Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), enero-febrero 2024,
Volumen 8, Número 1.

https://doi.org/10.37811/cl_rcm.v8i1

REGULACIÓN DEL MANEJO DE LA INTELIGENCIA ARTIFICIAL, CONSECUENCIAS Y DAÑOS A LA SOCIEDAD POR SU MAL USO

REGULATION OF ARTIFICIAL INTELLIGENCE MANAGEMENT, CONSEQUENCES AND DAMAGES TO SOCIETY DUE TO ITS MISUSE

Harold Patricio Loján Alvarado
Universidad Técnica de Machala, Ecuador

Oscar Efrén Cárdenas Villavicencio
Universidad Técnica de Machala, Ecuador

Universidad Técnica de Machala, Ecuador
Universidad Técnica de Machala, Ecuador

DOI: https://doi.org/10.37811/cl_rcm.v8i1.9596

Regulación del Manejo de la Inteligencia Artificial, Consecuencias y Daños a la Sociedad por su Mal Uso

Harold Patricio Loján Alvarado[1]
hplojan@utmachala.edu.ec
https://orcid.org/0009-0008-6869-261X
Universidad Técnica de Machala
Ecuador

Oscar Efrén Cárdenas Villavicencio
oecardenas@utmachala.edu.ec
https://orcid.org/0000-0001-6570-8040
Universidad Técnica de Machala
Ecuador

RESUMEN

La inteligencia artificial conocida por sus siglas IA ha traído grandes beneficios en los últimos años. Sin embargo, su implementación también conlleva riesgos que deben manejarse adecuadamente mediante regulaciones efectivas, especialmente en países como Ecuador donde el marco legal sobre IA es aún incipiente. Este artículo analiza la necesidad de establecer lineamientos éticos y legales sobre el uso de la IA, previniendo daños potenciales en la sociedad. Se exploran las consecuencias negativas que podría acarrear un mal uso de esta tecnología, haciendo hincapié en los vacíos legales actuales que permiten una aplicación inadecuada de la IA sin supervisión ni controles éticos. Dada la ausencia de leyes específicas en algunos países como por ejemplo Ecuador que regulen el desarrollo responsable de la IA, se proponen algunos lineamientos éticos preliminares contextualizados a la realidad local, que contemplan principios para evitar discriminación algorítmica, proteger la privacidad de datos, fomentar la transparencia y rendición de cuentas.

Palabras clave: regulaciones, inteligencia artificial, lineamientos, riesgos

[1] Autor principal.
Correspondencia: oecardenas@utmachala.edu.ec

Regulation of Artificial Intelligence Management, Consequences and Damages to Society Due to its Misuse

ABSTRACT

Artificial intelligence known by its acronym AI has brought great benefits in recent years. However, its implementation also entails risks that must be adequately managed through effective regulations, especially in countries like Ecuador where the legal framework on AI is still incipient. This article analyzes the need to establish ethical and legal guidelines on the use of AI, preventing potential harm to society. The negative consequences that could result from misuse of this technology are explored, emphasizing the current legal loopholes that allow inappropriate application of AI without supervision or ethical controls. Given the absence of specific laws in some countries such as Ecuador that regulate the responsible development of AI, some preliminary ethical guidelines are proposed contextualized to the local reality, which contemplate principles to avoid algorithmic discrimination, protect data privacy, promote transparency and accountability.

***Keywords*:** regulations, artificial intelligence, guidelines, risks

Artículo recibido 22 diciembre 2023
Aceptado para publicación: 20 enero 2024

INTRODUCCIÓN

La inteligencia artificial (IA), puede definirse como el medio por el cual las computadoras, robots y otros dispositivos realizan tareas que generalmente requieren de la inteligencia humana. Cuevas, A. Y. S., & Lino, A. E. (2020). El término IA también "se aplica a los sistemas que son capaces de analizar su entorno y pasar a la acción con cierto grado de autonomía con el fin de alcanzar objetivos específicos" Hueso, L. C. (2019). La IA se ha expandido rápidamente en años recientes, integrándose en muchas áreas como salud, transporte, finanzas, entretenimiento y más. Si bien la IA ha traído grandes beneficios a la sociedad, también conlleva riesgos potenciales si no es utilizada de forma ética y responsable (Cave y ÓhÉigeartaigh, 2018; Jobin et al., 2019).

