Miradas Críticas sobre la Tecnología: Auditoría de TI y Ética en IA

Marc A. González Figueroa
Maestría en Ciencias de Computadoras
Mentor: Alfredo Cruz, Ph.D.
Universidad Politécnica de Puerto Rico
Graduate Project EXPO, Mayo 2025

Resumen *— Este artículo integra, la Auditoría en Tecnología de la Información, la Ética en el uso de la Inteligencia Artificial y el impacto de estas herramientas en la sociedad. En primer lugar, integra los procesos de la Auditoría a los nuevos entornos de riesgo que presentan la IA. Para ello, describe la evaluación de controles, la recolección de evidencia y la presentación de hallazgos en un nuevo contexto. Hace hincapié en las auditorías de IA, en lo que se refiere a la ciberseguridad, y le da mucho peso a la integridad, la confidencialidad y la disponibilidad de los sistemas. Lo hace dentro de un marco de Ética, que no es específica de la IA, pero se presenta como muy necesaria y obligatoria para los auditores, por las implicaciones que tiene el uso de la IA y la falta de facultades que tienen para determinar si un algoritmo es justo o no.*

Términos *– Auditoría de TI, Ética en IA, Integridad de datos, Responsabilidad tecnológica.*

INTRODUCCIÓN

La forma en que la gente, las organizaciones y los gobiernos interactúan y gestionan la información y toman decisiones ha sido transformada de manera muy intensa en los últimos años. Esto es un efecto directo, por supuesto, del crecimiento acelerado de las tecnologías digitales. Pero no se trata solamente de la rapidez con la que estas tecnologías se están desarrollando: también su capacidad y la fuerza con la que "invaden" todos los espacios de nuestra vida están resultando absolutamente impresionantes. Así estamos viendo que, en cada instante, por cualquier asunto, son decididas innumerables interacciones a través de tecnologías digitales.

La auditoría de TI evalúa los sistemas computacionales de una entidad de forma sistemática y metódica, buscando asegurarse de que estos sistemas cumplan con los estándares de integridad, disponibilidad y confidencialidad. Es un proceso en el cual se detectan vulnerabilidades, se validan controles internos y se emiten recomendaciones que favorecen la toma de decisiones afinadas, con la mira en prevenir, por un lado, que se produzcan fallos técnicos de envergadura y, por otro, que se violen la seguridad y la privacidad de la información.

La ética en IA se ocupa, dentro de ese contexto, de establecer principios morales que protejan la integridad de los algoritmos, asegurando que estos actúen sin sesgos y, además, que respeten la privacidad, la equidad y los derechos humanos [1].

Esta propuesta de revisión integradora aborda los dos campos mencionados en su título: auditoría y ética tecnológica. La unión de estos dos campos del saber favorece el surgimiento de una cultura tecnológica que vale la pena ayudar y hacer crecer. Esta cultura tiene tres elementos en su favor. En primer lugar, es una cultura que se orienta a la rendición de cuentas. En segundo lugar, es una cultura que está, necesariamente, orientada a la prevención. Y, en tercer lugar, esta cultura es, sin duda, una cultura del bien común.

PLANTEAMIENTO DEL TEMA

El actual mundo de hiperconectividad, donde la automatización y la inteligencia artificial (IA) están profundamente integradas en procesos críticos de las organizaciones, plantea una problemática compleja que combina aspectos técnicos, éticos y de gobernanza. La creciente dependencia de los sistemas tecnológicos ha facilitado la llegada de avances y mejoras significativas en la eficiencia, en el análisis de datos y en la toma de decisiones que, por razones obvias, se espera sea más "inteligente" y más "automática". Sin embargo, este progreso tiene un lado oscuro que se refleja en riesgos latentes que requieren atención y que están conectados a dos aspectos en los que vale la pena reflexionar: uno, la

falta de rigurosidad en las auditorías de los sistemas de información y dos, la ausencia de marcos éticos bien definidos que guíen la implementación de la IA [2],[3].

