

DILEMAS ÉTICOS EN EL ESCENARIO DE LA INTELIGENCIA ARTIFICIAL

ETHICAL DILEMMAS IN THE ARTIFICIAL INTELLIGENCE SCENARIO

DILEMAS ÉTICOS NO CENÁRIO DE INTELIGÊNCIA ARTIFICIAL

Mario González Arencibia¹
Dagmaris Martínez Cardero²

Resumen

La importancia del tratamiento del tema asociado a los usos maliciosos de los resultados de la inteligencia artificial, está marcando el debate actual de los estudios que se realizan desde las instituciones académicas, por sus implicaciones negativas en el escenario social. En este sentido el propósito de este artículo, consiste en repensar los dilemas éticos que están ocurriendo en el marco de la sociedad digital. Con este propósito la pregunta es la siguiente: ¿Cómo eliminar los riesgos de un uso negativo del desarrollo tecnológico? El estudio permite, identificar los puntos críticos que están afectando la sociedad, con la idea de promover el debate en torno, a sus soluciones. La conclusión central es que la epistemología para la solución de los problemas sociales que está generando el inadecuado uso de la tecnología, está por encima de la naturaleza técnica, y exige reevaluar este fenómeno desde un enfoque multidisciplinario, con énfasis en el componente cultural.

Palabras claves: inteligencia artificial; dilemas éticos; autonomía de la tecnología; moral

Doi: <https://doi.org/10.15359/eyes.25-57.5>

Recibido: 12-03-2020. Reenvíos: 23-03-2020, 13-04-2020, 24-04-2020. Aceptado: 27-04-2020. Publicado: 12-05-2020.

¹ Profesor Titular en Universidad de las Ciencias Informáticas, Cuba. Economista, Dr. En Ciencias Económicas. Correo electrónico: mgarencibia@uci.cu ORCID: <https://orcid.org/0000-0001-9947-7762>

² Profesora auxiliar en Universidad de las Ciencias Informáticas, Cuba. Licenciada en Psicología. Correo electrónico: dagmarism@uci.cu ORCID: <https://orcid.org/0000-0001-9526-1112>

Mario González Arencibia y Dagmaris Martínez Cardero



Revista Economía y Sociedad by [Universidad Nacional](http://www.una.ac.cr) is licensed under a [Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Creado a partir de la obra en <http://www.revistas.una.ac.cr/index.php/economia>

Abstract

The importance of analyzing the malicious uses of artificial intelligence results is leading the current debate of studies conducted by academic institutions because of their negative implications in the social field. The purpose of this article is to rethink the ethical dilemmas under a digital society framework. Consequently, the important question here is how to eliminate the risks of the negative uses of technological developments. This study allows to identify the key points affecting the society in order to promote debate around the solutions. As the main conclusion, using epistemology to solve the social problems generated by the inadequate use of technology is above the technical nature and requires re-evaluating this phenomenon from a multidisciplinary approach, mainly focused on the cultural component.

Keywords: artificial intelligence; ethical dilemmas; technology autonomy; morality

Resumo

A importância do tratamento do tema associado aos usos maliciosos dos resultados da inteligência artificial marca o debate atual dos estudos realizados a partir das instituições acadêmicas devido a suas implicações negativas na cena social. Nesse sentido, o objetivo deste artigo é repensar os dilemas éticos que atualmente ocorrem no quadro da sociedade digital. Com este propósito, a questão é: como eliminar os riscos de uma utilização negativa do desenvolvimento tecnológico? O estudo permite identificar os pontos críticos que estão afetando a sociedade, com a ideia de promover o debate em torno das suas soluções. A conclusão central é que a epistemologia para a solução dos problemas sociais que o uso inadequado da tecnologia está gerando está acima da natureza técnica e exige uma reavaliação deste fenômeno a partir de uma abordagem multidisciplinar, com ênfase no componente cultural.

Palavras-chave: inteligência artificial; dilemas éticos; autonomia da tecnologia; moral



Introducción

Las preocupaciones por los dilemas éticos en el marco de la inteligencia artificial (IA), constituye un aspecto ampliamente debatido en las condiciones actuales (Mourelle, [2019](#)) y (Carreño, [2019](#)). Según Shaw ([2019](#)), estas inquietudes se centran en el empleo de herramientas de Big Data, y en los riesgos políticos que se derivan de su uso mal intencionado, al ser empleadas para el fomento de la manipulación y la distorsión de procesos sociales.

El hecho es que la aplicación ilegítima de la IA, está cuestionando los preceptos éticos en el campo de la ideología y la política, perturbando la legitimidad de los procesos democráticos. Ello se expresa particularmente en la manipulación de campañas electorales y sistemas políticos que no respondan a las aspiraciones de la hegemonía del capital.

Particular importancia tiene en este escenario, el interés por parte de comités académicos, organizaciones gubernamentales, y no gubernamentales, por el tema³, debido al carácter cada vez más social que generan. Sin embargo, la incertidumbre y la principal amenaza está en cómo su naturaleza privada desde los centros de poder mundial, emplea estos avances para ocasionar conflictos que generan distorsiones éticas (Samaniego, [2018a](#), [2018b](#)) y (Knaus, [2017](#)).

