el campo lleno de plantas verdes y

floridas: La casa de todos, el bien común explotado por pocos en franca negación a la participación de las mayorías. Recursos naturales que en seis días fueron creados por el Gran Arquitecto o Hacedor del Universo, y que bastó un solo día para que el creado ser humano hiciera para sí el derecho de uso hasta la extinción de los otros usuarios. Cuanto egoísmo centrado en la acumulación y el desperdicio en desmedro del prójimo donde el hambre y la miseria por doquier campea. Las sombras de la inequidad avanzan con la soledad sonora en el campo y por la selva de cemento citadinas, mientras los pocos que acumulan el capital cierran sus ojos ante la miseria humana, las pandemias y la desnutrición infantil (consecuencias del cambio climático y la actuación anti bioética del ser humano). Ya no hay gente bañándose en las fuentes, los ríos y lagos que bajan de las montañas y que atraviesan las ciudades, porque sus aguas son turbias y pestilentes, sin vida acuática ni flora circundante. Ya no hay aves en los vacíos parques citadinos, porque de árboles y matas carecen o simplemente porque sus frutos saben ácidos y sin nutrientes.

Oh, Dios de Jesús, cuanto daño se ha hecho a tu obra maestra, contaminando el agua, la tierra y el aire, cortando y quemando los bosques, extinguiendo a las especies menores y sobrepoblando con pocas especies mayores, solo para alcanzar el rédito económico, el poder político y hasta el religioso. Oh, Cristo del principio, la higuera que no te brindó frutos hoy está seca y vacía del todo, mientras las espigas de trigo que arrancaron tus discípulos en aquel sábado judaico hoy están vacías porque sus semillas infértiles al salir el sol se han quemado. Mira la quietud del sol al mediodía y luego la velocidad del viento en el atardecer del desierto. Mira la tempestad de la lluvia que destruye cuanto encuentra a su paso. Mira al ser humano que marca con sus manos, cicatrices sobre los inocentes rostros vivientes en la faz de la tierra. Mira tú obra Padre, donde la desolación y la angustia hacen brotar lágrimas de los cansados ojos de los sobrevivientes. Mira con amor paterno y haz el milagro de un nacimiento nuevo en la conciencia del ser humano, de los gobernantes y gobernados, para que se unan como hermanos en el cuidado y la gestión de un ambiente sano en medio de este cambio climático, donde las plantas

medicinales han sido contaminadas por la acción no ética por parte del ser con amor creado, el denominado ser humano. *Laudato si, mio signore* (Alabado seas mi Señor).

Referencias bibliográficas

- Abe, R y Ohtani, K. (2013). *Traditional botanical knowledge of medicinal plants in a quilombola community in the Atlantic Forest of northeastern Brazil*. <http://bit.ly/3WjwV56>
- Abad, D. y Poluche, B. (2019). *Comparación de los niveles de plomo y carbono presentes en las plantas medicinales adquiridas en tres mercados y en cultivos de tres zonas urbanas de Cuenca (Sur, Centro histórico y Norte*. Cuenca, Ecuador. Tesis de pregrado UPS.
- Acosta, J. y Salcedo, R. (2020). La medicina tradicional como práctica profesional reconocida por la OMS. *Revista Internacional de Salud Pública*, 15(3), 45-58.
- Aguirre, A. (2014). *Colección y caracterización morfológica de la diversidad de plantas medicinales de las provincias de Loja y Cotopaxi*. Universidad Central del Ecuador, Facultad de Ciencias Agrícolas.
- Akerele, O. (1993). Las plantas medicinales: un tesoro que no debemos desperdiciar. *Foro Mundial de la Salud*, 14(4), 390-395. <http://bit.ly/3WfwSHx>
- Almeida, R. (2000). *Uso de plantas medicinales*. Editorial Árbol S.A. de C.V.
- Bermúdez, A., Oliveira, M. A. y Velázquez, D. (2005). La investigación etnobotánica sobre plantas medicinales: Una revisión de sus objetivos y enfoques actuales. *INCI*, 30(8).
- Calle, D. y Zhindon, J. (2019). *Evaluación de la presencia de plomo a través de huertos hortícolas agricultura urbana en cinco colegios ubicados en el distrito Sur de Cuenca*. Tesis de pregrado UPS.
- Cardona, J. y Rivera, A. (2012). *Plantas medicinales en Colombia para el tratamiento de la diabetes y la anemia: Un enfoque etnobotánico*. Universidad Nacional de Colombia.
- Carrera, P. (2016). *Determinación microbiológica y de metales pesados en toronjil (Melissa officinalis) y taraxaco (Taraxacum officinale) expendidos en los diferentes mercados del Distrito Metropolitano de Quito*. Universidad Central del Ecuador.
- Cuéllar, A., Ramírez, J. y López, M. (2015). Las plantas como sumideros de gases de efecto invernadero y su papel en la mitigación del cambio climático. *Revista de Ecología y Medio Ambiente*, 23(2), 56-73.

- De la Torre, L., Navarrete, H., Muriel, P., Macía, M. J. y Balslev, H. (2008). *Enciclopedia de las plantas útiles del Ecuador*. Herbario QCA, Pontificia Universidad Católica del Ecuador. <https://bit.ly/3VF242w>
- Fernández, J. (2016). *Captura de carbono en hortalizas: Un análisis de su impacto ambiental en Ecuador*. Universidad Politécnica Salesiana.
- García, P. (2004). Los riesgos de toxicidad en la medicina tradicional: Una revisión de casos clínicos en América Latina. *Revista Latinoamericana de Salud Pública*, 12(1), 34-50.
- Gallegos-Zurita, M. y Gallegos-Z, D. (2017). Plantas medicinales utilizadas en el tratamiento de enfermedades de la piel en comunidades rurales de la provincia de Los Ríos – Ecuador. *Anales de la Facultad de Medicina*, 78(3), 315.
- Galarza, P. (2017). *Determinación de la concentración de elementos contaminantes y bacterias patógenas en un sistema de agricultura urbana compuesto por hortalizas, implementado en la Universidad Politécnica Salesiana*. Tesis de pregrado UPS.
- Gómez-Baggethun, E., Corbera, E. y Reyes-García, V. (2010). Traditional ecological knowledge and global environmental change: Research findings and policy implications. *Ecology and Society*, 15(4), 28.
- INEN (2016). *Norma Técnica Ecuatoriana NTE INEN 2392:2007. Requisitos de hierbas aromáticas*. Instituto Ecuatoriano de Normalización.
- Jijón, A. (2015). *Conocimiento tradicional etnobotánico del área de influencia del Bosque Protector Aguarongo, Azuay, Ecuador*. Investigación UDA.
- Królak, E. (2003). *The use of plants as bioindicators of heavy metal contamination in urban areas*. *Environmental Pollution*, 122(3), 37-45.
- Lara, R. y Sánchez, M. (2018). Eficiencia de extractos de plantas y metabolitos secundarios como alternativa a la resistencia a antibióticos. *Revista Internacional de Biotecnología y Salud*, 9(1), 58-74.
- Maregesi, S. y Mtewa, A. (2007). Ethnopharmacological survey of medicinal plants used in Tanzania for treating infectious diseases. *Journal of Ethnopharmacology*, 112(2), 234-243.
- Martín, M. (2017). *Análisis de contaminantes metálicos y minerales en plantas medicinales: Evaluación del riesgo para la salud humana*. Universidad de Granada.
- Nieves, D. y Solano, H., (2021). *Evaluación ambiental y etnobotánica de la flora medicinal en el Bosque Protector Aguarongo y las parroquias aledañas de Jadán, Zhidmad y Santa Ana*. Tesis de pregrado UPS.

- Oliveira, C., Lima, S., y Santos, M. (2012). Plantas medicinales en la farmacología moderna: Uso y comercialización de fitomedicamentos. *Revista Brasileña de Ciencias Farmacéuticas*, 18(2), 145-162.
- Onofre, J. (2013). *La medicina tradicional y sus nuevas perspectivas en la salud pública*. Editorial Universitaria Andina.
- Osinaga, B. C. (2014). *Monitoreo de diversidad vegetal y carbono en bosques andinos-protocolo extendido*. Editor CONDESAN.
- Pabón, G., Mendoza, R. y Arango, L. (2017). Transmisión del conocimiento etnobotánico en comunidades indígenas de la Amazonía ecuatoriana. *Revista de Estudios Amazónicos*, 14(1), 45-62.
- Pérez, G. (2008). *Valoración etnobotánica de especies medicinales en la región Andina: Un enfoque comparativo*. Universidad Nacional Autónoma de México.
- Zhiminaicela, J., Fernández, R. y Calle, P. (2020). Etnobotánica y cambio climático: Adaptación de comunidades rurales en Ecuador. *Revista Científica de Ciencias Ambientales*, 7(2), 89-104.

