DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

¿Quién responde por la inteligencia artificial? Ética, sesgos y responsabilidad en la investigación y la escritura académica

Gezabel Guzmán*
Aliber Escobar**

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) indica cuatro valores fundamentales que sientan las bases para que los sistemas de inteligencia artificial (en adelante, IA) trabajen por el bien de la humanidad. Estos son:

1. derechos humanos y dignidad humana,
2. vivir en sociedades pacíficas,
3. garantizar diversidad e inclusión,
4. florecimiento del medio ambiente así como de los ecosistemas.

* Profesora investigadora en la Academia de Comunicación y Cultura del Colegio de Humanidades y Ciencias Sociales de la UACM. Forma parte del Comité editorial de *Andamios. Revista de Investigación Social* y es miembro del Sistema Nacional de Investigadoras e Investigadores. ORCID: 0000-0001-9926-5401 Correo electrónico: gezabel.guzman@uacm.edu.mx

** Profesor investigador en la Academia de Expresión Oral y Escrita del Colegio de Humanidades y Ciencias Sociales de la UACM. Forma parte del Comité Editorial de *GOSOY. Revista del Centro Lacaniano de Investigación en Psicoanálisis* y es miembro del Sistema Nacional de Investigadoras e Investigadores. ORCID: 0000-0001-9163-0515 Correo electrónico: aliber.escobar@uacm.edu.mx

Específicamente en relación al enfoque de derechos humanos y dignidad humana, la ética*** en el uso de la IA debe tomar en cuenta:

1. Proporcionalidad e inocuidad, ámbitos referidos al uso que no cause daños e injusticias, a nivel individual como comunitario, que puedan repercutir en el entorno social, cultural, económico, político, natural, es decir, de la vida en general.
2. Seguridad y protección, que no sólo se refiere a la protección tecnológica, sino también el fomentar buenas prácticas, por ejemplo, en no crear noticias falsas, suplantación de identidad, o generar información que perpetúe discursos de discriminación.
3. Derecho a la intimidad y protección de datos.
4. Gobernanza en colaboración de múltiples partes interesadas, para incluir diversidad de voces, tomando en cuenta sectores considerados vulnerables y promover el diálogo entre países, sectores tecnológicos, académicos, así como, sociales.
5. Responsabilidad y rendición de cuentas, para tener mecanismos de supervisión, evaluación y medición de amenazas al bienestar medioambiental.
6. Transparencia y explicabilidad, que debe revelar cómo funciona la IA, por ejemplo, qué datos utiliza, qué lógicas generales sigue.
7. Supervisión y validación para que los sistemas automatizados no operen sin control humano.
8. Sostenibilidad, para usar la tecnología en pro de resolver problemas ambientales y sociales; pero el término se refiere también a transparentar y promover la reflexión en el costo ambiental por el uso de la inteligencia ambiental.
9. Sensibilización y educación, enfocada, entre otros aspectos, a la alfabetización mediática.

*** Antónimos de la palabra ética podrían ser inmoral, deshonesto, turbio. Mientras que sus sinónimos aluden a lo íntegro, decente, honrado. Pero ¿qué es la ética?, ¿qué implicaciones tiene ésta en el uso de la IA?, y ¿qué tiene que ver con el mundo de la escritura y la investigación académica? Para la Real Academia de la Lengua Española (2024), la ética es el conjunto de normas morales que rigen la conducta de la persona en cualquier ámbito de la vida. Esta definición nos sirve como marco en el presente artículo para distinguir la ética

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

10. Equidad y no discriminación, para promover el acceso de la IA desde ámbitos inclusivos y no violentos (UNESCO, 2024).

Sin embargo, la realidad muestra que esos valores se han quedado apenas en ideales, porque casi todos han sido ignorados, transgredidos o tergiversados de manera discrecional, engañosa u oculta, por las élites creadoras de las distintas IAs existentes. Como muestra claramente Karen Hao (2026) a partir de su investigación y las múltiples entrevistas que realizó a ejecutivos y empleados de *OpenAI, Microsoft, Anthropic, Meta, Google, DeepMind* y *Scale* durante los últimos siete años.

Por otra parte, como bien señalan Alberto Quintavalla y Jeroen Temperman, en un trabajo que han realizado para plantear a la IA cuestiones legales, éticas, tecnológicas y sociales para analizar tanto las áreas que actualmente permanecen oscuras, como el impacto, el daño y los riesgos de la IA en los derechos humanos, a pesar de que sus historias se solapan, ambos fenómenos –la IA y los derechos humanos– llevaron existencias bastante separadas durante sus primeros cincuenta o sesenta años, más o menos. Solo en la última década sus caminos han convergido, las dos fuerzas se han encontrado, se han apoyado mutuamente o, como también puede ocurrir, han entrado en conflicto, provocando enfrentamientos de menor o mayor envergadura (Quintavalla y Temperman, 2023).

Lo anterior también permite advertir que la IA no ha sido siempre usada de forma responsable, tomando en cuenta los requisitos establecidos legalmente tanto por las instancias internacionales como por los gobiernos. A partir del reconocimiento de lo anterior, en este artículo nos proponemos abordar dos problemas que desde los inicios de la IA permanecen irresolubles, pues como nos interesa analizar las posibilidades de la IA para investigar y escribir académicamente, debemos tener muy claro de quién es

de la moral. Sobre todo, porque como explica García-Clarck, citando a Velasco, la "ética es una reflexión crítica y racional sobre la moralidad vigente en una sociedad determinada" (2025, p. 5). Es decir, "la moral está relacionada con las conductas, normas y valores morales, vigentes en una sociedad, mientras que la ética se constituye por descripciones, explicaciones y evaluaciones sobre tales elementos del mundo moral" (García-Clarck, 2025, p. 6). Ambos conceptos se encuentran presentes de forma interrelacionada y se presentan de manera

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

la responsabilidad en el uso de la IA para la investigación y cómo lidiar con los sesgos inherentes a su utilización.