El efecto es que los problemas que ha estado generando el inadecuado uso de las tecnologías, muestran la falta de preparación para el examen de las consecuencias negativas de los errores ocasionados por los diseñadores de programas informáticos. La notoriedad del tema, ha conducido a que la IEEE⁴, –entidad de referencia mundial– esté avanzando en la creación de equipos de trabajo para enfrentar las contradicciones sociales que están emergiendo en el contexto de la IA (Torrero, [2019](#)).

Entre las principales amenazas que, desde el enfoque ético, se debate actualmente, en el entorno digital es el referido a la privacidad, en un contexto donde la tendencia ha sido a que la intimidad se ve cada vez más erosionada por la invasión a que ha sido sometida. Este fenómeno ha estado acompañado del empleo de sistemas de información, que la naturaleza humana le ha impregnado el poder de que puedan tomar de manera autónoma decisiones ejecutables en ocasiones riesgosas, presentándose un dilema moral (Shaw, [2019](#)).

La cuestión está en que las características positivas y negativas inherentes a los seres humanos se están reflejando en los resultados de la revolución científica derivada de los avances de la informática. Se trata de decisiones morales en el entorno digital de las personas, sobre cómo las emplean y la forma en que se relacionan entre sí (Ramos, [2018](#)). La otra argumentación es que

³ Ismail, N. ([2019](#))

⁴ Institute of Electrical and Electronics Engineers



una visión para el desarrollo e implementación de la IA exige tener claridad y consenso acerca de los valores a considerar, para su aplicación en el nuevo escenario tecnológico (Berberich, & Diepold, [2018](#)).

Una interrogante clave en este contexto es la siguiente: ¿Cómo eliminar los riesgos de un uso malicioso del desarrollo tecnológico en el ámbito de la inteligencia artificial? La prueba es que, desde la concepción de la ética y la moral, trabajar en la conciencia de sus desarrolladores es esencial. Atendiendo a esta inquietud, el propósito de esta investigación consiste en reflexionar sobre los dilemas éticos de los desarrollos tecnológicos en el marco de la sociedad digital.

Dilemas éticos

En general, la ética es una disciplina filosófica, que tiene un carácter normativo y práctico, sobre cómo actuar en el entorno social, donde se inserta el individuo, aplica los principios éticos para conciliar la moral⁵, los deseos y las capacidades de las personas para actuar de manera coherente. En su dimensión consecuencialista, una persona es ética si en cada acción que realiza mide los efectos que pueden producir sus actos, seleccionando aquellos que tienen resultados a favor de lo moral.

Desde la ética deontológica, una persona es ética solo si respeta los deberes, las obligaciones y los derechos afines con determinadas situaciones. Las personas con ética deontológica (conocida como ética de la obligación o del deber) actúan en correspondencia con las normas establecidas. Por otro lado, la ética de la virtud, plantea que una persona es ética, solo si piensa y actúa coherentemente desde los valores morales, cuya actuación debe nacer de manera intrínseca como factor de credibilidad (Cointe, Bonnet, & Boissier, [2016](#)) y (Boissier, Bonnet, Cointe, De Swarte, & Vallée, [2018](#)).

Los principios son el eje de referencia sobre los cuales se enfrentan los dilemas éticos, por lo que en cualquier análisis asociado al tema es importante partir de ellos. De acuerdo a un estudio sobre: “27 Principios fundamentales de la bioética”, toda aplicación de la ciencia y la tecnología debe partir del respeto y la protección a las personas, considerarlas con autonomía, evitando hacerles daño, y maximizando los beneficios, reduciendo los perjuicios, lo cual se traduce en la importancia de considerar las bases de la beneficencia y la no maleficencia, es decir, que, si se hace el bien, no es lógico causar un daño. (Curso OPE Andalucía, [2019](#)).

⁵ La moral consiste en un conjunto de reglas que describen el cumplimiento de un comportamiento dado, según las costumbres de un grupo o de una sola persona. Estas reglas asocian un valor o un desvalor a algunas combinaciones de acciones y contextos. Pueden ser específicos o universales, es decir, relacionados o no con un período, un lugar, un pueblo, una comunidad, etc. (Cointe *et al.*, [2016](#)).



Por consiguiente, generar el bien es una obligación de la ética, por lo que este hecho condiciona la necesidad de normas, que exigen que los riesgos de cualquier investigación sean evitables, y que esta, esté adecuadamente pensada, y que los investigadores sean competentes para llevarlas a cabo, garantizando el bienestar de sus participantes.

De estas afirmaciones, se concluye que un dilema ético irrumpe cuando el profesional afronta dos o más alternativas de actuación posibles, en un contexto en que están presentes valores y principios éticos; que, de violar estas premisas, el profesional involucrado está en una situación en la que está penado a cometer una arbitrariedad: sin importar lo que haga, hará algo “desacertado” o faltará a un deber u obligación, generando un problema ético (Kvalnes, [2019](#), p. 11) y (Castillero, [2016](#)).

De lo anterior, se deriva que el dilema ético es de naturaleza moral, lo cual exige garantizar el respeto a los principios fundamentales, que son los que facilitan la toma de decisiones, debido a que exige del análisis, de la reflexión y del pensar lo que es mejor para la persona. Así, un dilema ético genera un problema ético, al verse violados valores y principios, que generan una situación en la que los valores morales unido a los deberes se encuentran en conflicto, de forma que cualquier respuesta posible es moralmente intolerable.