El uso comprometido de la IA no solo es no ejercer prácticas ilegales, también es usar la IA de forma que no vulnere a las minorías ni a los derechos humanos y que no lleve al incremento de la brecha de desigualdad existente, ya sea a propósito o de forma accidental. Los riegos de la IA que aparecen ante la falta de transparencia traen consigo situaciones que pueden derivarse de defectos, errores o sesgos en la programación que atenten contra la privacidad, intimidad, dignidad, salud o patrimonio de las personas y la responsabilidad eventualmente podría estar distribuida entre distintos actores haciendo más complejo el problema. Clara, B. B., & Malbernat, L. R. (2021).

Este artículo analiza la necesidad de establecer marcos regulatorios efectivos sobre el desarrollo y uso de la IA. Se estudian las posibles consecuencias negativas de un mal uso de la misma y se proponen lineamientos éticos y legales para promover un avance responsable de esta tecnología. La IA ha demostrado un gran potencial para automatizar tareas, mejorar la eficiencia de procesos, personalizar servicios y contribuir al progreso científico y tecnológico (Esteva et al., 2019). El uso de IA por parte de empresas tecnológicas, instituciones académicas y agencias gubernamentales se ha disparado en años recientes. Se estima que la inversión global en IA superó los US$50 mil millones en 2020 y seguirá creciendo exponencialmente (Hao, 2022; Marr, 2022).

Si bien la IA ofrece beneficios significativos, también conlleva una variedad de riesgos éticos y sociales que no pueden ser ignorados (Corbett-Davies y Goel, 2018; Papernot et al., 2018). Por ejemplo, los algoritmos de aprendizaje automático pueden replicar y amplificar sesgos humanos si se entrenan con datos no representativos o discriminatorios (Binns, 2018). Otros peligros incluyen vulneración de

privacidad, pérdida de transparencia y agencia humana, sesgos en la toma de decisiones y disrupción de empleos, entre otros (Tegmark, 2017).

Dada la creciente influencia de la IA en numerosos ámbitos, resulta crucial manejar de forma proactiva sus riesgos asociados. Este artículo analiza la necesidad de establecer marcos regulatorios y lineamientos éticos efectivos sobre el desarrollo y implementación de sistemas de IA, dado que mediante una regulación adecuada, es posible potenciar los beneficios de la IA y mitigar sus perjuicios potenciales (Jobin et al., 2019). Simultáneamente, son pocas las empresas tecnológicas que han adoptado lineamientos éticos sólidos o que realizan auditorías rigurosas de sus algoritmos en busca de discriminación y otros daños (Corbett-Davies et al., 2017).

Esta falta de supervisión efectiva genera una responsabilidad difusa sobre los impactos negativos de la IA y dificulta la rendición de cuentas cuando surgen problemas (Cave y ÓhÉigeartaigh, 2018). Para colmar este vacío, se requieren esfuerzos conjuntos de múltiples actores, incluyendo empresas, gobiernos, academia y sociedad civil (Tegmark, 2017). Las iniciativas deben abarcar regulación legal, autorregulación industrial, estándares técnicos, educación en ética algorítmica y mayor participación pública en la gobernanza de la IA (Shokri et al., 2011). Con el marco regulatorio adecuado, es posible orientar el desarrollo de la IA hacia el bien común y alinearla con valores humanos fundamentales como justicia, dignidad y solidaridad. Este artículo aporta al debate actual proponiendo lineamientos concretos para la gobernanza ética de la IA.

METODOLOGÍA

El presente estudio utilizó una metodología cualitativa con alcance descriptivo para analizar los aspectos éticos y sociales relacionados con la inteligencia artificial (IA). La recolección de datos se realizó mediante dos técnicas:

Revisión bibliográfica

Se realizó una exhaustiva revisión de literatura utilizando las bases de datos académicas Scopus, Web of Science y IEEE Xplore, también se usaron páginas como MIT Technology Review para extraer informacion actualizada. Las palabras clave empleadas en las búsquedas fueron: "ética", "inteligencia artificial", "sesgos", "discriminación algorítmica", "privacidad", "transparencia", entre otras relevantes.

Se analizaron artículos, libros, informes y otras fuentes documentales publicadas en los últimos 5 años en inglés y español.

Entrevistas semiestructuradas

Se realizaron entrevistas semiestructuradas a cinco ingenieros expertos en IA y cinco abogados conocedores del tema. Las entrevistas exploraron perspectivas sobre vacíos éticos y legales en el campo de la IA. Se diseñó una guía de preguntas utilizando un lenguaje claro y técnico. Las sesiones fueron grabadas en audio previo consentimiento de los participantes y posteriormente transcritas para su análisis.