1. Inteligencia artificial: fundamentos, sesgos y responsabilidad ética en la investigación académica

Han pasado 70 años desde que el término "inteligencia artificial" (aparentemente, una *contradictio in adjecto*) fue acuñado por John McCarthy y Marvin Minsky, para nombrar una rama particular de las ciencias computacionales (Born, 2018). En este artículo vamos a partir de la definición un poco modificada por Rainer P. Born e Ilse Born-Lechleitner, de la que propuso Edward Feigenbaum en 1981, a saber:

> La IA es aquella parte de la informática que se ocupa del diseño de sistemas informáticos inteligentes, es decir, sistemas que muestran las características que asociamos con la inteligencia en el comportamiento humano –por ejemplo, la comprensión del lenguaje, el aprendizaje, el razonamiento, la resolución de problemas, etc.– (Born, 2018, p. VIII).

De hecho, es el cerebro humano, a partir del conocimiento de sus interconexiones neuronales en red, la base de inspiración para crear modelos o algoritmos (*machine learning* y *deep learning*) para hacer predicciones o tomar decisiones basadas en datos, lo cual permite a las computadoras hacer inferencias y aprender tareas sin que estén específicamente programadas para éstas (Stryker y Kavlakoglu, 2024; Copeland, 2026).

Actualmente, existen diferentes tipos de IA, por ejemplo, la "IA estrecha" empleada para realizar tareas específicas como asistentes de voz (Alexa,

constante en cualquier ámbito de la vida cotidiana, por ejemplo, ante la eutanasia, se evoca a la ética para que se aporten argumentos a favor o en contra de ésta con el fin de complejizar un acto que conlleva implicaciones sobre la vida y la muerte humana. Así, "el planteamiento teórico, valorativo y prescriptivo, que ofrece la ética, puede interpelar a los individuos o grupos involucrados en algún dilema moral y derivar en prácticas concretas" (García-Clarck, 2025, p. 7). Por tanto, la ética indudablemente desemboca en una acción, pero en una acción

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

Siri) y los chatbot de redes sociales usados para simular conversaciones humanas. O la "IA generativa" para crear contenido complejo, imágenes de calidad, video, audio, textos. Nos vamos a centrar en la generativa, porque lo que nos interesa es abordar la ética en la investigación y la elaboración de textos académicos, pero sobre todo la reflexión en torno a los sesgos presentes en la inteligencia artificial.

Sabemos que hoy la IA puede realizar la escritura de un artículo académico que asevere una serie de tesis muy polémicas, aunque, al no ser humana, todavía no puede asumir una posición moral sobre las mismas, ni tampoco puede ser obligada a responder en caso de que sean infamias. No obstante, a pesar de que no debe ser considerada "autora", en sentido humano, porque tradicionalmente la creación de una obra depende de un ejercicio intelectual, moral, original y singular de un ser humano (Penabad-Camacho, Morera-Castro y Penabad-Camacho, 2024), es verdad que, al tener ya décadas de entrenamiento humano, la IA hoy crea artículos, libros, piezas musicales, guiones de teatro, películas y series, audiovisuales, etc., tan semejantes a las grandes obras del acervo universal gracias a que ha sido alimentada con todas ellas, por lo que resulta muy difícil, sino imposible, distinguirlas de las que han sido realizadas por un humano.

Lo anterior es posible porque, si el aprendizaje automático de la IA se basa en el flujo de información que el propio ser humano introduce, es comprensible que sea capaz de reproducir fielmente los grandes logros civilizatorios que son ingresados sistemáticamente a su base de datos. Aunque, insistimos, no pueda hacerse cargo de las consecuencias de estos. Según Murray, el aprendizaje automático "es lo que permite que un buscador nos haga recomendaciones. Nos diga cómo llegar a una dirección o incluso nos permita traducir textos [...]. El aprendizaje automático, permite que las computadoras resuelvan los problemas mediante el reconocimiento de patrones en los datos" (2020, p. 143), que millones de seres humanos discriminan puntualmente.

humana, porque al igual que la moral es un fenómeno humano. Ante ello la IA no tiene ética, ni conciencia, es una herramienta que procesa datos proporcionados exclusivamente por los humanos que la programan y la utilizan. Por ello, podemos encontrar en la IA sesgos culturales y sociales completamente humanos, es decir, ésta es un reflejo de nosotros como empleadores y de ciertas sociedades, sobre todo, de aquellas dominantes a nivel global bajo el uso de determinado idioma, y de ciertos valores que respondan a la lógica del capital. Por su

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

Ahora bien, así como lo anterior permite entender cómo es que la IA puede replicar lo humano, también permite advertir que, a pesar de que es capaz de tener iniciativas provocadas por los algoritmos, al final sólo reproduce lo que obtiene de los humanos, sin que haya una deliberación o voluntad singular que la pueda hacer responsable de lo que dice o hace. Por tanto, lo que la IA reproduce son los valores y normas de las élites que la han creado. Mismos que, según Sam Altman, al haber sido cabildeados con ciertos gobiernos de todo el mundo, se imponen como imperativos morales que Bill Gates, Elon Musk, Dario Amodei, Ilya Sutskever, etc., buscan cumplir con objetivos falaces: cómo solucionar el cambio climático o eliminar la necesidad de trabajar, y para ello están arrasando con los recursos naturales y explotando a millones de personas. Es decir, son inmorales y además ocultan el verdadero objetivo del transhumanismo en el que, en un futuro no muy lejano, el mundo ya no dependerá de las personas. Dicho de otro modo, será un mundo amoral gobernado por la IA.