En consecuencia, un dilema ético se transforma en un problema ético, en tanto, contribuye a alterar las normas y el sistema de valores sociales, y estos al mismo tiempo, transitan y se convierten en un problema social, cuando pasan a producir afectaciones a una parte importante de la sociedad, y esta, asume una posición crítica ante el mismo demandando de una solución.

Inteligencia artificial como concepto

Según colectivo, Sara Mattingly-Jordan, Rosalie Day, Bob Donaldson, Phillip Gray, L. Maria Ingram, ([2019](#), p. 8), en su glosario de términos, *Ethically Aligned Design*, la inteligencia artificial se refiere a: “La capacidad de las computadoras u otras máquinas para simular o mostrar un comportamiento inteligente”, es decir es un sistema de símbolos, que simula el comportamiento humano, desde programas informáticos.

Dicho concepto es contrario, al supuesto de inteligencia humana que expresa según Macías B., Fernández, A., Méndez, C., Poch, J., & Sevillano ([2015](#)), la capacidad mental que permite planificar, razonar, pensar de modo abstracto, comprender y dar solución a problemas complejos, empleando la experiencia y aprendiendo de forma racional, como una actividad que permite transformar el entorno social.

El concepto de inteligencia humana planteado, es contrario al de inteligencia artificial, en la medida en que, se concibe a la inteligencia en sus múltiples dimensiones, que van en contra de



considerar la tecnología como un fenómeno que puede tomar decisiones razonadas. La cuestión sería, ¿Qué opinar del criterio de Berberich, N. & K. Diepold, (2018), en su artículo *The Virtuous Machine. Old Ethics for New Technology?*, quienes plantean que la ética podría ser un factor inherente a las máquinas, cuestión que está siendo arraigado en el debate actual sobre los aspectos éticos de la inteligencia artificial.

Es importante acotar a partir de lo anterior, varias cuestiones: primero, la inteligencia artificial como concepción, está fundamentada en la resolución de problemas con la finalidad de imitar el comportamiento humano desde lo cognitivo, y no así, desde lo emocional. La definición del término tecnologías inteligentes, ha ocasionado conflictos en la aplicación de estas, debido a que se le intenta dar funciones inherentes a la capacidad de manejar las emociones como lo hace el ser humano, entrando en contradicción entre el deber moral, y las circunstancias donde se implementa.

Ello cuestiona el argumento, de si efectivamente, se debería emplear el calificativo de inteligencia, a los actuales usos de la tecnología digital. Identificar dificultades es el primer paso de su superación, esto indica la necesidad del tratamiento de este debate considerando un enfoque ético, que coloque en el centro los aspectos asociados a la inteligencia emocional. Este examen deberá profundizar, en las potencialidades que tienen las denominadas tecnologías inteligentes de generar empatía y habilidades sociales para afrontar las tareas de la cotidianidad.

Segundo, los sistemas de inteligencia artificial en sí no son autónomos, son guiados por personas y su empleo depende de su naturaleza humana. Es inconsistente en este sentido, que estos sistemas llegarán a controlar el mundo. La malevolencia está en las inadecuadas conductas de quienes organizan estos procesos. Es decir, son los diseñadores de algoritmos con objetivos inadmisibles quienes manipulan lo que se denomina armas autónomas y maniobran con objetivos políticos elecciones presidenciales.⁶ Por lo que es significativo, enfrentar grupos de poder vinculados al mal empleo de la tecnología (López de Mántaras, 2019).

Tercero, es insuficiente el concepto de autonomía de la tecnología, esencialmente este término remite a la capacidad que tienen las personas de tomar decisiones y resolver problemas de manera independiente y racional, desde las herramientas de la inteligencia emocional. Simular datos a través de comportamientos humanos imitados por la tecnología, no es equivalente a la capacidad de pensar desde la teoría y la práctica experimentada por el hombre. El calificativo más cercano que se le podría añadir al tema de tecnología con autonomía es el de autómeta, lo que

⁶ Los resultados de las elecciones presidenciales de EE. UU. y el referéndum del Reino Unido, sobre la adhesión a la Unión Europea se vieron influenciados por el empleo de la IA, para detectar y dirigirse a votantes indecisos (Bryson & Winfield, 2017, p. 1).



se refiera a un proceso que desde la tecnología imita la figura y los movimientos de un ser animado.

Cuarto, los sistemas de inteligencia artificial como tecnologías en sí son amorales, no tienen capacidad para distinguir entre el bien y el mal, asumen la naturaleza de sus creadores (Bossmann, 2016). No será la tecnología en sí, la que acabará con las miserias humanas, tampoco dominarán el mundo, y mucho menos desde una visión catastrofista acabarán con el planeta, Internet de las Cosas no es quien pondrá en riesgo la privacidad del individuo, las armas y los denominados sistemas autónomos e inteligentes, son programados por el ser humano.

Quinto, una evaluación del cuestionamiento que se le realiza a los dilemas éticos que se generan desde la inteligencia artificial, sugiere reflexionar sobre el contenido mismo de esta categoría en su expresión filosófica, lo que hace pensar que las contradicciones sociales que se generan desde su inadecuado uso, no están en la tecnología en sí, sino en el actor que genera los procesos de algoritmización de los componentes humanos. Ello obliga a repensar desde la ética la relación sujeto-objeto que produce estos resultados.