Se realizó un reporte de las entrevistas, utilizando palabras claves de los resultados y sintetizando las mismas. Los resultados se presentan de forma descriptiva resaltando los hallazgos y discusiones más destacadas en torno a la ética y regulación de la IA. Estos elementos sugeridos permitirán a los lectores conocer las estrategias metodológicas, además de valorar su rigor y coherencia, así como la replicabilidad de los procedimientos y del estudio.

RESULTADOS Y DISCUSIÓN

Crecimiento y riesgos potenciales de la IA

El auge de la Inteligencia artificial

La IA ha experimentado un crecimiento exponencial en la última década, impulsado por los rápidos avances en capacidad computacional, disponibilidad de grandes conjuntos de datos y algoritmos de aprendizaje profundo (Esteva et al., 2019). Las inversiones globales en IA por parte del sector privado se duplicaron entre 2017 y 2021, superando los US$50 mil millones (Hao, 2022). Se espera que el mercado de IA alcance US$500 mil millones para 2024 (Marr, 2022).

Prácticamente todos los sectores han adoptado aplicaciones de IA:

- Salud: diagnóstico asistido, detección de tumores, monitorización de pacientes.
- Finanzas: análisis de riesgo crediticio, detección de fraude, trading algorítmico.
- Transporte: vehículos autónomos, logística optimizada, drones de entrega.
- Venta al por menor: recomendación de productos, predicción de demanda, automatización de inventarios.
- Gobierno: vigilancia, aplicación de leyes, servicios públicos.

Si bien la IA permite automatizar procesos, personalizar servicios y mejorar la toma de decisiones, también implica una variedad de riesgos potenciales (Corbett-Davies y Goel, 2018):

- Discriminación algorítmica y resultados injustos.
- Vulneración de privacidad mediante vigilancia masiva.
- Pérdida de transparencia y agencia humana.
- Desempleo tecnológico al automatizar trabajos.
- Sesgos y errores en sistemas opacos.
- Ciberataques que aprovechan vulnerabilidades de la IA.
- Uso de IA para manipulación y control social.

Estos peligros demuestran la necesidad de establecer regulaciones, controles éticos y una supervisión más estrecha sobre el desarrollo de la IA.

Sesgos algorítmicos

Uno de los principales riesgos éticos de la IA es la posibilidad de que replique y amplifique sesgos humanos (Jobin et al., 2019). Los algoritmos de aprendizaje automático detectan patrones en los datos con los que son entrenados. Si esos datos contienen sesgos, prejuicios o falta de representatividad, esos problemas se heredarán en los modelos resultantes (Binns, 2018).

Algunas fuentes comunes de sesgos algorítmicos son:

- Datos no representativos: muestras pequeñas o no aleatorias pueden llevar a generalizaciones erróneas.
- Variables proxy discriminatorias: el uso de atributos como el código postal como sustituto para la raza puede perpetuar la discriminación.
- Etiquetado de datos sesgado: la clasificación humana de los datos usados en entrenamiento también puede estar sesgada.
- Falta de diversidad en equipos de desarrollo: los sesgos inconscientes de los programadores pueden permear el diseño de algoritmos.

Los sesgos algorítmicos ya han provocado escándalos en contextos como calificación de exámenes, evaluación docente, predicción del crimen y publicidad online (Corbett-Davies et al., 2017). Supervisar, auditar y mejorar la calidad de los datos son pasos cruciales para abordar esta problemática en IA.

Riesgos de ciberseguridad.

Los sistemas de IA también implican nuevos desafíos para la ciberseguridad (Papernot et al., 2018). Algunos ejemplos son:

- Ataques de envenenamiento de datos: manipular los datos de entrenamiento para alterar el comportamiento de modelos.
- Suplantación (spoofing) de modelos de IA: engañar a un sistema haciéndose pasar por un modelo legítimo.
- Extracción de modelos: robar parámetros y arquitectura de modelos entrenados.
- Backdoors en modelos: insertar vulnerabilidades ocultas para comprometer modelos desplegados.
- Ataques adversarios: técnicas para engañar modelos y causar fallos.
- Los sistemas de IA deben ser diseñados teniendo en cuenta estas amenazas emergentes. Prácticas como cifrado, firma digital y blockchain pueden mejorar la ciberseguridad de la IA.
- 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. (Arencibia & Cardero, Dilemas éticos en el escenario de la inteligencia artificial, 2020).