A pesar de los avances en la ciberseguridad, muchas organizaciones aún no han incorporado prácticas robustas de auditoría de Tecnología de la Información (TI). Las auditorías de TI permiten no solo identificar vulnerabilidades, sino también asegurar el cumplimiento normativo y proteger la integridad, confidencialidad y disponibilidad de los datos. Sin embargo, en muchas instituciones, esas prácticas no se realizan de forma periódica y profunda, lo que deja expuestos a los sistemas ante ciberataques, pérdida de información, y errores de configuración que podrían evitarse con un enfoque preventivo y sistemático.

La Figura 1 contiene el Modelo CIA de Seguridad Informática, que está formado por los tres conceptos de seguridad, que son fundamentales para la protección de los sistemas de información. Estos conceptos son: Confidencialidad, Integridad y Disponibilidad.

Figura 1
El Triángulo de la CIA

La intersección de estos conceptos, o el espacio común que en ellos se encuentra, es su interdependencia. Es guardada con el mismo celo que cualquiera de los tres, y es tan vital como cualquiera de los tres. Es el aspecto que en el plano horizontal se cuela por la intersección de los tres conceptos y también por la parte superior e inferior de la figura.

Esa es la manera en que el plano y el espacio en el que se encuentran las intersecciones entre Confidencialidad, Integridad y Disponibilidad se comunican, en lo que podríamos llamar una figura plana. La base de la figura es la intersección de los tres. Y es tan vital como el mismo concepto de Seguridad. La Integridad. La Confidencialidad.

Tenemos el después de enforcarnos en el triángulo de la CIA el concepto del ciclo de un Auditor. En la Figura 2 representa el ciclo de un auditor, estructurado en cuatro fases principales que conforman un proceso continuo y sistemático.

Figura 2
Ciclo de Auditoría de TI

El ciclo inicia con la Planificación, etapa en la que se definen los objetivos, el alcance, los recursos necesarios y se identifican los riesgos tecnológicos relevantes. Le sigue la fase de Recolección de Evidencia, donde se examinan los controles, procesos y sistemas mediante técnicas como entrevistas, análisis de registros y pruebas técnicas. Posteriormente, en la fase de Informe, se documentan los hallazgos, se evalúa el nivel de cumplimiento con las políticas y normativas y se proponen recomendaciones específicas para corregir debilidades o mitigar riesgos. Finalmente, la fase de Seguimiento verifica que las acciones correctivas se hayan implementado y que los problemas identificados se hayan resuelto adecuadamente. Este ciclo cerrado permite a las organizaciones mantener un sistema de control actualizado y funcional, garantizando la mejora continua de su infraestructura tecnológica.

Por el contrario, la implementación de sistemas de IA en sectores sensibles, como el judicial, el militar y el biomédico, presenta un desarrollo ético inquietante [4]. Casos como los de algoritmos que discriminan por raza o género, o decisiones autónomas en contextos militares sin intervención humana, evidencian:

- La falta de supervisión en estos sistemas.
- La falta de principios éticos en el diseño de estos sistemas.

Además, muchos sistemas de IA se han entrenado con datos sesgados, perpetuando patrones de exclusión y desigualdad. Esto se ha documentado en ejemplos como el sistema COMPAS en Estados Unidos, el cual sobrestimaba el riesgo de reincidencia en personas afroamericanas, o en plataformas de reclutamiento automatizado que favorecen candidatos masculinos debido a datos históricos discriminatorios. Estos ejemplos ilustran cómo la falta de una evaluación ética desde el diseño, entrenamiento y despliegue de los algoritmos puede amplificar injusticias estructurales.

El dilema ético en la toma de decisiones entre seres humanos y sistemas de Inteligencia Artificial (IA) se representa en la Figura 3. En ella se presenta una balanza que simboliza la tensión entre la justicia, asociada a las decisiones humanas, y la precisión, representada por los algoritmos. Esta comparación gráfica resalta la tendencia de los humanos a considerar al tomar decisiones (típicamente judiciales) más factores morales, contextuales y empáticos que decisiones "automáticas" que tomarían por ejemplo una IA.

Por otra parte, los sistemas automatizados tienden a resolver los problemas considerando de forma predominante la eficiencia y la exactitud matemática, y la base de datos que alimenta sus decisiones es simplemente asombrosa por la "humildad" de que ella misma no tiene en cuenta ningún juicio ético. Para los problemas de vida o muerte, decidir en la frontera entre zonas moralmente equidistantes debería ser un asunto muy serio. Esta figura pone en evidencia la necesidad de encontrar un equilibrio entre ambos lados, ambos abordajes.