La eutanasia en Ecuador: análisis de la sentencia 67-23-In/24 de la Corte Constitucional

Jorge Iván Rengel Maldonado
Universidad Politécnica Salesiana
jrengel@ups.edu.ec
<https://orcid.org/0000-0002-3910-7345>

Introducción

Este artículo pretende analizar la eutanasia en el Ecuador, tomando como punto de partida la Sentencia 67-23-IN/24 de la Corte Constitucional. La eutanasia es un tema de constante discusión y controversia a nivel mundial y ha tomado realce en Ecuador a raíz del fallo de la Corte Constitucional, dictado como consecuencia del pedido formulado por Paola Roldán, quien padecía de una enfermedad neurológica progresiva conocida como esclerosis lateral amiotrófica (ALS en inglés, ELA en castellano).

La metodología hace un análisis documental, tanto de la referida sentencia de la Corte Constitucional, así como de los puntos de vista de diferentes autores que se han manifestado a favor o en contra de la eutanasia y del marco jurídico vigente en el país.

Entre los argumentos favorables a la eutanasia, encontramos temas como la autonomía del paciente, que sugiere que las personas tienen derecho a tomar sus propias decisiones, incluidas su vida o muerte; especialmente, cuando están padeciendo un gran sufrimiento debido a

enfermedades o lesiones incurables. Se puede considerar, además, el derecho a una muerte digna, que permitiría a las personas finalizar su vida evitando sufrimientos prolongados e innecesarios. Por último, la empatía con por la persona que sufre o padece dolorosa e inhumanamente, como consecuencia de una enfermedad o lesión incurable.

Por otro lado, entre los argumentos en contra, se puede destacar el valor de la vida humana, como algo sagrado, especialmente, desde la perspectiva religiosa o moral. El posible abuso que se pudiera hacer de esta figura, malas interpretaciones o presiones sobre las personas por diferentes razones; a todo esto, se suma la idea de los cuidados paliativos, como una alternativa a la eutanasia, con la finalidad de aliviar el sufrimiento.

En medio de estas discusiones, la sentencia 67-23-IN/24 se erige como una resolución clave de la Corte Constitucional del Ecuador. Esta sentencia aborda la acción pública de inconstitucionalidad, acompañada de una solicitud de suspensión del Artículo 144 del Código Orgánico Integral Penal, COIP, promulgado por la Asamblea Nacional; demandada por Paola Roldán. El Art. 144 del COIP tipifica el homicidio simple; textualmente, señala: “La persona que mate a otra será sancionada con pena privativa de libertad de diez a trece años” (COIP, 2025). Mediante la mencionada sentencia, la Corte Constitucional resuelve sobre seis aspectos fundamentales (Corte-Constitucional, 2024):

- a. Declara la constitucionalidad condicionada, de la conducta tipificada en el Art. 144 del COIP, cuando se cumplan los supuestos expuestos en la mencionada decisión;
- b. Declara la constitucionalidad aditiva del Art. 6 del Código de Ética Médica;
- c. Declara la inconstitucionalidad del Art. 90 del Código de Ética Médica, que queda fuera del ordenamiento jurídico;
- d. Dispone que el Defensor del Pueblo, en el plazo de 6 meses desde la notificación de la sentencia, prepare un proyecto de ley que regule los procedimientos eutanásicos;

- e. Dispone que el Ministerio de Salud, en el plazo máximo de 2 meses desde la notificación de la sentencia, prepare un reglamento para la aplicación de la eutanasia activa voluntaria; y,
- f. Dispone que la Asamblea Nacional, en plazo máximo de 12 meses, desde la presentación del proyecto de ley, referido en el numeral 4, precedente, “conozca, discuta y expida” la ley que regule los procedimientos eutanásicos.

En definitiva, la sentencia 67-23-IN-24 de la Corte Constitucional, legitima la eutanasia en Ecuador, condicionada a circunstancias específicas, determinadas en su texto. Este artículo analizará, en lo esencial, los antecedentes, dentro del marco de la acción planteada por la señora Paola Roldán Espinosa, que dan origen a la sentencia. Se revisará la reflexión que hace la Corte Constitucional en sus aspectos más relevantes y, sucintamente, la decisión adoptada por la mayoría de los jueces, que abre en Ecuador la posibilidad de acceder a la eutanasia a las personas quienes cumplan con las especificidades contenidas de la resolución, lo que abre un nuevo debate.

Antecedentes del Caso 67-23in

La accionante, señora Paola Roldán Espinosa, con fecha 08 de agosto de 2023, propone ante la Corte Constitucional del Ecuador, la acción pública de inconstitucionalidad, junto a una solicitud de suspensión del Art. 144 del COIP, emitido por la Asamblea Nacional. La causa es conocida por el Juez Constitucional Enrique Herrería Bonnet. Se admite la acción con fecha 29 de septiembre de 2023 por parte del Tercer Tribunal de la Sala de Admisión de la Corte Constitucional y se niega la solicitud de suspensión provisional del Art. 144 del COIP. De la misma manera, se dispuso que la Asamblea Nacional, el Presidente de la República y el Procurador General del Estado, intervengan defendiendo o impugnando la Constitucionalidad del Art. 144 del COIP (Corte-Constitucional, 2024).

Adicionalmente, se resolvió que el Pleno de la Corte conozca del caso para otorgarle una priorización; esto es, que se atienda de manera

más expedita el pedido contenido como parte de la acción presentada y, así, el 9 de septiembre de 2023, el Pleno de la Corte Constitucional acepta la priorización. Con fecha 20 de noviembre de 2023 se convoca a audiencia a las partes procesales: Sra. Paola Roldán Espinosa y sus abogados: Farith Simón Campaña, Ramiro Ávila Santamaría y Pablo Encalada; Yolanda Salgado Guerrón, abogada de la Presidencia de la República; Edgar Fabián Lagla Toapanta, asesor de la Coordinación General de Asesoría Jurídica de la Asamblea Nacional y a los amici curiae seleccionados (Corte-Constitucional, 2024). Sobre este último punto cabe anotar que hubo muchas personas que se presentaron con su amicus curiae; entre ellos, la Corte escogió a quienes pudieron participar en la audiencia.

Argumentos de los sujetos procesales y planteamiento de los problemas jurídicos.

En esta parte el enfoque se centrará en el análisis crítico-jurídico de los principales argumentos utilizados por la accionante, señora Paola Roldán; así como el planteamiento de los problemas jurídicos fundamentales, previos a la resolución emitida por la Corte Constitucional.

- Principales fundamentos de la acción y pretensión y su relación con los problemas jurídicos fundamentales planeados en la sentencia.

La acción de inconstitucionalidad que desemboca en la sentencia 67-23-IN/24, menciona que la norma impugnada (Art. 144, COIP), ‘infringe’ los siguientes derechos:

- A la dignidad;
- Al desarrollo de la personalidad;
- Al fomento de la autonomía y disminución de la dependencia;
- A la integridad física y prohibición de tratos crueles, inhumanos y degradantes;
- Derecho a morir dignamente. (Corte-Constitucional, 2024)

Con respecto a la dignidad, la accionante, Paola Roldán, manifiesta:

10.1 [La dignidad es trasgredida] si es que no prevalece el fin que las personas puedan decidir sobre sus vidas, en uso de su autonomía y libertad y se imponen fines ajenos que provienen del Estado, la ética, la religión y los valores que una persona no comparte. (Sentencia 67-23-IN/24, 05/02/2024, página 6)

Al respecto, se considera importante mencionar que, todo derecho, cualquiera que sea, tiene un límite: mi derecho termina donde comienza el derecho de otro; esta idea determina los límites de la libertad individual, idea que se atribuye a Jean Paul Sartre:

Los límites de la libertad individual son un tema central en el ámbito legal. Si bien gozamos de una serie de derechos y libertades reconocidos, es importante tener en cuenta que estos derechos no son absolutos y deben ser ejercidos dentro de los límites establecidos por la ley. El principio de convivencia en el marco legal implica que nuestras acciones deben estar en línea con el ordenamiento jurídico y respetar los derechos y libertades de los demás. Esto significa que nuestra libertad individual termina donde comienza la libertad y los derechos de los demás. (Todoinfolegal.com, n.d.)