Riesgos socioeconómicos

Más allá de los aspectos éticos y técnicos, la IA también conlleva riesgos sociales y económicos que requieren atención (Tegmark, 2017):

- Concentración del poder de mercado en grandes empresas tecnológicas.
- Disrupción de empleos y desplazamiento de trabajadores.
- Aumento de desigualdad si los beneficios se acumulan en pocas manos.
- Manipulación política mediante personalización algorítmica de contenidos.
- Efectos desestabilizadores sobre mercados financieros dominados por IA.
- Dinámicas geopolíticas cada vez más condicionadas por carreras tecnológicas en IA.

Abordar estos desafíos obliga a pensar en la IA no sólo en términos técnicos, sino dentro de contextos sociales, políticos y económicos más amplios.

Consecuencias Negativas de un mal uso de la IA

Un mal uso de la IA puede tener graves consecuencias para los individuos y la sociedad (Cave y ÓhÉigeartaigh, 2018):

A. Vulneración de derechos humanos: los sistemas de IA podrían usarse para vigilancia masiva, manipulación psicológica, represión política, etc.
B. Discriminación: La IA podría replicar prejuicios humanos y llevar a decisiones injustas sobre préstamos, contratación laboral, libertad condicional, etc.
C. Pérdida de privacidad: Grandes volúmenes de datos personales utilizados para entrenar modelos de IA pueden llevar a abusos de privacidad.
D. Desempleo: La automatización de trabajos mediante IA puede perturbar mercados laborales y generar desempleo.
E. Armas autónomas letales: El uso de IA en armas puede hacer la guerra más probable al distanciar a los humanos del ciclo de decisiones

Estos son sólo algunos ejemplos de los daños potenciales que podrían resultar de una IA mal utilizada. Los riesgos se magnifican por la opacidad y complejidad de muchos sistemas de IA actuales.

Lineamientos Éticos y Regulatorios

Para mitigar los riesgos, es esencial establecer lineamientos éticos y marcos regulatorios sobre el desarrollo y uso de la IA (Shokri et al., 2011):

- Promover la transparencia en los sistemas de IA, para que las decisiones sean explicables.
- Monitorear y auditar los sistemas de IA para detectar discriminación y otros daños.
- Diseñar la IA con enfoque de privacidad desde el inicio.
- Crear comités de ética de IA en empresas y gobiernos.
- Prohibir ciertos usos dañinos de la IA, como en vigilancia masiva.
- Exigir evaluaciones de impacto algorítmico antes de implementar sistemas críticos de IA.
- Establecer responsabilidad legal por daños causados por sistemas de IA.
- Invertir en educación sobre ética de IA.

Para que la IA opere de manera eficiente, se requieren volúmenes sustanciales de datos para entrenar y perfeccionar los algoritmos. Por ende, resulta fundamental establecer políticas y marcos legales bien

definidos en relación con la recopilación y el empleo de datos. Estos lineamientos requerirán esfuerzos conjuntos de múltiples actores como empresas, gobiernos, academia y sociedad civil para ser efectivos.

Diferentes puntos de vista frente al uso de la IA

A. Resumen de entrevistas con los ingenieros del área de Tecnologías de la Información

Estas son las preguntas que fueron tratadas en la entrevista:

1. ¿Cuáles son en su opinión los principales riesgos éticos que enfrenta actualmente el desarrollo de la IA?

   Los entrevistados coinciden que los más preocupantes son la posibilidad de sesgos algorítmicos que lleven a resultados discriminatorios o injustos, la amenaza a la privacidad por el uso de enormes volúmenes de datos, y la opacidad de muchos sistemas que dificulta la auditoría y la rendición de cuentas.

2. ¿Cree que hace falta una mayor regulación gubernamental de la IA?

   Sin duda, pero no debe ser la única aproximación. La autorregulación de la industria y los estándares técnicos también son importantes. Y se requiere una mayor participación de la sociedad civil en las discusiones sobre cómo moldear el futuro de la IA

3. ¿Qué medidas concretas se podrían tomar para fomentar un desarrollo ético de la IA?

   Por ejemplo, realizar rigurosas auditorías algorítmicas buscando sesgos; invertir más en técnicas para explicar y dar transparencia a los sistemas de IA; diseñar procesos internos de revisión ética de proyectos de IA; capacitar exhaustivamente a los equipos en temas de ética y derechos humanos.

4. ¿Cómo ve el panorama a futuro? ¿Es optimista?