Usos maliciosos de la inteligencia artificial que se transforman en dilemas éticos

El estudio de un colectivo de autores de diversas universidades norteamericanas⁷, titulado: *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*, realizado por Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D., (2018); destaca aspectos positivos que se podrían transformar y emplear con fines nocivos y, difíciles de impedir, lo cual se puede observar en cuatro campos: identidad, invasión de los entornos, intimidad y seguridad.

Invasión de los entornos

- La IA tiene nuevas vulnerabilidades sin resolver, como es el caso de ataques de envenenamiento de información (introducción de datos de formación que hacen que un proceso de aprendizaje cometa errores), ejemplos contradictorios (entradas diseñadas para ser clasificadas erróneamente, por los sistemas de aprendizaje automático), y la explotación de fallas en el diseño de los objetivos autónomos.
- La cotidianidad en el entorno físico, exige de comunicación frecuente, así como tomar decisiones sobre la base del comportamiento social que se aprecia, con la automatización

⁷ Miles Brundage, Shahar Avin, Jack Clark, Helen Toner, Peter Eckersley, y Ben Garfinkel, Allan Dafoe, Paul Scharre, Thomas Zeitoff, Bobby Filar, Hyrum Anderson, Heather Roff, Gregory C. Allen, Jacob Steinhardt, Carrick Flynn, Seán Ó h Éigeartaigh, Simon Beard, Haydn Belfield, Sebastian Farquhar, Clare Lyle, Rebecca Crotoft, Owain Evans, Michael Page, Joanna Bryson, Roman Yampolskiy, Dario Amodei.



de estas tareas, los sistemas de IA permiten a los actores mantener su anonimato y experimentar un mayor grado de distancia psicológica de las personas a las que impactan. Por ejemplo, alguien que emplea un sistema de armas autónomo para cometer un asesinato, en lugar de usar una pistola, impide tanto la necesidad de estar presente en la escena, como la de mirar a su víctima.

- La capacidad de producir nuevos peligros, debido a la posibilidad que tienen estas tecnologías de rastrear las fragilidades con facilidad y tomar el control de vehículos, dispositivos, robots o cualquier infraestructura digital, es trascendente. Y, de otra parte, hace pensar en cómo estas debilidades podrían recrudecerse, al proporcionar la posibilidad de empleo de *software* malicioso, a grupo de personas, o a individuos aislados, que antes no accedían a esta tecnología.

Una de sus características esenciales es que son sistemas eficientes y escalables,⁸ es decir cuanto más se empleen, más conocimiento incorporan a su modo de actuación. Existen *softwares* maliciosos y *ransomwares* inteligentes que, con estos aprendizajes, se propagan, coordinando ciberataques globales con análisis de datos avanzados para personalizar estos, con las consecuencias implícitas de este peligro.

Intimidad

- Los programas que buscan fragilidades en dispositivos y redes pueden aplicarse para agredir y aprovechar estas vulnerabilidades, por ejemplo, la capacidad de producción de imágenes puede dar lugar a reemplazos de identidad o la publicación de contenido falso.
- La automatización de tareas imposibilita que factores psicológicos como la empatía entren en juego a la hora de tomar decisiones, a lo que se añade que en este contexto se refuerza el anonimato.
- La capacidad de generar imágenes, texto y audio sintéticos podría utilizarse para hacerse pasar por otros en línea, o para influir en la opinión pública mediante la distribución de contenido generado por inteligencia artificial, a través de canales de medios sociales. Investigadores de ZeroFox demostraron que un sistema de *phishing* de automatizado podría reproducir *tweets* en la plataforma de medios sociales Twitter basados en los

⁸ La eficiencia de un sistema de IA consiste en que una vez entrenado y desplegado, puede ejecutar una tarea de forma más rápida que un humano. La escalabilidad está determinada por las posibilidades que posee de completar mayor cantidad de acciones, aumentar la potencia de cálculo a la que tiene acceso o hacer copias que le permitirían desarrollar más actividades. Por ejemplo, un sistema de reconocimiento facial típico es eficiente y escalable; una vez desarrollado y entrenado, puede ser aplicado a muchas cámaras diferentes por mucho menos que el costo de contratar analistas humanos para hacer el trabajo equivalente (Brundage *et al.*, 2018).



intereses de un usuario, logrando un alto índice de clics en un enlace que podría ser malicioso.

- Con el manejo de datos se pueden hacer predicciones del comportamiento y preferencias de las personas, lo que, al mismo tiempo, aumentan las posibilidades de persecución o control (Bryson, & Winfield, [2017](#)).

Seguridad

En relación con la seguridad, el uso malicioso de la inteligencia artificial amenaza:

- a) La seguridad digital: por ejemplo, a través de máquinas de formación de delincuentes para piratear o manipular socialmente a las víctimas en niveles humanos o sobrehumanos de rendimiento.
- b) La seguridad física: por ejemplo, agentes no estatales que convierten en armas a los aviones teledirigidos de consumo.
- c) La seguridad política: por ejemplo, mediante la eliminación de la vigilancia, la elaboración de perfiles y la represión de la privacidad, o a través de campañas de desinformación automatizadas y selectivas.