Ese futuro es el que, como bien señala Lorenzo Ramírez, Raymond Kurzweil proyecta en el año 2045:

> En el que espera que la inteligencia artificial (IA) haya adquirido el desarrollo necesario para alcanzar lo que denomina como la 'singularidad', en el momento en el que la tecnología se creará a sí misma e incluso se adentrará en la biología y la neurociencia prácticamente de forma autónoma, con el mínimo componente humano. (Ramírez, 2025, p. 324)

Por otro lado, hoy la inteligencia artificial es omnipresente en el ámbito académico. Y puede ser muy útil como una herramienta de asistencia técnica para traducir textos, corregir estilos, procesar datos, redactar borradores, corregir redacción, ortografía y generar gráficos. Así como para desempeñarse como un asistente de investigación que identifique información específica

parte, para Brugger (2005), la ética va más allá de una ciencia social descriptiva de lo moral que describe y analiza las valoraciones fácticas de otros. Más bien es una disciplina filosófica que emplea un método para reflexionar sobre los datos de la experiencia moral (conceptos y principios morales), y no consiste en un empirismo o en el recurso de un sentimiento no analizable, sino en el uso de la razón. Es decir, entre otros aspectos, permite generar una

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

en múltiples textos. O como herramienta de búsqueda para que autores encuentren rápidamente literatura sobre temas concretos, mientras realizan otras actividades. Asimismo, puede resumir artículos académicos en minutos y ayudar a que editores y revisores puedan mejorar la redacción, detectar similitudes, verificar la integridad de imágenes, comprobar la originalidad de un documento y muchas cosas más.

Sin embargo, hay revistas como *Science* y *Nature* que prohíben su uso; incluyendo el empleo de aplicaciones *Large Language Models* (LLM) para, entre otros aspectos, realizar traducción de textos. Aunque redes regionales como el Sistema Regional de Información en Línea para Revistas Científicas de América Larina, el Caribe, España y Portugal (Latindex) y *Scientific Electronic Library Online* (SciELO), se encuentran abiertas a su utilización bajo ciertos marcos, como la responsabilidad, las obligaciones y la transparencia de los autores y las autoras. Es decir, exigen la responsabilidad pública de autoría, y para ello requieren que todo el material citado se atribuya correctamente, así como indicar expresamente cualquier uso de IA, lo cual busca, en la medida de lo posible, controlar la transparencia y la integridad en la investigación.

En consecuencia, muchas revistas actualmente colocan políticas para el uso de la IA al momento de presentar un artículo, que incluyen declaración de uso y que promueven la no manipulación de datos, el no plagio y sobre todo, que reconocen el valor de la supervisión humana para asegurar la originalidad, la coherencia, la calidad y la identificación de sesgos. Lo que indica que la ética en la investigación que utiliza IA es un terreno espinoso y todavía parcialmente incógnito. De hecho, Michael James Heron y Pauline Helen Belford partieron de un libro de ficción publicado por Richard G. Epstein (1997) a finales del siglo pasado, sobre el caso de un robot asesino y los problemas éticos que ocasiona sobre la responsabilidad y la complicidad de las instancias involucradas, para abordar un estudio de caso de la ética académica en la IA.

reflexión razonada de las decisiones que toman las personas. Ante ello, el ser humano al ser un ser social, realiza acciones en el marco de la moral, dicho actuar debe ser razonado, consciente y conlleva responsabilidad en cada individuo. Así, la responsabilidad en el uso de la inteligencia artificial recae sobre la persona que la emplea, aunque para realizar un uso razonado de la IA es vital conocer cómo opera ésta y qué implicaciones conlleva en el

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

En su texto, aclaran una cuestión que nos parece fundamental cuando hablamos de la ética tecnológica, a saber, que "[l]a responsabilidad ética y moral respecto a la tecnología –cómo hablamos de ella y cómo la utilizamos– siempre reside en el componente humano de la ecuación" (Heron y Belford, 2025, p. 13). Dicho de otra forma, que mientras no se demuestre lo contrario, no debemos caer en la ilusión de que la IA puede llegar a ser independiente de sus creadores y responsable de "sus" errores. Pues, actualmente, la IA no es una "entidad" equivalente a un ser humano ni independiente del mismo. Prueba de ello es que, a pesar de que películas como *Mission: Impossible - The Final Reckoning* (McQuarrie, 2025), reproducen la ilusión de que llegará un momento en que la IA adquirirá total independencia del ser humano, la imagen que de ella han hecho sus propios programadores es de una tecnología que ha sido creada –por ellos mismos– "a su imagen y semejanza" (Hao, 2026, p. 35). Es decir, que la IA no se ha creado sola, ni tampoco puede volverse en contra de la humanidad por voluntad propia. Recordemos que el famoso computador inteligente llamado HAL 9000 en *2001: A Space Odyssey* (Kubrick, 1968) o la inteligente MOTHER en *Alien* (Scott, 1979), se negaban a obedecer las órdenes de los tripulantes de las naves porque habían sido programados previamente para una misión, pero no por voluntad propia.