   Indican que ven una creciente conciencia de los desafíos éticos en IA, tanto en la industria como en la academia. Pero queda mucho por hacer. Superar estos retos requerirá un esfuerzo sostenido a largo plazo, involucrando a múltiples actores. Debemos asegurarnos de desarrollar la IA de forma compatible con los valores humanos fundamentales.

B. Resumen de entrevistas con los abogados conocedores de IA

1. ¿Cuáles diría que son los principales vacíos legales actuales con respecto al uso de la IA?

   Piensan que una de las áreas más críticas es la de responsabilidad legal. Cuando un sistema de IA causa algún daño, a menudo no está claro quién debería ser legalmente responsable. Las leyes actuales no estaban diseñadas pensando en la IA.

2. ¿Cómo debería abordarse estos temas?

   Se necesitan leyes que establezcan claramente quién es responsable cuando un algoritmo, robot o sistema de IA autónomo cause daños. Esto podría incluir responsabilidad compartida entre diversos actores como los desarrolladores, las empresas que los implementan y los usuarios finales.

3. Más allá de la responsabilidad, ¿en qué otras áreas ven vacíos legales importantes?

   Otro tema clave es la privacidad y el uso de datos personales. Las regulaciones actuales no contemplan adecuadamente los enormes volúmenes de datos que se usan para entrenar sistemas de IA, incluyendo información personal muy sensible. Se requieren nuevas normas para equilibrar innovación e intimidad. También hacen falta leyes que aborden posibles sesgos algorítmicos y discriminación en IA, especialmente en áreas sensibles como contratación, préstamos y decisiones de justicia penal. Y normas para un uso ético de la IA en vigilancia, cuidado de la salud, vehículos autónomos, entre otras áreas de alto impacto.

4. ¿Cree que las empresas tecnológicas deberían autorregularse más mientras se desarrolla este marco legal?

   Sin duda, la autorregulación responsable de la industria es esencial, pero no basta por sí sola. Se necesita un enfoque integral con cooperación público-privada y participación de la sociedad civil para gobernar éticamente estas tecnologías poderosas. Todos los actores deben involucrarse en este debate.

Riesgos emergentes de la IA

La rápida adopción de sistemas de IA en diversos ámbitos también conlleva nuevos riesgos que deben ser estudiados y abordados (Corbett-Davies y Goel, 2018). Una categoría importante son los problemas relacionados con sesgos y equidad algorítmica. Varios estudios han demostrado que los algoritmos de

aprendizaje automático pueden replicar sesgos presentes en los datos usados para entrenarlos, derivando en discriminación contra ciertos grupos (Binns, 2018; Corbett-Davies et al., 2017; Jobin et al., 2019). Por ejemplo, un sistema de evaluación docente en EEUU fue encontrado generando calificaciones más bajas para profesores de minorías, debido a sesgos en los datos históricos (Corbett-Davies et al., 2017). Superar estos sesgos requiere un cuidadoso diseño de los sistemas de IA, utilizando técnicas de aprendizaje justo y auditando rigurosamente los modelos.

Otra categoría de riesgos surge del uso intensivo de datos por parte de la IA, especialmente información personal sensible. Varios autores han llamado la atención sobre las amenazas a la privacidad derivadas de entrenar algoritmos con enormes cantidades de datos privados (Cave y ÓhÉigeartaigh, 2018; Papernot et al., 2018; Tegmark, 2017). Se requieren salvaguardas técnicas y regulatorias para prevenir abusos.

Asimismo, la adopción acelerada de la IA conlleva desafíos geopolíticos y una posible "carrera armamentística" por la IA entre potencias (Cave y ÓhÉigeartaigh, 2018). Esta dinámica podría desestabilizar el equilibrio global de poder y aumentar los riesgos de conflictos. Promover la cooperación internacional es clave para guiar la IA hacia fines pacíficos y evitar una competencia tecnológica descontrolada. Asimismo, la adopción acelerada de la IA conlleva desafíos geopolíticos y una posible "carrera armamentística" por la IA entre potencias (Cave y ÓhÉigeartaigh, 2018). Esta dinámica podría desestabilizar el equilibrio global de poder y aumentar los riesgos de conflictos. Promover la cooperación internacional es clave para guiar la IA hacia fines pacíficos y evitar una competencia tecnológica descontrolada.

CONCLUSIONES

La IA tiene un gran potencial para beneficiar a la sociedad, pero sólo si se desarrolla y utiliza de forma ética y responsable. Los riesgos y daños potenciales analizados en este artículo obligan a establecer regulaciones, lineamientos éticos y controles claros para guiar el avance de la IA tanto en el sector público como privado.