Identidad

- El examen facial preciso se puede implementar a sistemas de armas autónomas, lo que es significativo si se valoran sus potencialidades, por el hecho de que la IA en reconocimiento de imágenes ha pasado de conceptualizar correctamente alrededor del 70 % de las fotografías a una identificación casi perfecta del 98 %, superior al 95 % de precisión humana (Brundage, *et al.*, [2018](#)).

Por consiguiente, el estudio de Brundage, *et al.*, ([2018](#)), sobre los usos maliciosos de la inteligencia artificial –en los campos antes presentados–, sintetiza potencialidades que podrían ser perjudiciales para la sociedad, sobre las que es importante pensar en cualquier alternativa, ante los daños que puedan causar. Ello sugiere repensar de manera permanente los escenarios de la seguridad física, y la ciberseguridad, con la finalidad de proteger la soberanía nacional de los Estados, evitando la posibilidad de que aparezcan amenazas que afecten a países y regiones.

Sobre los riesgos de los usos maliciosos de la tecnología, es notorio el caso del *Deepfakes*,⁹ falsificación a la perfección, cuyo contenido no ético consiste en manejar procedimientos de

⁹ Se elaboran videos de líderes estatales, como el del caso de Obama, que parecen hacer comentarios que nunca se hicieron en realidad, con las consecuencias implícitas de este hecho.



aprendizaje profundo para inventar situaciones falsas, la cuales pueden ser desarrolladas en audiovisuales y en textos. Por ejemplo, se emplea un escrito fundamentándose en predicciones del acontecimiento que se podría producir después (Paniagua, [2019](#)).

Las redes neuronales, y las técnicas como el *deep learning machine*, y el *learning* avanzado, posibilitan que los sistemas artificiales descifren patrones y pongan al descubierto, que pueden crear vulnerabilidades éticas. Es en este escenario, en el que se pueden identificar los riesgos asociados a la construcción y a la ejecución de los sesgos humanos que se concretan en sistemas algorítmicos, los cuales, empleados indebidamente desde la inteligencia artificial, causan efectos perjudiciales para la sociedad.

En esta situación se ponen en evidencia las *Deep Fake* (falsificaciones profundas) mediante la manipulación de imágenes convertidas en videos, las cuales han traído peligros para la soberanía nacional de países y regiones enteras, creando múltiples desórdenes informativos, al punto de que estos impactos se están manipulando ideológicamente, para el replanteo de pensamientos como el fin de la verdad, cuestión que está dirigido a atacar creencias, convicciones, sentimientos, símbolos, todo ello basado en una gran capacidad económica y tecnológica. El propósito es someter ideológicamente al oponente y la construcción de un campo de batalla a distancia, con herramientas de IA de gran alcance y precisión, que reducen la conciencia de defensa.

De estos ejemplos, se podría concluir parcialmente, que los fallos que producen los desarrollos tecnológicos de la IA, inadecuadamente programados, colocan en riesgo la dignidad y la vida humana, la libertad de expresión, el derecho a un empleo e incluso la salud. Tómense de ejemplo, el caso de Arkansas (EE. UU.), donde un error de codificación de un algoritmo ocasionó un descenso injustificado de la cobertura de salud, para pacientes con parálisis cerebral, asignándoseles una menor cantidad de horas de cuidados intensivos, de las que en realidad debían planificarse (Lecher, [2018](#)).

Jonnathan Penn, especialista del Centro Berkman Klein de Harvard University, en el Seminario HUMAINT: “Impacto ético, legal, social y económico de la inteligencia artificial”, organizado por el Centro Común de Investigación de la Comisión Europea, en febrero de 2019, en Sevilla, presentó evidencias de quebrantamiento de la intimidad en Facebook, al plantear como ejemplo, que un examen de 1000 aplicaciones Android ejecutado por la organización *Privacy International*, reveló que el 61 % de esa pesquisa se reproduce instantáneamente, para cualquier usuario de las redes sociales, independientemente del interés de este.

También, Google divulga datos sensibles de sujetos con dificultades médicas asociadas a cáncer, impotencia, adictos a la droga. Estos hechos han sido denunciados en el ámbito de la región europea, a solicitud de University College London, el navegador web privado Brave y Open Rights Group, organizaciones que acusan a esta empresa de transgredir el Reglamento General de



Protección de Datos (GDPR), por favorecer el éxodo masivo de información íntima (Paniagua, [2019](#)).

A todos estos análisis se unen concepciones de obtención de supremacía mediante el uso del dominio monopólico de la tecnología que plantean: “Quien lidere la Inteligencia Artificial gobernará el mundo” (Gigova R., [2017](#)). Esta frase podría tener una connotación optimista o catastrófica. La primera dirigida a eliminar los problemas más acuciantes de la humanidad (hambre, pobreza, medio ambiente, desigualdades sociales), la segunda, fundamentada en acciones no éticas, como los ciberataques a países. Lamentablemente, líderes mundiales destacan las posibilidades que tiene la inteligencia artificial para amenazar intereses de empresas, ciudades, países y continentes, como un ejercicio de dominio para desequilibrar la balanza de poder mundial (Samaniego, [2018b](#)).