Lo citado implica que, en definitiva, se puede ejercer cualquier derecho, mientras éste no vulnere el ejercicio del derecho de los demás: *mi libertad termina donde empieza la del otro*. En este contexto, el Estado como ente creador de las normas jurídicas, ha generado límites que todos debemos respetar. Por otra parte, se debe entender que el argumento sobre la *dignidad transgredida*, expuesto por la accionante, no es un absoluto que se pueda esgrimir frente a los demás; pues la dignidad de los otros, sus creencias, sus convicciones y las normas jurídicas, son para todos. No se puede, entonces, superponer los anhelos, los deseos, dependiendo de la situación de una persona, frente a los derechos, deseos y voluntad de los demás. Son las normas jurídicas las que ponen un orden en la sociedad y es el Estado el llamado a hacer cumplir el orden que deviene de las normas jurídicas.

No obstante, hay otro argumento recurrido por la accionante que se considera importante destacar y es el siguiente (Corte-Constitucional, 2024):

La aplicación del tipo penal de homicidio simple a quienes asisten a una persona que padece intensos sufrimientos físicos o emocionales por una enfermedad o lesiones graves para que tengan una muerte por piedad, (sic) atenta contra la dignidad de las personas enfermas [e] impide ejercer su derecho a decidir hasta cuándo vivir [...] castigando a quienes contribuyen desde un conocimiento especializado a cumplir la voluntad del sujeto pasivo y negándoles así la posibilidad de contar con atención médica profesional. (Sentencia 67-23-IN/24, 05/02/2024, página 6)

En este caso el argumento es quitar la vida a alguien por ‘piedad’; palabra que, en el contexto de lo expuesto por la accionante implica que alguien se compadece del ‘intenso sufrimiento físico o emocional causado por una enfermedad o lesión grave’; pero, sin dudarlo, implica un condicionamiento, que está sujeto a un análisis integral de la situación física y emocional de la persona cuyo padecimiento sería ‘aliviado’, a través de la actuación de un profesional de la salud, quien con pleno conocimiento de causa actúa por compasión o piedad. Se encuentra entonces, siempre presente, el contexto de que no se trata de cualquier situación sino una muy peculiar *intenso sufrimiento físico o emocional causado por una enfermedad o lesión grave*, que va a ser ponderado por un profesional de la salud.

Al referirse a la autonomía y disminución de la dependencia, la señora Roldán, plantea los siguientes argumentos (Corte-Constitucional, 2024):

La Constitución reconoce que la autonomía y la independencia son fundamentales con relación a las personas con discapacidad; sin embargo, puede ocurrir que, en ciertas circunstancias de una enfermedad o lesión grave, tenga como consecuencia inevitable la pérdida de autonomía y el aumento creciente de la dependencia. La autonomía no se restringe al aspecto físico, pues implica tomar decisiones sobre cómo vivir y hasta cuándo vivir de acuerdo con lo que la persona considera que es bueno, deseable y posible. La persona autónoma es dueña de su ser, tiene soberanía sobre su vida y su cuerpo, y tiene autoridad propia. (Sentencia 67-23-IN/24, 05/02/2024, página 7)

Es indudable que la autonomía de una persona se puede ver limitada, de una manera abrumadora, por determinadas enfermedades o

lesiones graves. Esto es absoluto y real; pero, afirmar que la autonomía implica: tomar decisiones sobre cómo vivir y hasta cuándo vivir de acuerdo con lo que la persona considera que es bueno, deseable y posible, implica caer nuevamente en una generalización que no es real. En este caso, la autonomía estaría sujeta a las condiciones de que se ha visto limitada por enfermedades o lesiones graves; no un *tedium vitae* o algún tipo de depresión, que ameritarían otro tipo de tratamiento. No se puede caer, entonces, en una generalización.

Es interesante revisar el siguiente argumento sobre la dependencia total, traído por la accionante (Corte-Constitucional, 2024):

La dependencia total, permanente, para toda necesidad biológica, humana y emocional [...] puede considerarse como contradictoria con la ‘disminución de la dependencia’ que se convierte en un objetivo imposible y que pone a las personas en una posición que en muchos casos es considerada por ellas como una vida humillante e indeseable. (Sentencia 67-23-IN/24, 05/02/2024, página 7)

La dependencia total crea una situación de vulnerabilidad emocional, a tal punto que quien la vive busca, humanamente, terminar con ese padecimiento. Más adelante, la señora Roldán, al argüir acerca del *dolor intenso por una enfermedad* manifiesta lo siguiente (Corte-Constitucional, 2024):

Sin detrimento de esto, señala que existe la obligación de toda autoridad pública de impedir o hacer cesar amenazas o vulneraciones a la integridad personal. A su criterio, “el padecer un dolor intenso por una enfermedad, que la persona considere insoportable, se convertiría en una vulneración a la integridad en el ámbito privado, si se impide la posibilidad de poner fin a esos sufrimientos mediante una muerte digna”. (Sentencia 67-23-IN/24, 05/02/2024, p. 9)

Efectivamente, frente al padecimiento de una enfermedad o lesiones graves se genera, en muchos casos, una dependencia total y permanente para suplir todas o casi todas las necesidades que implica tener una naturaleza humana, ligada a su biología o emociones y esto, indudablemente, puede llegar a ser algo indeseable o humillante, que determina que la

vida digna, como derecho, ha concluido para alguien. El derecho a una vida digna forma parte de los denominados derechos humanos, que son reconocidos en general, tanto en normas jurídicas al interior de los países, como en instrumentos internacionales como la Declaración Universal de Derechos Humanos de 1948, la cual, en el numeral 1, de su artículo 25, en la parte pertinente, señala: “Toda persona tiene derecho a un nivel de vida adecuado que le asegure, así como a su familia, la salud y el bienestar, y en especial la alimentación, el vestido, la vivienda, la asistencia médica y los servicios sociales necesarios [...]” (ONU, 1948).

De la misma forma, dentro de la denominada *Declaración de los Derechos Humanos Emergentes*, del Instituto de Derechos Humanos de Cataluña, en el acápite llamado: I. Marco General: valores y principios, sub-acápite II, valores y, con respecto a la vida digna, se acota lo siguiente (Institut-de-Drets-Humans-de-Catalunya, 2009):

Por otra parte, *si bien el individuo quiere vivir por encima de todo, quiere al mismo tiempo vivir bien, quiere una vida de calidad*. El desarrollo científico y técnico, unido al desarrollo económico, puede estar al servicio de la vida humana e incluso no humana, pero también puede desarrollarse en detrimento de dicho valor. La defensa de un medio ambiente sano y equilibrado, la reclamación del derecho a una muerte digna, el valor que damos a la ciudad, ponen de manifiesto que la vida que valoramos hoy es una vida de calidad (énfasis del autor).

De igual manera, sobre la autonomía, la accionante, Paola Roldán, hace la siguiente reflexión argumentativa, relacionándola con la idea de la *muerte digna* (Corte-Constitucional, 2024):

La autonomía se manifiesta en el consentimiento libre y voluntario, que es uno de los requisitos indispensables para ejercer el derecho a la muerte digna [...] por ello podría considerarse que, en el Ecuador, el tipo penal del homicidio simple interfiere de manera desproporcionada en la autonomía y la autodeterminación de cada persona para escoger un plan de vida y en los eventos asociados a la muerte digna, también de elegir el momento y modo en que desean terminar su existencia, en el marco de respeto a su dignidad. (Sentencia 67-23-IN/24, 05/02/2024, página 7)

Es cuestionable la proposición expuesta al final del párrafo citado pues, la idea de la muerte digna siempre va a estar condicionada, pues no se trata de una decisión (*escoger el momento y el modo*: parece más un suicidio) y nada más. Está condicionada directamente a la idea de enfermedad o lesiones graves, que produzcan intensos sufrimientos; solamente, cuando se cumplan estas condiciones se podría considerar primero la idea de una muerte digna y luego, la idea de que la acción del facultativo de salud, no está contraviniendo el Art. 144 del COIP. No obstante, más adelante, la señora Roldán, manifiesta también con respecto a la muerte digna y, haciendo alusión a una sentencia previa de la Corte Constitucional, lo siguiente (Corte-Constitucional, 2024):