Por consiguiente, no debemos confundirnos en relación con la instancia responsable en el trabajo de investigación. Si hoy alguno de nuestros asistentes de IA comete algún error o delito de derechos de autoría, quien debe pagar, además de sus creadores, son sus usuarios. Es decir, los investigadores. Pues, al ser una tecnología que carece de calidad moral, nosotros somos los responsables de cualquier imprecisión, plagio u omisión.

ámbito social, económico, cultural y ecológico. Por ejemplo, los microchips que alimentan la IA necesitan elementos de tierras raras, que a menudo se extraen de formas destructivas para el medio ambiente. Los centros de datos, incluidos los operados por proveedores de servicios en la nube, producen residuos que contienen mercurio y plomo que contaminan mantos acuíferos y entran en el subsuelo dañando poblaciones y ecosistemas. Sumado a ello, se emplea un uso desmedido de agua para enfriar componentes eléctricos con los cuales

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

2. La dimensión humana en la construcción de los sesgos en los sistemas de inteligencia artificial

De este modo, la IA que ha sido entrenada durante décadas para realizar investigación y coadyuvar a la escritura, comete los mismos errores con los cuales se le programó, sea por ignorancia, descuido, cansancio o prejuicios por parte de los alimentadores, o falta de pericia por parte de los usuarios y usuarias. De ahí que tenga los mismos sesgos humanos, tanto de programadores como de sus creadores. Muchas veces orgánicos (producto de quienes la han alimentado) y otras tantas intencionales; de acuerdo con la ideología dominante de sus creadores. Según Gabriela Arriagada-Bruneau, Claudia López y Marcelo Mendoza:

> El sesgo implica un error; en otras palabras, una desviación entre las observaciones y la realidad objetiva. Los sesgos también deben tener un componente sistemático; son errores sistemáticos y no meras incidencias aleatorias. Además, cuando hablamos de sesgo, también aludimos a una tendencia injustificada o irrelevante a favor o en contra de ciertas ideas o entidades frente a otras (2026, p. 48).

Lo anterior permite aclarar que, así como sucede en la vida real, en la IA los sesgos son ineludibles; por ejemplo, aquellos de clase, género y racialización. Así, los discursos que son generados por la IA "no provienen únicamente de una conciencia neutra, o desde una racionalidad del algoritmo totalmente pura. Se trata de un reflejo del propio ser social [...] que, desde luego, es desigual, patriarcal y colonial" (Sánchez y Guerrero, 2026, p. 183). En ese sentido, al ser alimentada por seres humanos, en diferentes condiciones materiales, en diversas situaciones psicológicas y con múltiples niveles edu-

opera la IA. Además, su uso genera una gran emisión de gases de efecto invernadero por la quema de combustibles fósiles, sin hablar del daño a los ecosistemas y a ciertas poblaciones que causa toda la producción tecnológica que requiere el funcionamiento de centros de datos (UNEP, 2025). Ante ello, es fundamental hablar de una ética social, la cual para Brugger (2005), alude a una ética política, una ética de la vida, que impacta en el ejercicio profesional como la medicina, el periodismo, los editores, la propia escritura y en el uso de

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

cativos, no es posible que la IA sea objetiva, ni mucho menos que carezca de errores, por ello, transmite creencias de la propia cultura que los produce. Así, la IA no sólo reproduce las ideas y las tendencias subjetivas de quienes la alimentan, sino que, al procesarlas y sistematizarlas, es muy fácil que se desvíe de "la realidad" que pretende acumular y representar para producir otra en su lugar. Una realidad sesgada.

Podemos identificar, al menos, tres contextos de origen de los sesgos: 1) social, 2) técnico y 3) cognitivo. Los sociales, "tienen su origen en los contextos sociales, culturales e institucionales en los que se generan y utilizan los datos. Estos sesgos reflejan y perpetúan los prejuicios y las desigualdades existentes en la sociedad" (Arriagada, *et al.*, 2026, p. 48). Es por lo anterior que si la IA utiliza gente con condiciones precarizadas de países como India, Venezuela o Kenia –entre muchos más–, se puede esperar el reflejo de cierta sociedad y una cultura determinada, con todo lo que eso implique. Sobre todo, porque "los sistemas de IA pueden replicar sesgos sociales existentes, pues la mayoría de ellos se construyen usando conjuntos de datos que reflejan el mundo real [...] imperfecto, injusto y discriminatorio" (Freres en Sánchez y Guerrero, 2026, p. 184). Así, el capitalismo digital reflejado en la IA es posible gracias a la explotación de ciertas poblaciones, por ello, muchos sesgos, que de ésta emanan sostienen la desigualdad.

Por su parte, los sesgos técnicos "se derivan de los métodos, procesos y herramientas específicas que se utilizan para desarrollar e implementar sistemas de IA. [...] están relacionados con las limitaciones algorítmicas y con las decisiones técnicas de los desarrolladores, como la selección de datos y el diseño de modelos" (Arriagada *et al.*, 2026, p. 48). Aquí, quienes realizan prácticas de explotación, al preocuparse sólo por sus intereses en avances tecnológicos, utilidades económicas y políticas, regularmente utilizan métodos bajo una lógica capitalista de despojo. De esta forma, según Deborah Raji, "para que se construyan sistemas de IA, a menudo existe un costo humano oculto" (Hao, 2026, p. 94). Lo que implica que los procesos que se siguen para en-

la IA. Donde, además, no puede separarse la ética individual y la social, entre ambas existe un nexo indisoluble, como lo existe también entre la persona y el colectivo, entre uno y el otro. Por ello, para Lévinas (en Lascevena y Fattah, 2023), la existencia comienza cuando nos vemos interpelados por el rostro del otro. Ahí, frente al otro, como realidad diferente, es donde se nos presenta la posibilidad del descubrimiento. Bajo esta postura, la ética está

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

trenarla no siempre son estandarizados, seguros, ni están transparentados, ya que las decisiones técnicas de dirigentes y creadores muchas veces no son consensuadas, declaradas, ni mucho menos, convenientes para el mundo.