Los desarrollos de la tecnología, actualmente exigen marcos regulatorios en relación con enfrentar sus usos maliciosos, por consiguiente, la integración de estos a la carrera armamentista genera preguntas que cuestionan este uso: ¿Es correcto este empleo de la tecnología?, ¿Es moral?¹⁰ ¿Cuánta inseguridad genera? Este cuestionamiento se conecta con la idea de desarrollar la ética en los avances de la IA. (Scherer, [2016](#)). De esto se deriva, la urgencia de soluciones que es necesario encontrar frente a los dilemas éticos planteados.

Debate en torno a las soluciones éticas en el escenario de la IA

¿Qué plantea el debate en torno a la ética en el contexto planteado? En lo que respecta con la solución de los problemas sociales generados por los usos maliciosos examinados, existen diferentes criterios:

Un grupo de autores Laichter, L. ([2019](#)), Yu, H., Shen Z., Miao, C., Leung, C., Lesser, V. & Yang, Q., ([2018](#)), Boissier, O., *et al.*, ([2018](#)), Varangaonkar, A. ([2018](#)), Conitzer, Sinnott, Schaich, Deng & Kramer ([2017](#)), Cointe, *et al.*, ([2016](#)), coloca la atención en relación con generar alternativas desde la propia tecnología, dedicándose a construir variantes éticas desde esta. Es decir, que el rol de corregir los dilemas se lo otorgan al componente artefactual. Esta postura enarbola un exagerado fetichismo tecnológico, a partir de la creencia, de que desde los artefactos se resolverán los problemas generados por los sesgos humanos que generan comportamientos maliciosos. Un análisis crítico de esta visión, indica la necesidad de asumir la tecnología como complemento, y no como centro para resolver los usos negativos de la ciencia.

Otra arista del debate, plantea la pregunta siguiente: ¿Resolver los problemas que genera el inadecuado empleo de la IA, tiene una solución política? Samaniego ([2018a](#)). En este sentido, se

¹⁰ Monasterio Astobiza, A. ([2019](#)).



asume que la solución es política, en un plano en que la seguridad y la prevención, son los ejes que debe promover el debate. En esta dirección, Brundage, *et al.* (2018), en su investigación indica, que legisladores y académicos deben de actuar de cohesionada, a partir de que la investigación tecnológica debe ser coherente con los aspectos que generan efectos positivos y perjudiciales desde la inteligencia artificial.

Desde una visión con mayor integralidad que la arista anterior, el Institute of Electrical and Electronics Engineers (IEEE) (2019), en su Informe Ethically aligned design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, promueve las ideas siguientes:

- Implicar a la sociedad civil, fabricantes y gobiernos como colaboradores, para establecer y recomendar normas y códigos éticos en todos los espacios.
- Instaurar estándares y órganos regulatorios que vigilen el desarrollo de la IA.
- Trasladar las obligaciones legales de las sociedades actuales al mundo de la inteligencia artificial.
- Garantizar que la IA siempre está guiada por el control y el juicio humano, y nunca al revés, ello requiere mecanismos de control y evaluación, creando organizamos, que certifiquen que un sistema de algoritmo es justo antes de distribuirlo.

Para lograr lo anterior, el Informe Ethically aligned design sugiere: (IEEE, 2019)

- Controlar la actividad de la IA: se recomienda para los expertos, tener acceso a los datos de los sistemas autónomos que aclaren su intencionalidad, cómo funcionan o cómo interaccionan con el usuario humano.
- A desarrolladores: acercar sus resultados a la diversidad de leyes, normas y principios culturales que existen en el mundo.
- A legisladores: trabajar con transparencia en aspectos como la responsabilidad, las obligaciones y la rendición de cuentas, en un escenario en que la industria de *software* debe saber a qué ajustarse.

Finalmente, el informe Ethically aligned design recomienda (IEEE, 2019):

- Educar a legisladores y gobiernos para que la sociedad en su conjunto se enfrente a los desafíos de forma clara, dejando a un lado las confusiones sin fundamento.
- Involucrar a las personalidades con mayor credibilidad en el debate y darle la máxima difusión posible.
- Educar en ética y seguridad, para sensibilizar a la sociedad sobre los riesgos de un mal uso de la tecnología.



Lecturas a considerar

Las acciones orientadas a contrarrestar los problemas examinados, señalan que se está asumiendo una conciencia proactiva en el espacio digital para la búsqueda y perfeccionamiento de las soluciones: empresas como Facebook y Google laboran para concretar y perfeccionar sus principios y códigos morales. Microsoft ha organizado el Instituto de Investigación AI Now, con el propósito de asegurar que los sistemas respondan a las demandas sociales. Gobiernos como el de Reino Unido, España, Unión Europea, se han pronunciado por la organización de proyectos de investigaciones, para enfrentar las dificultades derivadas del uso indebido de la tecnología (Paniagua, E., [2019](#)).

De lo anterior, se deduce que la clave para enfrentar los usos maliciosos de la tecnología está en asumir un enfoque integral, que incluya los aspectos, culturales, históricos, morales, políticos, filosóficos, teniendo como centro el componente educativo por encima del instructivo. Esto indudablemente requiere, de sentido de pertenencia con el enfoque humano que debe derivarse de las ciencias que originan el desarrollo tecnológico, en un contexto social en que los valores no se vean interferidos en las aplicaciones que se ofrecen en este espacio (Alves, [2020](#)).