En cuanto a la muerte digna, refiere que es un “derecho de quienes padecen y han sufrido enfermedades graves” y señala que la Corte Constitucional lo reconoció en la sentencia 679-18-JP/20 y acumulados al precisar que “el derecho al disfrute pleno de la salud implica la mejora de las capacidades y potencialidades para que la vida de la persona con enfermedad sea lo más plena posible [...]” y que estas capacidades y “potencialidades para la vida” también implican “la consideración de una *muerte natural digna*, sin dolor ni padecimiento” (énfasis en el original). En tal sentido, señala que el paciente tiene derecho a decidir y definir su comprensión del nivel más alto de salud posible en el curso de su enfermedad hasta su muerte, por lo que, puede “optar por detener y cambiar el tratamiento con medicamentos”. (Sentencia 67-23-IN/24, 05/02/2024, página 9)

Sobre éste, dentro del “Planteamiento de los Problemas Jurídicos”, números 31 y 32, parte pertinente, la Corte Constitucional, realiza la siguiente aclaración que es interesante con relación al *reconocimiento al derecho a la muerte digna*, al que se menciona la accionante (Corte-Constitucional, 2024):

31. En el caso 679-18-JP/20 y acumulados se abordó el derecho a la vida digna y a la salud, contenidos principalmente en el artículo 66 numeral 2 de la CRE, con ocasión del acceso a medicamentos y la Corte no reconoció el derecho a la muerte digna. De esta manera, se planteó que el “más alto nivel posible de salud contribuye a poner las condiciones dignas de vida y dignas de muerte por sobre la vida en circunstancias de padecimiento y miseria”.

(CCE, sentencia 679-18-JP/20 y acumulados, 5 agosto de 2020, párr. 89) Lo que esclareció este Organismo es *que una salud óptima contribuye a que las condiciones de vida y de muerte no estén rodeadas de sufrimiento y de dolor*, por lo tanto, el acceso a medicamentos que ayuden a combatir dichas afecciones resulta trascendental. Incluso, se señaló que los pacientes pueden “optar por detener y cambiar el tratamiento con medicamentos”, lo cual se relaciona con la capacidad de cada persona para acceder a otros tratamientos, medicamentos o probar otro tipo de cuidados sin que aquello guarde relación con acceder a un procedimiento eutanásico para morir en caso de padecer una enfermedad grave e incurable o una lesión corporal grave e irreversible que provoque intenso sufrimiento (énfasis en el original).

32. En consecuencia, se desestima el cargo relativo a la inconstitucionalidad de la norma por ser contraria al derecho a la muerte digna, pues, como se evidenció ut supra, dicho derecho no se ha reconocido a través de la jurisprudencia de esta Corte, como esgrime la accionante [...] (Sentencia 67-23-IN/24, 05/02/2024, página 9)

Categoricamente, la Corte *niega el reconocimiento de la muerte digna como derecho* a través de la sentencia 679-18-JP/20 y acumulados de 05/08/2020. Sobre la idea de ‘muerte digna’ y el alargamiento de la vida de forma innecesaria, es importante, también acotar y citar que, en el documento “Declaración de los Derechos Humanos Emergentes”, del Instituto de Derechos Humanos de Cataluña, antes referido, dentro de su Título I, “Derecho a la Democracia Igualitaria, Artículo 1. Derecho a la existencia en condiciones de dignidad”, se refiere a que (Institut-de-Drets-Humans-de-Catalunya, 2009): “Todos los seres humanos y las comunidades tienen derecho a vivir en condiciones de dignidad”; y, luego, desarrollando esta gama de derechos, señala, en su numeral 7, puntualmente, con respecto a la “muerte digna”, lo siguiente: “(...) 7. *El derecho a una muerte digna, que asegura a toda persona el derecho a que se respete su voluntad de no prolongar artificialmente su vida, expresada en un testamento vital o documento similar formalizado con las debidas garantías*” (énfasis del autor).

No prolongar la vida artificial e innecesariamente, se complementaría, cuando se cumplan las condiciones de enfermedad o lesión grave

que genere el padecimiento de intenso sufrimiento en la persona. Es decir, cuando no exista realmente un estado de vida digna, conforme lo indicado en líneas precedentes. Es muy importante destacar, además, que, *en cuanto al libre desarrollo de la personalidad* y sobre la base del Art. 66 de la Constitución de la República del Ecuador, la accionante, hace la siguiente reflexión, que se conecta, con el párrafo anterior (Corte-Constitucional, 2024):

Quando una persona padece intenso sufrimiento físico o emocional por una enfermedad grave, en el ejercicio de su libre desarrollo de la personalidad, debería poder decidir libremente ponerles fin y escoger los medios para hacerlo, siempre que respete el derecho de terceras personas. *El fin es morir dignamente y el medio es el procedimiento de muerte asistida pues 'la lucha por la vida tiene que hacerse con el menor dolor posible, con la búsqueda de la paz con la mejora de la calidad de vida durante la enfermedad y hasta la muerte y respetando la voluntad del paciente'* (énfasis del autor). (Sentencia 67-23-IN/24, 05/02/2024, página 7)

Morir dignamente, es una aspiración humana universal y, más aún, es válida y consecuente con las limitaciones propias de cualquier enfermedad degenerativa como el ELA. Dentro del “Planteamiento de los Problemas Jurídicos”, desarrollado a través de la sentencia, la Corte Constitucional realiza la siguiente acotación (Corte-Constitucional, 2024):

El libre desarrollo de la personalidad para el ejercicio del derecho a la muerte digna podría verse afectado por la injerencia del ejercicio tradicional de la medicina, por los artículos 6 y 90 del Código de Ética Médica, por creencias religiosas, por el uso innecesario del derecho penal y por la presunta afectación al ejercicio de derechos de terceras personas. (Sentencia 67-23-IN/24, 05/02/2024, página 13)

Esto implica que existen factores externos que generan alguna injerencia o que pueden influir o ser determinantes al momento de querer hacer efectiva (aplicar en la vida real), la idea de la muerte digna, planteada por la accionante; factores como: creencias religiosas, el propio Código de Ética Médica; las normas jurídicas o derechos de terceros. Por otra parte, y de manera complementaria, es importante destacar el argumento que

expone la señora Roldán sobre el derecho a la integridad personal y la prohibición de tratos crueles (Corte-Constitucional, 2024):

No permitir morir dignamente en condiciones de padecimientos físicos y emocionales por una enfermedad grave optando autónoma y libremente para que un tercero capacitado le ayude a morir en condiciones de dignidad en aplicación del tipo penal de homicidio simple, es un atentado al derecho a la integridad personal por permitir condiciones de vida con dolores crueles, inhumanos y en condiciones degradantes. (Sentencia 67-23-IN/24, 05/02/2024, página 8)

Afirmación con la que se concuerda; pues, no compadecerse de un sufrimiento ajeno que va más allá de cualquier límite de tiempo o tolerancia física es simplemente inhumano y transgrede la idea de una vida digna y autonomía real del individuo. No obstante, la Corte Constitucional desestima este argumento (Corte-Constitucional, 2024): “por no reunir los requisitos señalados en la LOGJCC (Ley Orgánica de Garantías Jurisdiccionales y Control Constitucional): certeza, claridad, pertinencia y especificidad.” (Sentencia 67-23-IN/24, 05/02/2024, página 15)

Más adelante, la accionante hace alusión a que la Constitución de la República “reconoce los derechos a la inviolabilidad a la vida y a la vida digna” (Corte-Constitucional, 2024):

Provocar la muerte no es punible en determinadas circunstancias (sic) la ‘inviolabilidad’ de la vida tiene excepciones, por ejemplo, cuando se excluye la antijuridicidad en casos de legítima defensa, por estado de necesidad cuando se produce la muerte durante las hostilidades en conflicto armado o cuando el Estado permite la muerte por piedad como ha sucedido en muchos países. (Sentencia 67-23-IN/24, 05/02/2024, página 9)

Se refiere en este párrafo a las excepciones: casos en que la muerte de una persona no es punible y busca encasillar estas causas a la eutanasia; generando un símil entre éstas y aquella; considerándola, también un caso de excepción; al que habría que añadir: “siempre y cuando se cumplan los condicionamientos” referidos con anterioridad. Complementa

el argumento expuesto por la accionante la siguiente idea, que, también, es importante analizar debidamente en su alcance:

La dimensión de la vida digna exige obligaciones positivas de hacer para que las personas puedan tener buen vivir, calidad de vida y el máximo bienestar físico y emocional posible y cuando no es posible garantizar estas condiciones y si una persona se encuentra en circunstancias excepcionales como padecer intensos sufrimientos físicos o emocionales por una enfermedad o lesión física grave, el Estado debe ofrecer las condiciones para una muerte digna. Por tanto, el COIP no debe penar aquello que la Constitución reconoce como el ejercicio de un derecho que se deriva de la dignidad. (Sentencia 67-23-IN/24, 05/02/2024, página 9)

Para el autor, es una reflexión válida que concuerda con las ideas que se han sustentado e intercalado a las citas de los argumentos presentados por la señora Roldán; pues respalda la idea de generar una excepción que devino, luego, en la parte resolutive de la sentencia, materializada como ‘constitucionalidad condicionada’; esto es: ligada al cumplimiento de las ‘condiciones’ previas, referidas varias veces con anterioridad.