Por su parte, a los sesgos cognitivos son:

> Errores humanos de juicio o razonamiento que pueden afectar al funcionamiento y a la implementación de los sistemas de IA. [...] Se producen cuando los sistemas de IA reflejan las limitaciones cognitivas y las heurísticas de sus desarrolladores y usuarios, cuando imitan los patrones de toma de decisiones humanos, o cuando los seres humanos toman decisiones metodológicas sesgadas que contribuyen a la existencia de sesgos técnicos, como el sesgo de selección. (Arriagada *et al.* 2026, p. 48)

Como se puede ver, este tipo de sesgo involucra tanto a creadores como a alimentadores, que es posible dividir en dos: aquellos de quienes ya hemos hablado, que son pagados y muchas veces explotados para ingresar y discriminar contenidos; y aquellos que no somos pagados, pero que alimentamos a la IA de manera gratuita e inadvertida. Por ejemplo, cuando resolvemos una prueba CAPTCHA, supuestamente diseñada para "demostrar" que somos humanos, seleccionando todas las imágenes que incluyen una bicicleta; cuando damos "me gusta" a ciertas imágenes en redes sociales, o cuando no nos despegamos de nuestro móvil ni siquiera mientras dormimos. Como bien advierte Daniel Estulin, "[a]cceder a tu cuerpo las 24 horas del día, los siete días de la semana, genera muchos datos" (2020, p. 17), mismos que resultan ser lo más valioso para la IA y por ello, los almacena –sin consultarnos– en base de macrodatos.

Con respecto a los creadores, este tipo de sesgos pueden ser problemas que desarrolladores producen al no programar correctamente para que la IA aprenda, lea, cree contenidos, o extraiga patrones adecuadamente o

sustentada en las relaciones humanas y amerita la responsabilidad absoluta por el otro. En ese sentido, Lévinas "considera que la ética es esencial para flexionar sobre la particularidad de cada ser, reflejada en el rostro, siendo precisamente esas peculiaridades las que los diferencia de los otros, haciéndolos únicos y elementales. Asimismo, la ética permite ir al encuentro del Otro" (en Lascevena y Fattah, 2023, p. 84). Por ello, para Lévinas, esta ética

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

incluso problemas causados porque programan a la IA para "robar" contenidos sin pagar derechos y sin declarar que lo están haciendo, para al final ser descubiertos y denunciados. Así mismo, se presentan cuando alimentadores o usuarios, seleccionamos de manera incorrecta un contenido y la IA lo reproduce automática e irreflexivamente.

Por todo lo anterior, los sesgos son tan ineludibles que resulta imprescindible reconocerlos como una parte esencial de la producción académica apoyada en IA. Muchos de éstos responden a nuestras propias ideas que consideramos certezas y que, por tanto, no cuestionamos. Sólo generando un (auto)pensamiento crítico y teniendo mucho cuidado es que podremos evitar cometer errores, plagios y confusiones que nos lleven a situaciones complicadas en las que nosotros seremos quienes debamos responder.

3. Complejizar la ética: hacia una comprensión conceptual y política de la inteligencia artificial

Entonces, si queremos utilizar la IA para nuestras investigaciones, "[h]ay que examinar en qué medida estas innovaciones tecnológicas interactúan con nuestras expectativas de autogobierno e igualdad, [pues] [a]ntes que normativo, el desafío al que nos enfrentamos es conceptual" (Innerarity, 2025, p. 24). Es decir, quizá haya que trascender la normativa meramente ética que encontramos en todos los reglamentos, declaraciones de valores, requisitos, exigencias, imperativos, etc., en el uso de la IA, porque, al desconocer la complejidad tecnológica de ésta, pareciese que lo único que nos exigen es ser responsables cuando hay muchas más cosas en juego que no estamos registrando. Por tal motivo, Daniel Innerarity (2025) introduce una aproximación política que lleva más allá el imperativo ético y evidencia el control privado de la información como lo realmente importante.

Así, desplaza la discusión del sesgo ético que demanda la rendición de cuentas en lugar de la discrecionalidad y que exige el respeto de los valores

de la experiencia humana se da en la proximidad con el otro y en su cuidado, donde el "otro" puede ser también el mundo natural con toda su vastedad de vida. Y la proximidad es entendida no sólo como cercanía o aproximación, sino como aquello que hace posible un acto de obligación, de colaboración, de justicia, sólo ahí puede ocurrir el acontecimiento ético. Al respecto, no podemos ser indiferentes al daño ecológico y social que causa el uso de la IA.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