Ello induce a repensar las prácticas tecnológicas que distorsionan la realidad, reconociendo y enfrentando, las contradicciones y las incertidumbres éticas y epistemológicas desde un pensamiento crítico y analítico, en beneficio de la formación de una postura axiológica, y que ello permita ganar credibilidad moral y liderazgo de aquellos que se encargan del diseño del software con fines sociales, en un escenario, donde se proteja: la intimidad, confidencialidad, privacidad y la soberanía nacional de países y regiones enteras.

El efecto es que el científico se desarrolle como un sujeto moral, en empatía con la sociedad y vinculado con esta. Guiado por un código civil que se concrete en sus propuestas tecnológicas, en un ámbito que no se altere la naturaleza, su medio ambiente, y donde no se afecten los sujetos sociales. Tomando conciencia de que la construcción algorítmica en el escenario digital debe estar orientada a mejorar la sociedad y no a afectarla. Que siempre sea el individuo, el sujeto responsable de actuación de la tecnología, en lo concerniente a sus objetivos, selección y operación final. Es decir, que el control social sea significativo. Por lo que la IA no debería reemplazar las decisiones de los seres sociales, cuando se trata de cuestiones éticas.

Comentarios finales

Las aplicaciones de la inteligencia artificial, tienen un gran potencial para el desarrollo social y están revolucionando la industria de *software*, sin embargo, por su doble uso, ello exige de un empleo regulado y responsable de estas, hecho que es cada vez más importante debido a que se



están proyectando cuestiones éticas sobre la forma en que se está empleando la tecnología, y cómo está afectando la sociedad.

Las nuevas realidades que ofrecen las tecnologías están mostrando la necesidad de una transformación en los patrones jurídicos, éticos y morales, lo cual debe convertirse en una actividad y modos de actuación guiados por la práctica social que debe adjudicarse cada individuo, y la sociedad en general. Ello deberá contribuir a enfrentar las amenazas del contexto actual, por lo que la ciencia de los datos, debería repensarse en función de proteger los intereses humanos, y en contra de los fines comerciales que deterioran el comportamiento humano.

Los riesgos tecnológicos son el resultado de déficits éticos, lo que muestra la discapacidad de conectar de manera civilizada la ciencia con el ser humano; así como, la falta de habilidades sociales para compromisos sobre esta conexión, pensando globalmente y actuando localmente, y en el nivel de individuos.

La realidad examinada está indicando nuevas tareas al pensamiento ético, por lo que los riesgos a que pudiera estar sometida la sociedad, requieren colocar la mirada en una trayectoria diferente, a la que está orientando el uso malicioso de las tecnologías. Este, afecta la manera en que se construye y gestiona el ecosistema digital, así como, a la forma en que se diseña y distribuyen sus resultados, lo que requiere de políticas gubernamentales y respuestas institucionales y humanísticas.

Las ciencias naturales están siendo empleadas para la distorsión de los valores sociales, lo que requiere alertarlas para que los resultados de su investigación no se orienten al servilismo de prácticas mal intencionadas, e impedir que se legitimen acciones con impactos negativos para la sociedad. El hecho es que la práctica de la ciencia desde la IA, puesta en función de entorpecer los acontecimientos, se está transformando en un freno del pensamiento, cuestión que la desacredita.

Referencias

Alves, M. (2020). The Natural Fallacy in a Post-Truth era. A perspective on the natural sciences' permeability to values. *EMBO Reports*, 21(2), 1-4. Doi: <https://doi.org/10.15252/embr.201949859>

Berberich, N., & Diepold, K. (2018). *The Virtuous Machine - Old Ethics for New Technology?* Recuperado de <https://arxiv.org/abs/1806.10322>



- Boissier, O., Bonnet, G., Cointe, C., De Swarte, T., & Vallée, T. (2018). *Ethics and autonomous agents. Building ethical collectives.* Recuperado de <https://ethicaa.greyc.fr/media/files/ethicaa.delivvable.5.pdf>
- Bossmann, J. (2016). Top 9 ethical issues in artificial intelligence. En *World Economic Forum*. Recuperado de <https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/>
- Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation.* (Primera ed.). Future of Humanity Institute, University of Oxford, Arizona State University. Recuperado de <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf>
- Bryson, J. & Winfield, A. (2017). Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems. *Computer*, 50(1), 116-119. Doi: <https://doi.org/10.1109/MC.2017.154>
- Carreño, R. (2019). *¿Qué sucede con las plataformas digitales latinoamericanas?* Recuperado de <https://www.linkedin.com/pulse/qu%C3%A9-sucede-con-las-plataformas-digitales-ren%C3%A9-carre%C3%B1o>
- Castillero, O. (2016). Dilemas éticos: qué son, tipos y 4 ejemplos que te harán pensar. *Psicología y mente.* En *Psicología y mente.* Recuperado de <https://psicologiaymente.com/psicologia/dilemas-eticos>
- Cointe, N., Bonnet, G., & Boissier, O. (2016). Ethical Judgment of Agents' Behaviors in Multi-Agent Systems. En *ACM Digital Library.* Recuperado de <https://dl.acm.org/doi/10.5555/2936924.2937086>
- Conitzer, V., Sinnott, W., Schaich, J., Deng, Y., & Kramer, M. (2017). *Moral decision making frameworks for artificial intelligence.* Recuperado de <https://users.cs.duke.edu/~conitzer/#publications>
- Curso OPE Andalucía. (2019). *Principios fundamentales de la bioética.* Recuperado de <https://aulaplusformacion.es/wp-content/uploads/2018/09/demo-OPE-2019-andalucia-3.pdf>
- Gigova, R. (2017). Who Vladimir Putin thinks will rule the World. En *CNN.* Recuperado de <https://edition.cnn.com/2017/09/01/world/putin-artificial-intelligence-will-rule-world/index.html>