Para finalizar el análisis de los principales argumentos de la accionante, señora Paola Roldán; es fundamental revisar lo que expone sobre el *test de proporcionalidad*, que, como indica, debe (Corte-Constitucional, 2024): “verificar si la medida objeto de examen de constitucionalidad tiene (i) un fin constitucionalmente válido, (ii) es idónea, (iii) necesaria y (iv) proporcional” (Sentencia 67-23-IN/24, 05/02/2024, página 10). Al correr este proceso de *validación de proporcionalidad*, concretamente sobre los puntos “iii” y “iv” y generar el análisis correspondiente la accionante señala (Corte-Constitucional, 2024):

La necesidad. – La hipótesis del tipo penal ‘homicidio simple’ es matar a otra persona. La muerte a una persona, cuando la conducta es típica, antijurídica y culpable, es una medida necesaria para garantizar la inviolabilidad de la vida y el tipo penal [resulta] necesario. *Los cuidados paliativos y la suspensión o alteración de tratamientos médicos no interrumpen los dolores y no protegen la vida, sino que impiden la muerte digna.* En este sentido, el impedir la muerte digna podría considerarse una medida innecesaria

y por tanto inconstitucional (énfasis del autor) (Sentencia 67-23-IN/24, 05/02/2024, página 10).

Al respecto, se manifiesta el desacuerdo con la afirmación que se ha destacado y que se ha citado textualmente; pues no se debe generalizar a todos los casos; sino, que se debe primero, analizar su pertinencia en caso de enfermedad o lesión grave. Como se manifestó, caso contrario, tal generalización implicaría que, para cualquier situación, se deberían simplemente ‘descartar los cuidados paliativos’, argumentando que ‘no interrumpen dolores’, ‘no protegen la vida’ o que ‘impiden una muerte digna’. Lo que sí es correcto, es generar una excepción condicionada a la regla general contemplada en el Art. 144 del COIP, como en efecto, lo hizo la Corte Constitucional en la parte resolutive de la sentencia; cuyo análisis contextual se profundizará más adelante.

Para concluir esta parte del análisis, la accionante, con respecto a la *proporcionalidad*, manifiesta (Corte-Constitucional, 2024):

La proporcionalidad propiamente dicha. – Los derechos en juego son la vida digna del sujeto que quiere la muerte digna y la libertad de la persona que asiste a esa persona. La gran diferencia entre la muerte provocada por piedad con el homicidio simple es que el titular del bien jurídica vida pide y clama la muerte, esa persona no puede considerarse víctima sino un sujeto de derechos. En cambio, en el homicidio simple el titular del derecho a la vida muere en contra de su voluntad y es una víctima. (Sentencia 67-23-IN/24, 05/02/2024, página 10)

Por otra parte, se considera como válido y muy relevante este análisis, pues hace una clara diferenciación entre un hecho *típico, antijurídico y culpable* (delito), considerando la regla de que alguien causa la muerte de una persona y es considerado un delito; opuesta la *excepción* que sería sentir piedad o compasión por alguien que, cumplidas las condiciones de padecer una enfermedad o lesión grave y, voluntariamente, decide terminar con su vida, mediante la intervención de un facultativo médico a quien no se puede imputar por este hecho, el cometimiento de un delito.

Finalmente, con relación al “Planteamiento de los Problemas Jurídicos” es muy importante citar la reflexión que hace la Corte Constitucional en el acápite 35 de la sentencia en el cual, con mucha precisión y firmeza, señala (Corte-Constitucional, 2024):

La accionante manifiesta que el sufrimiento que produce la enfermedad grave e incurable o la lesión corporal grave podrá ser físico, en otros supuestos, señala que puede ser solo psíquico y, en otras partes de su demanda, indica que deben concurrir los dos elementos. Al respecto, este Organismo aclara que la presente sentencia y sus efectos se circunscribirán en que el padecimiento debe ser intenso, por lo que, los requisitos deben atender a cuestiones extremas de sufrimiento y superar un umbral de razonabilidad. Además, es indispensable que *el dolor intenso sea provocado por una lesión corporal irreversible que revista de gravedad o de una enfermedad que sea grave e incurable* (énfasis en el texto original) (Sentencia 67-23-IN/24, 05/02/2024, p. 16).

Se considera muy atinada esta reflexión de la Corte Constitucional, pues logra circunscribir la situación y análisis del pedido de la accionante, exclusivamente al caso de *enfermedad o lesión grave e incurable que genere un dolor intenso*, para evitar posibles distorsiones en la interpretación y alcance del pedido y, entonces, encasilla y analiza el problema jurídico, basado en la siguiente pregunta (Corte-Constitucional, 2024):

¿La aplicación de la sanción prevista en el tipo penal de homicidio es incompatible con los derechos a la vida digna y al libre desarrollo de la personalidad en el supuesto en el que (i) un médico ejecute la conducta tipificada en el artículo 144 del COIP cuando (ii) una persona, expresando su consentimiento inequívoco, libre e informado (o a través de su representante cuando no pueda expresarlo), solicite acceder a un procedimiento de eutanasia activa (iii) por el padecimiento de sufrimiento intenso proveniente de una lesión corporal grave e irreversible o una enfermedad grave e incurable? (Sentencia 67-23-IN/24, 05/02/2024, página 16)

Breve análisis de la parte resolutive de la sentencia

Primera parte de la sentencia

La Corte Constitucional en la parte resolutive de la sentencia declara la constitucionalidad condicionada del Art. 144 COIP. Pero, antes de avanzar se explica rápidamente, ¿qué es la constitucionalidad condicionada? Desde la doctrina, se encuentran algunas apreciaciones con respecto a la constitucionalidad condicionada. Martínez (2000), señala, al referirse a los diferentes tipos de sentencias, las que tienden a modular su contenido:

[T]ienen que ver con el contenido de lo acusado y, en ellas la Corte, en vez de retirar del ordenamiento la disposición acusada (inconstitucionalidad) o mantenerla sin modificaciones (constitucionalidad), emite un fallo que altera parcialmente ese contenido. [...] el tribunal restringe el alcance normativo de la disposición acusada, ya sea limitando su aplicación (no se aplica en tales casos), ya, sea limitando sus efectos (no sé, aplica tal o cual consecuencia). Estas sentencias interpretativas suponen entonces que se expulsa una interpretación de la disposición, pero se mantiene una eficacia normativa a la misma. Es de lejos, la técnica más usual [...]

Más adelante, el mismo autor, cita un fallo de la Corte Constitucional de Colombia (C-496/94) que hace alusión directa a la constitucionalidad condicionada y le da un contexto preciso para poder entender en qué casos es factible que se dé (Martínez, 2000):

De esta manera, las sentencias de constitucionalidad condicionada nacen del profundo respeto de la Corte Constitucional por las normas emanadas del poder legislativo, puesto que la Corte busca conservar en el ordenamiento jurídico a las leyes. En efecto, sólo cuando una norma legal no admite ninguna interpretación razonable que sea conforme con la Carta, procede esta Corporación a retirarla del ordenamiento legal. En cambio, cuando son posibles tales interpretaciones conformes a la Constitución, esta Corporación siempre ha mantenido las normas dentro del ordenamiento, puesto que es natural que la Corte siempre busque conservar las disposiciones emanadas del Legislador. *Por ello esta Corporación ha dicho que, si una “disposición legal admite varias interpretaciones, de las cuales*

algunas violan la Carta, pero otras se adecúan a ella, entonces corresponde a la Corte proferir una constitucionalidad condicionada o sentencia interpretativa que establezca cuáles sentidos de la disposición acusada se mantienen dentro del ordenamiento jurídico y cuáles no son legítimas constitucionalmente. En este caso, la Corte analiza la disposición acusada como una proposición normativa compleja que está integrada por otras proposiciones normativas simples, de las cuáles algunas, individualmente, no son admisibles, por lo cual ellas son retiradas del ordenamiento (C-496/94)” (énfasis del autor) (Martínez, 2000).