en lugar del abuso de poder, para advertirnos que, además de eso, es preciso hacer el examen crítico y la articulación entre lo fáctico y lo normativo que nos proporcionen un marco teórico capaz de permitirnos comprender aquellas cuestiones de lo tecnológico que rebasan el ámbito de lo humano, por ello, no pueden ser circunscritas a lo ético tal como lo hemos sostenido en el apartado uno. Si ya desde el principio de este artículo señalamos que la IA no tiene moral, por tanto, no conviene detenernos en discusiones de ese orden, ni mucho menos en tratar de juzgarla –como hacen en la película *Le procès du chien* (Dosch, 2024), en la que someten a juicio a un perro que muerde a una mujer–; al mismo tiempo, es urgente entender qué es lo que la rige, para poder intervenir adecuadamente. Lo anterior quiere decir que, en lugar de preocuparnos solamente por la transparencia en su uso y en tratar de disminuir los sesgos en el manejo de la IA (los cuales, cabe señalar, pueden reproducir todo tipo de discursos normativos, patriarcales, conservadores, raciales; discursos xenófogos, violentos, ultraconservadores, capitalistas depredadores), hay que construir condiciones teóricas para poder comprenderla y utilizarla de manera más responsable en nuestras investigaciones. Pues, independientemente de que sea cierto que la responsabilidad reside en el componente humano, en realidad, cuando usamos IA la ética se complejiza en nuestra relación con la tecnología. Dicho de otro modo, así como un sesgo representa un verdadero problema cuando no tenemos ni la advertencia ni los elementos para poder identificarlos y moderarlos, pareciese que en esta época la responsabilidad no es reductible a la mera utilización de un dispositivo, sino al conocimiento tanto de su funcionamiento como de las implicaciones que tiene.

Una vez convenido lo anterior, quizá lo que resulta más importante al momento de mirar el sesgo de la IA es entender que funciona de manera diferente de nuestro cerebro. Pues si la racionalidad algorítmica no es humana, eso significa que carece, todavía, de la posibilidad de ser reflexiva y ese trabajo sigue siendo exclusivamente humano. En consecuencia, podemos sostener

Por otra parte, siguiendo a Lévinas, la ética en la escritura de un artículo académico interpela al autor o autora en cada palabra, entre sus líneas está la posibilidad de un encuentro con quien lea el texto y, por tanto, con el otro. Así, es un acto de colaboración con la comunidad científica el compartir fuentes, datos, metodologías, resultados, y dar cuenta abiertamente del uso de la IA, porque, como explica Gómez-Vargas (2025), la inteligencia artificial no

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

que la información que genera la inteligencia artificial y que es utilizada en nuestras investigaciones está plagada de sesgos, por lo que, el reto no sólo es que las personas autoras declaren de forma transparente el empleo de la IA, o que aprendan a manejarla correctamente, sino que estén advertidas de todas las implicaciones que tiene el uso de ésta en sus investigaciones.

Pues al ser la IA una tecnología relativamente nueva y desconocida, hace falta todavía mucho trabajo para entender sus procesos, métodos, lógicas, así como, sus encrucijadas que nos permitan posicionarnos éticamente frente a los sesgos que produce o reproduce e identificar cómo se puede sostener y cuáles son las condiciones de posibilidad de que una IA pueda ser juzgada como ética. Pues, si es verdad lo que Innerariry advierte acerca de que "[l] a algoritmización requiere pensar muchas categorías socioculturales, como sujeto, acción, responsabilidad, conocimiento o trabajo" (2025, p. 29), también requiere cuestionarnos si pueden ser propias y posibles en una IA. Ese sería otro de los retos, pues pondría sobre la mesa si en verdad podemos confiar en la IA a pesar de su irreflexividad y su amoralidad, o si podemos crear una IA, no solo reflexiva, si no moral (es decir, humana, no artificial); como la del protagonista de la película *Robocop* (Verhoeven, 1987). Que, por cierto, no era totalmente artificial, sino medianamente humana y, por ello, capaz de limitar moralmente sus propios actos ser capaz de autogobernarse. Cabría recordar que la IA tiene una capacidad increíble de aprender de la propia "experiencia" humana que se le comparte, al grado de que ya puede fingir, así como también mentir. Tal como muestra ejemplarmente la película *The Artifice Girl* (Ritch, 2022).

Inconclusiones. La inteligencia artificial como campo rico de investigación

Llevar los valores fundamentales promovidos por la UNESCO (2024), respecto del uso de la IA en revistas académicas para orientar la produc-

puede recaer solo en prohibiciones, sino en abrir posibilidades de acción, reflexión y crítica, delineando otros posibles horizontes de pensamiento y ejecución. Pero, valdría investigar cómo aminorar el propio daño medio ambiental que causa su uso y así propiciar "un uso *virtuoso* de la inteligencia artificial, ulterior a la de una máquina que devenga sustituto de la inteligencia humana. De ahí que se demanda que la intervención de la IA funja como un

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

ción y difusión del conocimiento, amerita –como hemos revisado– varios aspectos. Primero, que autores declaren de manera transparente su uso. Como lo explican Penabad-Camacho, Morera-Castro y Penabad-Camacho, "la inteligencia artificial es una herramienta en expansión por lo que, en vez de bloquear sus posibles aplicaciones en el entorno de publicación científico-académica, se propone transparentar, normalizar y regular su uso" (2024, p. 5). Pero, como hemos argumentado en este artículo, es vital reconocer las implicaciones que tiene su empleo y complejizar la ética en nuestra relación con la tecnología.

Segundo, para aminorar los sesgos presentes en la IA, los procesos editoriales, además de realizar protección de datos de participantes, pueden promover la evaluación por pares basada en el respeto, la retroalimentación constructiva y la colaboración. Fomentando la participación de autores, revisores e integrantes de comités editoriales provenientes de distintas disciplinas, géneros, territorios y contextos culturales, promoviendo además la diversidad lingüística. Tercero, las publicaciones pueden priorizar investigaciones que contribuyan al bienestar social, fomentando el diálogo interdisciplinario, la representación de poblaciones diversas, incluyendo sectores considerados vulnerables e investigaciones que visibilizan problemáticas sociales y ambientales donde los aportes contribuyan a la injerencia social. Como explica Spinak, la introducción de la IA "ha cambiado el escenario de la redacción, creación y producción de trabajos de investigación generando una revolución del mundo editorial académico de manera irreversible" (2023, s/p). Ante esto, las personas que integran comités editoriales de revistas académicas tienen una responsabilidad ética y un compromiso social frente a los desafíos que amerita el uso de la IA.