- Institute of Electrical and Electronics Engineers. (2019). *Ethically aligned design A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*. (Primera ed.). IEEE Standards Association. Recuperado de <https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html>
- Ismail, N. (2019). AI ethics: Time to move beyond a list of principles. En *Information Age*. Recuperado de <https://www.information-age.com/ai-ethics-principles-123479114/>
- Knaus, C. (2017). Ethical dilemmas and artificial intelligence. En *Gulf Time*. Recuperado de <https://www.gulf-times.com/story/566080/Ethical-dilemmas-and-artificial-intelligence>
- Kvalnes, Ø. (2019). *Moral Reasoning at Work. Rethinking Ethics in Organizations*. (Second ed.). Norway: Department of Leadership and Organizational Development. Doi: <https://doi.org/10.1007/978-3-030-15191-1>
- Laichter, L. (2019). *Ethics of Autonomous Systems. Annotated bibliography of recommended materials*. Recuperado de http://larslaichter.com/assets/docs/Ethics_of_Autonomous_Systems.pdf
- Lecher, C. (2018). What happens when algorithm cuts your health care. En *The Verge*. Recuperado de <https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy>
- López de Mántaras, R. (2019). Inteligencia artificial: temores, realidades y dilemas éticos. Revista Interactiva. En *Interactiva*. Recuperado de <https://interactivadigital.com/opinion-marketing-digital/inteligencia-artificial-temores-realidades-y-dilema-eticos/>
- Macías, C., Fernández, A., Méndez, C., Poch, J., & Sevillano, B. (2015). Inteligencia humana. Un acercamiento teórico desde dimensiones filosóficas y psicológicas. *Revista de Información Científica*, 91(3), 577-592. Recuperado de: <http://www.revinfcientifica.sld.cu/index.php/ric/article/view/235/1350>
- Mattingly-Jordan, S., Day, R. Donaldson, B., Gray, P., & Ingram, I. (2019). *Ethically Aligned Design. First Edition Glossary. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*. (Primera ed.). IEEE Standards Association. Recuperado de https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e_glossary.pdf



- Monasterio, A. (2019). Ética para máquinas: similitudes y diferencias entre la moral artificial y la moral humana. *Revista Internacional de Éticas Aplicadas*, 30, 129-147. Recuperado de <https://www.dilemata.net/revista/index.php/dilemata/article/view/412000295>
- Mourelle, D. (2019). Amazon contra Microsoft: las tecnológicas entran en la industria de defensa. En *El Orden Mundial*. Recuperado de <https://elordenmundial.com/amazon-microsoft-tecnologicas-industria-de-defensa/>
- Paniagua, E. (2019). El dilema de la tecnología ética. En *Innovadores*. Recuperado de <https://innovadores.larazon.es/es/not/el-dilema-de-la-tecnologia-etica>
- Ramos, M. (2018). La ética digital se consolida como un aspecto clave en la toma de decisiones del CIO. En *CIO*. Recuperado de <https://www.ciospain.es/liderazgo/la-etica-digital-se-consolida-como-un-aspecto-clave-en-la-toma-de-decisiones-del-cio>
- Samaniego, J. (2018a). La tecnología tiene implicaciones éticas: tenemos que dejar de mirar hacia otro lado. En *Blog Lenovo*. Recuperado de: <https://www.bloglenovo.es/la-tecnologia-abre-grandes-dudas-eticas-tenemos-dejar-mirar-hacia-lado/>
- Samaniego, J. (2018b). Solo la inteligencia artificial puede ganar la ciberguerra del futuro. En *Blog Lenovo* Recuperado de: <https://www.bloglenovo.es/amenazas-de-la-inteligencia-artificial/>
- Scherer, M. (2016). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*, 29(2), 354-400. Doi: <https://doi.org/10.2139/ssrn.2609777>
- Shaw, J. (2019). Artificial Intelligence and Ethics. Ethics and the dawn of decision-making machines. En *Harvard Magazine*. Recuperado de <https://harvardmagazine.com/2019/01/artificial-intelligence-limitations>
- Torrero, J. (2019). Ética e inteligencia artificial: Adán y Eva de silicio mordiendo la manzana prohibida. En *Nobbot*. Recuperado de <https://www.nobbot.com/firmas/etica-e-inteligencia-artificial-adan-y-eva-de-silicio/>
- Varangaonkar, A. (2018). The ethical dilemmas developers working on Artificial Intelligence products must consider. En *Packt*. Recuperado de <https://hub.packtpub.com/ethical-dilemmas-developers-on-artificial-intelligence-products-must-consider/>
- Yu, H., Shen Z., Miao, C., Leung, C., Lesser, V. & Yang, Q. (2018). *Building Ethics into Artificial Intelligence*. Recuperado de: <https://arxiv.org/pdf/1812.02953.pdf>