Múltiples actores, incluyendo empresas tecnológicas, gobiernos, academia y sociedad civil, deben involucrarse activamente en crear marcos regulatorios efectivos que maximicen los beneficios de la IA y minimicen sus perjuicios a nivel local y global. La regulación efectiva del manejo de la inteligencia

artificial es esencial para salvaguardar a la sociedad de consecuencias y daños no deseados.

Queda un largo camino por recorrer para garantizar un futuro de la IA alineado con los valores humanos fundamentales como justicia, dignidad y derechos individuales. Las iniciativas propuestas en este artículo son un primer paso, pero se requerirán esfuerzos sostenidos y discusiones inclusivas con diversas perspectivas para desarrollar una gobernanza ética robusta de estas poderosas tecnologías que ya están transformando nuestras sociedades de maneras sin precedentes.

Los principios éticos deben ser el norte que guíe la investigación, desarrollo e implementación de la IA, para el beneficio de toda la humanidad.

REFERENCIAS BIBLIOGRAFICAS

Cuevas, A. Y. S., & Lino, A. E. (2020). Inteligencia artificial un peligro latente. Monografía de Grado. Perú: Universidad Mayor de San Marcos. https://www.academia.edu/30039422/Inteligencia_artificial_un_peligro_latente.

Hueso, L. C. (2019). Riesgos e impactos del Big Data, la inteligencia artificial y la robótica: enfoques, modelos y principios de la respuesta del derecho. Revista general de Derecho administrativo, (50), 1-37. https://www.iustel.com/v2/revistas/detalle_revista.asp?id_noticia=421227

Clara, B. B., & Malbernat, L. R. (2021). RIESGOS, DILEMAS ÉTICOS Y BUENAS PRÁCTICAS EN INTELIGENCIA ARTIFICIAL. XXIII Workshop de Investigadores en Ciencias de la Computaci´on, (págs. 155-159). Argentina. https://sedici.unlp.edu.ar/bitstream/handle/10915/119977/Ponencia.pdf-PDFA.pdf?sequence=1&isAllowed=y

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2019). Deep learning in medical image analysis. IEEE Access, 7, 29742-29773. https://doi.org/10.1109/ACCESS.2019.2905215

Hao, K. (2022, January 12). In 2020 companies probably spent over $50 billion on AI—we still have no idea how to make it ethical. MIT Technology Review. https://www.technologyreview.com/2022/01/12/1044880/ai-ethics-investments/

Marr, B. (2022, January 17). The 5 biggest artificial intelligence (AI) trends in 2022. Forbes. https://www.forbes.com/sites/bernardmarr/2022/01/17/the-5-biggest-artificial-intelligence-ai-

trends-in-2022/

Corbett-Davies, S., & Goel, S. (2018). The measure and mismeasure of fairness: A critical review of fair machine learning. arXiv preprint arXiv:1808.00023. https://arxiv.org/abs/1808.00023

Cave, S., & ÓhÉigeartaigh, S. (2018). An AI race for strategic advantage: Rhetoric and risks. In Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (pp. 36-40).

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. Conference on Fairness, Accountability and Transparency, 81, 149-159. https://dl.acm.org/doi/abs/10.1145/3287560.3287573

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic decision making and the cost of fairness. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 797-806). https://doi.org/10.1145/3097983.3098095

Arencibia, M. G., & Cardero, D. M. (2020). Dilemas éticos en el escenario de la inteligencia artificial. Economía y Sociedad vol.25 n.57 Heredia Jan./Jun. 2020. http://dx.doi.org/10.15359/eys.25-57.5.

Papernot, N., Song, S., Mironov, I., Raghunathan, A., Talwar, K., & Erlingsson, Ú. (2018). Sok: Security and privacy in machine learning. 2018 IEEE European Symposium on Security and Privacy (EuroS&P), 399-414. https://doi.org/10.1109/EuroSP.2018.00035

Tegmark, M. (2017). Benefits & risks of artificial intelligence. Future of Life Institute. https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/

Shokri, R., Theodorakopoulos, G., Le Boudec, J. Y., & Hubaux, J. P. (2011). Quantifying location privacy: The case of sporadic location exposure. In Proceedings of the 11th international symposium on privacy enhancing technologies (pp. 57-76). https://doi.org/10.1007/978-3-642-22263-4_4