De esta forma, se puede afirmar que “una decisión de la Corte Constitucional que limita el contenido de una disposición legal en cuestión se conoce como constitucionalidad condicionada. En lugar de declararla inconstitucional por completo, la Corte establece condiciones o restricciones específicas en el ordenamiento jurídico para preservarla” (Cvik, 2024), como la constitucionalidad condicionada del Art. 144 del COIP, dictaminada por la Corte Constitucional de Ecuador, lo que permitió la eutanasia activa bajo determinados presupuestos. En síntesis, el propósito de esta figura es equilibrar la protección de los derechos constitucionales con la necesidad de mantener la regulación legal en situaciones específicas.

Ahora, *¿cuáles son las circunstancias que impone la Corte Constitucional del Ecuador a este caso?* De acuerdo con la sentencia, objeto de este análisis, se podrían sintetizar de la siguiente manera:

a. Primera circunstancia

Siempre y cuando no sea sancionado el médico que ejecute la conducta tipificada en el Art. 144 del COIP. *Condicionando*, este hecho a que (Corte-Constitucional, 2024): una persona solicite un proceso de eutanasia activa, *expresando su consentimiento inequívoco y libre e informado* o a través de su representante cuando no pueda informarlo.

b. Segunda circunstancia

Por el padecimiento de un sufrimiento intenso, proveniente de (Corte-Constitucional, 2024): una lesión, necesariamente corporal, grave

e irreversible o de una enfermedad que sea grave e incurable. En ambos casos, bajo estos supuestos (lesión y enfermedad descritos), la Corte Constitucional considera que la norma es incompatible con los derechos a (Corte-Constitucional, 2024): la vida digna y al libre desarrollo de la personalidad (autonomía).

La inviolabilidad de la vida admite excepciones cuando busca proteger otros derechos (Corte-Constitucional, 2024) (énfasis del autor). Para respaldar esta afirmación, la Corte Constitucional señala, el acápite 67 de la sentencia, lo siguiente (Corte-Constitucional, 2024):

Bajo esta proposición, la Corte IDH en el caso referido ha reiterado que se violará el derecho a la vida cuando “la privación se hubiera producido de manera arbitraria [...] por ser producto de la utilización de la fuerza de forma ilegítima, excesiva o desproporcionada”. De modo que aun “cuando la protección del derecho a la vida es un valor primordial dentro de la Constitución, este no puede ser interpretado de forma absoluta”. *Por consiguiente, la normativa convencional, constitucional y legal establecen supuestos en los que no es punible la privación de la vida cuando esta privación no es arbitraria o ilegítima.* (énfasis del autor) (Sentencia 67-23-IN/24, 05/02/2024, página 25).

Complementa esta reflexión, luego de citar, también los casos de aborto en que éste se haya practicado para evitar un peligro para la salud o vida de la mujer embarazada o, cuando el embarazo es consecuencia de violación, referidos en el Art. 150 del COIP; el contenido del acápite 70 ibidem, que señala (Corte-Constitucional, 2024):

De los ejemplos referidos, se colige una mínima intervención del Derecho Penal en razón de que el derecho a la vida no fue privado de forma arbitraria, pues la ejecución de la conducta se encuentra justificada de conformidad con la ley y en atención a la protección de otros derechos constitucionales. (Sentencia 67-23-IN/24, 05/02/2024, p. 25)

Termina el análisis sobre esta excepción a la *inviolabilidad a la vida*, manifestando, en el acápite 72 de la sentencia lo siguiente (Corte-Constitucional, 2024):

El derecho previsto en el artículo 66 número 1 de la CRE protege el derecho a la vida desde su dimensión de subsistencia y se encuentra resguardado por el artículo 144 del COIP frente a privaciones arbitrarias e ilegítimas. No obstante, en el supuesto planteado no se evidencia una conducta con dichas características, *ya que la privación de la vida ocurre con la aquiescencia y solicitud expresa del titular del bien jurídico, quien requiere acceder al procedimiento eutanásico por padecer intenso dolor como consecuencia de una lesión corporal grave e irreversible o una enfermedad grave e incurable*. En consecuencia, resulta controvertida la aplicación de la sanción contemplada en la norma al sujeto activo, ya que en el fondo no se protege la vida de un acto ilegítimo y arbitrario por las características del supuesto abordado (énfasis añadido) (Sentencia 67-23-IN/24, 05/02/2024, páginas 25 y 26).

En el contexto, la Corte discurre sobre las excepciones que contempla el derecho a la vida y considera que, en el caso planteado, cuando existe la aceptación, a través de voluntad expresa de quien quiere acceder al procedimiento de la eutanasia activa cuando padece un dolor intenso “como consecuencia de una lesión corporal grave e irreversible o una enfermedad incurable”, no existe una conducta penal cometida por el médico que, por un sentido de “compasión”, “ayuda” a morir a alguien bajo las circunstancias expuestas; tal conducta no puede ser sancionada bajo la esfera del Derecho Penal.

Se hace aquí una pausa, para entender rápidamente: *¿qué es la eutanasia activa? Y ¿cómo se diferencia de la eutanasia pasiva?* “La eutanasia pasiva se refiere a la supresión o no aplicación de medidas que mantienen o podrían mantener a una persona con vida, y como resultado, la persona fallece” (BBC-News-Mundo, 2019). Por su parte, “la eutanasia activa es en la interrupción o rechazo de los tratamientos médicos “que conllevan a acelerar la muerte de manera que la causa del deceso siempre será la enfermedad subyacente¹” (Corte-Constitucional, 2024) (Sentencia 67-23-IN/24, 05/02/2024, p. 18)

1 Colombia, Corte Constitucional, Sentencia C-233/21, 22 de julio de 2021.

En este caso, no se interviene activamente en el proceso de la muerte. Por ejemplo, se pueden suspender tratamientos médicos o retirar dispositivos de soporte vital. En contraste, la eutanasia activa implica acciones deliberadas para causar la muerte de manera rápida y eficaz, a petición expresa y reiterada del paciente. Es crucial destacar que la eutanasia activa se lleva a cabo a petición expresa y reiterada del paciente y se considera un acto de compasión por el sufrimiento insoportable causado por una enfermedad que no se puede curar (*Expreso*, 2024).

En el caso de lo resuelto por la Corte Constitucional, a través de la sentencia 67-23-IN/24, lo que se prevé es la eutanasia activa, que implica que es el paciente, quien, al cumplir los condicionamientos previos, requiere su aplicación; es decir, solicita que un facultativo médico aplique un procedimiento, previamente explicado al paciente, para terminar con su vida. La propia Corte, dentro del acápite 43 de la sentencia, hace alusión al contexto de la *eutanasia activa* en los siguientes términos (Corte Constitucional, 2024):

[...] la *eutanasia activa* es el procedimiento que a petición de parte o por un representante en caso de que el paciente no pueda expresar su voluntad es realizado por un médico para poner fin a la vida de quien padece sufrimiento insoportable proveniente de una lesión corporal grave e irreversible o una enfermedad grave e incurable. (Sentencia 67-23-IN/24, 05/02/2024, p. 18)

Ahora bien, de acuerdo con la sentencia, *¿cómo se llega a la eutanasia activa?* Quien “padece intenso dolor como consecuencia de una lesión corporal grave e irreversible o una enfermedad grave e incurable” (Corte Constitucional, 2024), *expresa su consentimiento*: 1) inequívoco (que no implique ninguna ambigüedad; esto es, que no exista posibilidad de error, indeterminación o fallo y 2) libre e informado (con pleno conocimiento de causa por parte de quien lo acepta.)

Otorga, además la posibilidad de que este consentimiento se genere a través de un representante cuando no pueda expresarlo el paciente directamente; al respecto, considero que, la forma idónea de ejecutar esta

voluntad sería mediante un documento escrito previamente por quien lo solicita, presentado por un tercero; por ejemplo, un *testamento vital*; pero: ¿qué es esto del testamento vital?