No cabe duda de que la regulación humana es un reto, por ello, "la experiencia de uso de la IA debe estar acompañada, desde el ejercicio de cada persona, por las mejores prácticas, respeto a los derechos humanos y conocimiento ético, bioético, disciplinar y profesional" (Penabad-Camacho; Morera-Castro y Penabad-Camacho, 2024, p. 6). Además de tecnológico,

complemento –aditivo– de la inteligencia humana" (p. 206). Por ello, se requiere un manejo crítico y responsable de la inteligencia artificial. Aunque, el mundo de la IA no es neutro, en éste operan relaciones de poder "puesto que dichos dispositivos pertenecen a empresas que poseen principalmente intereses comerciales" (Gómez-Vargas, 2025, p. 200). Es decir, la IA responde al giro ideológico del gobierno actual en Estados Unidos de América y *Silicon Va-*

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

científico, algorítmico, matemático, político, económico, etc. Pero ¿cómo lograrlo tomando en cuenta los estándares de evaluación académica que además de ignorar muchas cuestiones relacionadas con la IA, priorizan la premura, la producción y la publicación de artículos, sostenidos en un tiempo acelerado, en revistas de alto impacto? y ¿cómo advertirnos sobre los sesgos que puedan (podamos) tener los propios autores?

Así que, más que establecer imperativos que no sabemos si la IA puede cumplir, porque no es humana, habría que aprender cómo funciona y cuáles son sus cualidades académicas, para saber en qué fases de la investigación utilizarla y en qué otras no. Pues si bien en muchas áreas la IA ya puede sustituirnos, es verdad que en la academia todavía no lo puede hacer y en la medida en que "no hemos encontrado el equilibro adecuado entre el control humano y los beneficios de la automatización" (Innerarity, 2025, p. 30), podemos fácilmente deslizarnos en un vértigo que nos dé la impresión de que estamos utilizando la IA de forma adecuada y ética, mientras que en realidad estamos abusando y engañando a los demás sobre nuestro trabajo. Pues hay que recordar que así como hace mucho tiempo, antes de la IA, había investigadores que publicaban libros con la suficiente profundidad y rigor teóricos, sin la necesidad de hacer uso de herramientas que les permitieran ahorrarse el trabajo, hoy hay millones de personas cuyo ímpetu de competencia, éxito y reconocimiento les ha llevado a publicar múltiples artículos y libros sin haber investigado ni trabajado lo suficiente y hay muchos más que, por la premura de nuestra época, están dispuestos a leerlos sin verificar sus trayectorias, sólo por la cantidad de "me gusta" que tienen en sus redes sociales y por el capital social del que se han hecho en esta época.

Las condiciones de la academia ya no son las mismas del siglo pasado. Ante ello, una población que lleva décadas acostumbrándose a obtener el conocimiento de una IA y del *mainstream*, puede creer que ya no necesita docentes, leer textos físicos, tener encuentros y discusiones cara a cara dentro de las aulas, construir el conocimiento de forma conjunta y reflexiva, o

lley, así como a guerras por la hegemonía cibernética, comercial y bélica del orden mundial donde los "*Big Tech* son el prestador de servicios de la soberanía del incipiente imperio [...] afincado en la misoginia, el desprecio por la vida y el racismo" (Emmelhainz, 2026, s/p), con varios fines, entre éstos, modelos de censura, manipulación política, campañas de influencia sobre la opinión pública, extractivismo intelectual. Así, el uso de la IA exige reconocer que

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

escribir un artículo científico, una nota periodística, un ensayo, un poema, un libro, o una tesis sin ayuda de la IA.

La investigación en el campo de la IA ha experimentado cambios significativos en cuanto a su enfoque, metodología y marcos teóricos desde sus inicios. Se trataba de "alcanzar una IA 'fuerte', es decir, una forma de inteligencia artificial que fuera prácticamente indistinguible de la inteligencia humana. Sin embargo, [...] el punto en el que nos encontramos podría describirse con mayor precisión como la creación de una serie de tecnologías que proporcionan una ilusión de inteligencia" (Natale, 2021, p. 3).

Finalmente, desde que se inauguró la línea de investigación sobre la IA, uno de los problemas principales fue conocido como "el problema del observador", concepto referido a "cómo reaccionan los seres humanos al presenciar cómo las máquinas muestran inteligencia" (Natale, 2021, p. 43), mismo que ya indicaba que ninguna inteligencia puede ser reconocida por sí misma, sino en función de una otredad. Por esa razón es que, más de medio siglo después, seguimos intentando no sólo comprender la manera en que las computadoras expresan su "inteligencia", sino intentando ayudarles a ser inteligentes; aunque no hemos logrado obtener una que rebase la ilusión e iguale a los humanos en todos los ámbitos, porque, por más sesgos humanos que le introduzcamos, la singularidad sigue siendo inimitable. Pues es verdad que, porque la IA responde y reacciona como si fuera una persona, hay mucha gente que cree que las operaciones de ésta ya son equivalentes a las de la inteligencia humana, pero lo humano es algo que va más allá, una condición singular e irrepetible, ampliamente mucho más compleja a que una máquina.

toda decisión sobre su empleo implica responsabilidades humanas. De ahí la importancia de desarrollar desde la escritura y la investigación académica una mirada crítica, cuestionar el extractivismo, fomentar una ética compleja preocupada por la colectividad, y sobre todo, una praxis por el cuidado de la vida en general, del mundo común, sólo así se podrá contribuir al fortalecimiento, y no al deterioro, de la vida compartida.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

Fuentes consultadas

Arriagada, G., López, C. y Mendoza, M. (2026). *Ethics in Artificial Intelligence and Information Technologies*. CRC Press.