El testamento vital o documento de voluntades anticipadas, es un instrumento que expresa la voluntad de una persona con respecto a su tratamiento médico, así como el destino de su cuerpo u órganos en caso de enfermedad o fallecimiento. Podría incluir instrucciones como: aceptar o rechazar tratamientos paliativos específicos; establecer límites a su tratamiento médico; tomar una decisión sobre la donación de órganos; o, designar un representante que pueda interpretar sus voluntades o tomar decisiones que no se encuentran en su testamento vital. (Testamento Vital, s/f)

De acuerdo con el literal “a”, del Art. 5 del Reglamento del procedimiento para la aplicación de la eutanasia activa voluntaria y avoluntaria (Registro-Oficial-538, 2024), de 8 de abril de 2024, este instrumento (testamento vital), debe ser otorgado ante Notario Público.

Dentro de la sentencia, objeto de este análisis, en su acápite 45, sobre la alusión a otras formas de eutanasia; a saber: eutanasia activa voluntaria y eutanasia activa avoluntaria, la Corte, textualmente, explica (Corte-Constitucional, 2024):

A la luz de las definiciones expuestas, este Organismo considera que la eutanasia activa voluntaria, activa avoluntaria y pasiva se fundamentan en la voluntad del paciente. En la *eutanasia activa voluntaria* el paciente expresa la decisión de morir a través de un procedimiento eutanásico. En la *eutanasia activa avoluntaria* no se puede conocer la voluntad del paciente por la imposibilidad de expresarla, por ejemplo, en los casos en los que las personas se encuentran en estado vegetativo, coma permanente y, en su lugar, quien consciente es un representante². Y, en la *eutanasia pasiva* quien toma la decisión libre, responsable e informada respecto a negarse a recibir un tratamiento médico que lo conduce a morir, es el paciente. En los tres supuestos, el acto eutanásico no persigue aliviar el

2 Colombia, Corte Constitucional de Colombia, Sentencia C-233/21, 22 de julio de 2021, pie de página 144.

sufrimiento, sino poner fin a la vida para terminar con el padecimiento insoportable ocasionado por *una lesión corporal grave e irreversible o una enfermedad grave e incurable*. (Sentencia 67-23-IN/24, 05/02/2024, página 18) (énfasis del autor)

Estas definiciones y su contexto aluden al particular hecho de poner fin a la vida del paciente en virtud de terminar con el padecimiento insoportable ocasionado por la lesión o enfermedad grave, irreversible e incurable.

Ahora, puntualizando: ¿en qué casos se puede aplicar la eutanasia de acuerdo con la sentencia? El primer caso se da frente al padecimiento de sufrimiento intenso, proveniente de (Corte-Constitucional, 2024): *una lesión necesariamente: corporal, grave, irreversible*. El segundo caso es una enfermedad que sea grave e incurable (las denominadas como: catastróficas, raras y huérfanas o de alta complejidad), que igualmente genere el padecimiento de sufrimiento intenso para el paciente. En ambos casos (bajo estos supuestos), de acuerdo con la sentencia de la Corte Constitucional, la norma es incompatible con los derechos (Corte-Constitucional, 2024): a la vida digna y libre desarrollo de la personalidad (autonomía); conforme se analizó previamente.

Es importante considerar que la Corte ha dispuesto *la objeción de conciencia de los médicos en el proceso de eutanasia activa*. *¿Qué es o, qué implica, en términos generales, la objeción de conciencia?*

La objeción de conciencia es el derecho a no cumplir una tarea u obligación debido a razones religiosas o éticas. Se produce cuando una persona siente que el deber en cuestión atentaría contra sus convicciones personales. Es un derecho constitucional reconocido en diversos marcos legislativos y se entiende que el sujeto puede actuar de acuerdo con sus convicciones dentro de los límites fijados por la ley. (Grande Yeves, 2019)

En el Art. 3 del Reglamento del procedimiento para la aplicación de la eutanasia activa voluntaria y avoluntaria, de 8 de abril de 2024, dentro

del marco de definiciones que éste determina, encontramos como objeción de conciencia lo siguiente (Registro-Oficial-538, 2024): “Es el derecho constitucional de los profesionales a no atender aquellas demandas de actuación que resultan incompatibles con sus propias convicciones. No podrá ser restringido bajo ninguna circunstancia”. La Corte Constitucional, dentro de las consideraciones finales, acápite 5, de la sentencia, al respecto, manifiesta (Corte-Constitucional, 2024):

El respeto y salvaguarda a la objeción de conciencia del sujeto calificado (médico). El artículo 66, número 12 de la CRE reconoce el *derecho a la objeción de conciencia*. Con fundamento en lo anterior, “[n]adie puede ser objeto de medidas restrictivas que puedan menoscabar la libertad de conservar [...] o de cambiar de religión o de creencias”³

Así, la objeción de conciencia es un derecho que permite a una persona abstenerse de participar en actividades, servicios o prácticas que van en contra de sus convicciones éticas, morales o religiosas, por lo que, puede actuar de manera consecuente con sus ideas, valores y principios sin hacer daño a terceros. Este reconocimiento tiene una íntima relación con otros derechos como la libertad de pensamiento, de religión, de expresión e inclusive con el libre desarrollo de la personalidad, pues todos “guardan un vínculo sustancial e indisoluble al ser indispensables para el desarrollo de la personalidad y como garantías de la protección de la dignidad humana”⁴. (Sentencia 67-23-IN/24, 05/02/2024, p. 36)

En definitiva, se otorga, a través de sentencia, al facultativo médico el derecho a la objeción de conciencia, con respecto a la aplicación o no del pedido del paciente con respecto a la ejecución de la eutanasia activa.

Segunda parte de la sentencia

En esta se declara la constitucionalidad aditiva del Art. 6 del Código de Ética Médica. ¿Qué es la constitucionalidad aditiva? También conocida como sentencia aditiva, es cuando el Tribunal Constitucional,

3 Convención Americana sobre Derechos Humanos, art. 12.

4 CCE, Dictamen 3-22-OP/22, 03 de octubre de 2022, párr. 22.

en este caso la Corte Constitucional, decide que una disposición legal es inconstitucional sin alterar su texto. Esto da la posibilidad de ampliar o extender el contenido de la normativa, incluyendo su aplicación a situaciones que no fueron contempladas en la disposición original o ampliando sus consecuencias jurídicas (Derecho-UNED, 2024). En pocas palabras, es un método para declarar que un artículo es inconstitucional sin anularlo, sino más bien ampliando su alcance dentro del marco constitucional.

Otra definición de sentencia aditiva, del Diccionario Panhispánico de Español Jurídico (RAE, 2023), manifiesta que es una “[s]entencia manipulativa que, sin incidir en el texto de una disposición legal, establece la inconstitucionalidad de un precepto, produciendo el efecto de ampliar o extender su contenido normativo, permitiendo su aplicación a supuestos no contemplados expresamente en la disposición, o ampliando sus consecuencias jurídicas”. Cabe acotar que las llamadas sentencias manipulativas, son, en general, sentencias interpretativas.

En el hilo que nos ocupa, la Corte Constitucional, declara a constitucionalidad aditiva del Art. 6 del Código de Ética Médica; lo que implica esto, es que, su aplicación estará sujeta condiciones, conforme se explica a continuación.

El Art. 6 del Código de Ética médica, señala: “[e]l médico que desde que es llamado a atender a un enfermo, se hace responsable de proporcionarle todos los cuidados (médicos) necesarios para que recupere su salud”. Que luego añade: “su presponsabilidad mayor será la conservación de la vida del enfermo”. (énfasis añadido) Entonces, se prevé, a través de la constitucionalidad aditiva, la siguiente excepción para esta regla:

Excepto cuando la persona solicite acceder a un procedimiento de *eutanasia activa*; en caso de (Corte-Constitucional, 2024): a) padecimiento de intenso sufrimiento, proveniente de un sufrimiento corporal, grave e irreversible, b) padecimiento de una enfermedad grave e incurable. Para esto se requiere que el paciente otorgue su *consentimiento*: a) inequívoco, b) libre e informado. O, a través de su representante cuando no pueda hacerlo personalmente.

Tercera parte de la sentencia

Declara la *inconstitucionalidad* del Art. 90 del Código de Ética Médica, que queda fuera del ordenamiento jurídico; es derogado a través de la sentencia, que lo declara inconstitucional, conforme lo dispuesto en el acápite 110, página 35, de la sentencia 67-23-IN/24. El Art. 90, ibidem, señala que el médico no está autorizado para abreviar la vida del enfermo: “su misión fundamental frente a una enfermedad incurable será aliviarla mediante los recursos terapéuticos del caso”

Cuarta parte de la sentencia

Dispone que el *defensor del pueblo*, en el plazo de 6 meses, desde la notificación de la sentencia: *prepare un proyecto de ley que regule los procedimientos eutanásicos* (Corte-Constitucional, 2024). Es importante anotar que, con fecha 28 de junio de 2024, la Defensoría del Pueblo remitió, mediante Oficio Nro.DPE-DPE-2024-0565-O, el Proyecto de Ley que regula los procedimientos eutanásicos en cumplimiento de la sentencia de la Corte Constitucional No. 67-23-in/24, a la Asamblea Nacional, para el tratamiento legislativo correspondiente.