Born, R. (2018). *Artificial Intelligence: The Case Against*. Routledge Library Editions, Taylor and Francis.

Brugger, W. (2005). Ética. En *Diccionarios de Filosofía*. pp. 222-226. Herder.

Diccionario de la Lengua Española. (2024). *Ético, ca*. Disponible en: https://dle.rae.es/%C3%A9tico#H3yay0R

Dosch, L. (Dir.). (2024). *Le Procès du Chien*. Bande à Part Films y Atelier de Production.

Emmelhainz, I. (2026-07-05). México y la IA. En *La Jornada*. Disponible en: https://www.jornada.com.mx/noticia/2026/07/05/opinion/mexico-y-la-ia

Epstein, R. (1997). *The Case of the Killer Robot: Stories about the Professional, Ethical, and Societal Dimensions of Computing*. Wiley.

Estulin, D. (2020). *Metapolítica*. Penguin Random House Grupo Editorial, S.A.U.

García-Clarck, R. (2025). *La ética contemporánea y el problema de la violencia*. Torres Asociados.

Gómez-Vargas, C. (2025). ¿Schopenhauer como educador o la inteligencia artificial como educadora? La filosofía de la educación en el contexto de las inteligencias artificiales. En *Revista Latinoamericana de Filosofía de la Educación*. Vol. 12. Núm. 24. pp. 194-210. DOI: https://doi.org/10.63314/PAAV1231

Hao, K. (2026). *El Imperio de La IA*. Ariel.

Heron, M. y Belford, P. (2025). *A Case Study for Computer Ethics in Context: the Scandal in Academia*. CRC Press.

Innerarity, D. (2025). *Una teoría crítica de la inteligencia artificial*. Ensayo.

Kubrick, S. [dir]. (1968). 9000 en *2001: A Space Odyssey*. Metro Golwyn-Mayer (MGM).

Lascevena, M. y Fattah, J. (2023). La ética levinasiana: un encuentro de alteridad en los tiempos contemporáneos. En *Encuentros. Revista*

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

de Ciencias Humanas, Teoría Social y Pensamiento Crítico. Núm. 19. pp. 83-93. DOI: https://doi.org/10.5281/zenodo.8270641

Mc'Quarrie, C. (2025). *Mission: Impossible - The Final Reckoning*. Paramount Pictures, Skydance Media y TC Productions.

Murray, D. (2020). *La masa enfurecida. Cómo las políticas de identidad llevaron al mundo a la locura*. Titivillus

Natale, S. (2021). *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*. Oxford University Press.

Penabad-Camacho, L. Morera-Castro, M y Penabad-Camacho, M. (2024). Guía para uso y reporte de inteligencia artificial en revistas científico-académicas. En *Revista Electrónica Educare*. Núm. 28. pp. 1-41. Disponible en: https://www.revistas.una.ac.cr/index.php/EDUCARE/article/view/19830/31511

Quintavalla, A. y Temperman, J. (2023). *Artificial Intelligence and Human Rights*. Oxford University Press.

Ramírez, L. (2025). *El diablo está entre nosotros: los agentes del caos en el nuevo orden mundial*. La esfera de los libros.

Ritch, F. (Dir.). (2022). *The Artifice Girl*. Paper Street Pictures, Blood Oath y Last Resort Ideas.

Sánchez, S. y Guerrero, R. (2026). El algoritmo también tiene género. La acumulación originaria de datos y el patriarcado digital en América Latina. En *Tramas y Redes. Revista del Consejo Latinoamericano de Ciencias Sociales*. Núm. 10. pp. 181-196.

Schapire, A. (2021). *La traición progresista*. Península.

Scott, R. (Dir.). (1979). *Alien*. Brandywine Productions y Scott Free Productions.

Spinak, E. (2023-08-30). Inteligencia artificial y comunicación de investigaciones. En *ScieELO en Perspectiva*. Disponible en: https://blog.scielo.org/es/2023/08/30/inteligencia-artificial-y-comunicacion-de-investigaciones/

Stryker, C. y Kavlakoglu, E. (2024). ¿Qué es la IA? En *IBM*. Disponible en: https://www.ibm.com/mx-es/think/topics/artificial-intelligence

Programa de las Naciones Unidas para el Medio Ambiente (UNEP). (2025-11-13). *La IA plantea problemas ambientales. Esto es lo que el mundo puede hacer al respecto*. Disponible en: ONU

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

https://www.unep.org/es/noticias-y-reportajes/reportajes/la-ia-plantea-problemas-ambientales-esto-es-lo-que-el-mundo-puede

UNESCO. (2024). Ética de la inteligencia artificial. En *Inteligencia artificial y tecnologías emergentes*. Disponible en: https://www.unesco.org/es/artificial-intelligence/recommendation-ethics

Verhoeven, P. (Dir.). (1987). *Robocop*. Orion Pictures y Hawkeye Entertainment.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350