Quinta parte de la sentencia

Dispone que el Ministerio de Salud, en el plazo máximo de 2 meses, desde la notificación de la sentencia; *prepare un reglamento para la aplicación de la eutanasia activa voluntaria* que tendrá vigencia hasta la aprobación de la Ley (Corte-Constitucional, 2024). Es importante acotar, que dicho reglamento se generó por parte Ministerio de Salud, con fecha 08/04/2024, mediante Acuerdo Ministerial No. 59, que expide el Reglamento del Procedimiento para la Aplicación de la Eutanasia Activa Voluntaria y Avoluntaria. Publicado en el Tercer Suplemento N.º 538 - Registro Oficial 12/04/2024 (Registro-Oficial-538, 2024). Por lo que actualmente se encuentra vigente y contiene, en efecto, el procedimiento para la aplicación de eutanasia, en caso de las personas que quieran acogerse a la misma, cumpliendo los requisitos allí establecidos.

Sexta parte de la sentencia

Dispone que la Asamblea Nacional, en el plazo máximo de 12 meses, desde la presentación del proyecto de ley; *conozca, discuta y expida una ley que regule los procedimientos eutanásicos* (Corte-Constitucional, 2024). Como se mencionó en el acápite precedente este proyecto de ley se encuentra en conocimiento de la Asamblea Nacional.

Al respecto, en el acápite “Conclusiones y recomendaciones” del Informe técnico-jurídico no vinculante, No.0226-INV-UTL-AN-2024, suscrito por el Coordinador General de la Unidad de Técnica Legislativa de la Asamblea Nacional, de 23 de julio de 2024, textualmente, señala:

El “Proyecto de Ley que Regula los Procedimientos Eutanásicos”, sujeto a análisis, *no cumple* con los requisitos formales establecidos en los artículos 136 de la Constitución y 56, número 3 de la Ley 16 Orgánica de la Función Legislativa, en relación con que el Proyecto de Ley debe presentarse con la expresión clara de los artículos que con la nueva ley se derogarían o se reformarían.

Sobre la base de lo expuesto, la Unidad de Técnica Legislativa recomienda al Consejo de Administración Legislativa: a) considerar, los criterios establecidos en el presente Informe y b) no calificar, “Proyecto de Ley Orgánica que Regula los Procedimientos Eutanásicos”. Lo que implica que, se debería devolver el proyecto de Ley a la Defensoría del Pueblo para que se acojan las observaciones de la Unidad de Técnica Legislativa, previa su aceptación al trámite legislativo pertinente, de conformidad con la Ley, con lo que el plazo de 12 meses dado por la Corte a la Asamblea Nacional, empezaría a correr una vez que el Defensor del Pueblo subsane tales observaciones.

Conclusiones

La sentencia No. 67-23-IN/24 abre la puerta a la eutanasia en Ecuador, convirtiendo al país en el noveno en el mundo en incorporar esta

práctica en su ordenamiento jurídico. El fallo autoriza únicamente la eutanasia activa voluntaria, entendida como el acto médico intencional para provocar la muerte a petición del paciente. Asimismo, reconoce y protege la objeción de conciencia del profesional de la salud implicado.

Su aplicación se circunscribe a supuestos estrictos: pacientes que sufran un dolor intenso e intolerable derivado de una lesión corporal grave e irreversible o de una enfermedad grave e incurable. Para hacer efectivo el mandato judicial, el Ministerio de Salud emitió el Acuerdo Ministerial No. 59, que reglamenta el procedimiento para la eutanasia activa voluntaria y avoluntaria. Paralelamente, en agosto de 2024 se encuentra en trámite en la Asamblea Nacional el Proyecto de Ley que regulará estos procedimientos, a fin de dotar de seguridad jurídica y claridad al proceso.

Corresponde al legislador precisar, con rigor y equilibrio, las condiciones y garantías que salvaguarden la dignidad del paciente y el respeto a la autonomía individual. Es esencial difundir que la eutanasia no constituye un recurso indiscriminado, sino una medida extrema frente a un sufrimiento que no admite otro alivio.

El caso de Leonardo Aníbal Caiza, diagnosticado con esclerosis lateral amiotrófica, marca el primer procedimiento eutanásico conforme a la normativa vigente. Aún es temprano para prever los efectos sociales y éticos de esta nueva realidad jurídica; persisten más interrogantes que certezas, muchas de ellas arraigadas en la esfera íntima de cada persona.

Referencias bibliográficas

- Allianz Assistance (2023) “¿Qué es un testamento vital y cómo hacerlo?”. <https://bit.ly/3KNysxM>
- BBC News Mundo (7/06/2019). Qué son la eutanasia pasiva y activa y en qué se diferencian del suicidio asistido. <https://bit.ly/3IOmK5r>
- Clínica de la Universidad de Navarra (2023) *Diccionario médico: autonomía del paciente*. <https://bit.ly/46REiWp>
- Corte Constitucional Ecuador (05/02/2024). Sentencia 67-23-IN/24. <https://bit.ly/3JJKLLk>

- Correa Montoya, L. (2021). Muerte Digna. Lugar Constitucional y Núcleo Esencial de un Derecho Humano Emergente. *Opinión Jurídica*, 20(41), 127-54. <https://doi.org/10.22395/ojum.v20n41a4>
- Cvik, J. (16/02/2024). Homicidio Simple: Un Análisis de su Condicionalidad Constitucional. Meythaler y Zambrano Abogados. <https://bit.ly/4m2t9aQ>
- Derecho UNED (2024). Clases y características de las Sentencias del Tribunal Constitucional. <https://bit.ly/4p7QTgB>
- Diario El Universo* (20/11/2023). “Lo que vivo es doloroso, solitario y cruel, dice Paola Roldán en la audiencia de su pedido de acceso a la eutanasia. <https://bit.ly/4m6r4Lc>
- Diccionario Panhispánico de Español Jurídico (2023). <https://bit.ly/48OWEK6>
- Expreso* (08/02/2024). ¿Qué es la eutanasia y cómo se aplicaría en Ecuador? <https://bit.ly/4phcHXe>
- Grande Yeves, P. (2019). La objeción de conciencia en España. *La mente es maravillosa*. <https://bit.ly/4npR5q9>
- INCMNSZ (28/12/2017). “¿Qué son los cuidados paliativos?”. Gobierno de México. Disponible: <https://bit.ly/45RClDl>
- Institut de Drets Humans de Catalunya (2009). Declaración Universal de Derechos Humanos Emergentes. <https://bit.ly/3I8E3hk>
- Martínez Caballero, A. (Jan/June 2000). Tipos de sentencias en el control constitucional de las leyes: La experiencia colombiana. *Estudios Socio-Jurídicos*, 2(1). Bogotá. <https://bit.ly/4mZNwHa>
- MedicinePlus (2022). “¿Qué son los cuidados paliativos?”. <https://bit.ly/4nYCL7J>
- Ministerio de Salud (08/04/2024). Acuerdo Ministerial No. 59, que expide el Reglamento del Procedimiento para la Aplicación de la Eutanasia Activa Voluntaria y Avoluntaria. Publicado en el Tercer Suplemento N.º 538 - Registro Oficial 12/04/2024.
- Naciones Unidas (s/f). Declaración Universal de los Derechos Humanos. <https://bit.ly/4gQ7AcH>
- Romeo Casabona, C. (2021). La ley orgánica reguladora de la eutanasia y la adaptación del código penal. *BioLaw*, (2), 283-314. <https://bit.ly/48bazKn>
- Testamento Vital (s/f). <https://bit.ly/4gMwERA>
- Todoinfolegal.com (s/f). Los límites de la libertad: donde empiezan tus derechos y terminan los míos. <http://bit.ly/3Kt67wN>
- Unidad de Técnica Legislativa de la Asamblea Nacional (23 de julio de 2024). Informe técnico-jurídico no vinculante, No.0226-INV-UTL-AN-2024. <https://bit.ly/3JPPtqO>