Figura 3
Dilemas Éticos en la Inteligencia Artificial

Estos factores combinados ponen de manifiesto una crítica carencia en la integración de la auditoría de TI con la ética de la IA. No es suficiente con asegurar que los sistemas funcionen, hay que garantizar que lo hagan de una manera justa, explicable y responsable.

Esta carencia se ve acentuada por una ausencia de marcos regulatorios universales y una falta de políticas públicas adaptadas al vertiginoso avance de las tecnologías. En la inteligencia artificial se está desarrollando algo a una velocidad mayor que nunca, pero se está evaluando a una velocidad muy inferior, se les está poniendo a funcionar sin que haya nadie verificando si esas evaluaciones están dando buenos o malos resultados y cuán responsables son esos resultados.

La Tabla 1 muestra ejemplos auténticos de sesgos algorítmicos en los sistemas de inteligencia artificial utilizados por reconocidas organizaciones. El caso de COMPAS ilustra cómo un sistema judicial sobrestimó el riesgo de reincidencia en personas afroamericanas. El algoritmo de reclutamiento de Amazon favorecía a candidatos hombres, lo cual es un reflejo de los sesgos presentes en los datos históricos. Por último, el “chatbot Tay” de Microsoft terminó propagando contenido ofensivo tras interactuar con usuarios en redes sociales. Estos casos, pues, muestran por qué es

importante incorporar principios éticos y mecanismos de control en el diseño y el uso de algoritmos.

Tabla 1
Casos Reales de Sesgos Algorítmicos

Caso	Descripción	Resultado
Compas	Sistema de risco criminal usadas de Estados	Sobre estimado el riesgo de recibimiento
Amazon Hiring tool	Algoritmo de reclutamiento de empleados por Amazon	Favorecían los candidatos masculinos
Microsoft Tay Bot	Un chatbot lanzado por Amazon	El chatbot propago contenido ofensivo

Por lo tanto, en este trabajo se plantea un problema que no es solo técnico, sino más bien de estructura institucional: la falta de una cultura integrada de auditoría y ética en la tecnología, que permita prevenir, mitigar y corregir el impacto de los sistemas automatizados en las personas y comunidades, y que en su mayoría desempaña el rol de los sistemas de justicia. Ya sea al permitir o impedir hacer auditorías efectivas que realmente impacten de manera positiva en el diseño de los sistemas y en sus efectos sobre las personas, los sistemas de justicia determinan si hay una cultura de auditoría y justicia en la tecnología o no.

OBJETIVO DEL PROYECTO

La meta principal de este proyecto es analizar y poner de relieve la importancia de integrar la auditoría de Tecnologías de la Información y los principios éticos en el desarrollo y uso de los sistemas de Inteligencia Artificial. Se trata de un trabajo de concienciación, no solo sobre los riesgos técnicos, sino también sobre los morales que acarrea la automatización y el uso de algoritmos no supervisados. Esos riesgos hacen necesaria la implementación de controles que garanticen la transparencia, la seguridad, la equidad y la responsabilidad en los entornos digitales [5]. Pero no es solo la concienciación lo que busca este proyecto.

También pretende ofrecer herramientas conceptuales y prácticas que permitan evaluar de forma integral los sistemas tecnológicos. Y lo que este proyecto también busca es, a través de esas evaluaciones, que se considere no solo la robustez técnica de esos sistemas, sino también su impacto social.

En eso el proyecto tiene un especial énfasis en la protección de los derechos humanos, la privacidad de los datos y la toma de decisiones. El hacer esto: concienciar, ofrecer herramientas y hacer evaluaciones con esas herramientas, es parte del trabajo que este proyecto se adjudica. Y lo que se busca a través de todo esto es fortalecer en las universidades una sana cultura tecnológica.

PREGUNTAS SOBRE LA INVESTIGACIÓN

Las siguientes preguntas fueron piezas claves para hacer el desarrollo de la investigación en los temas de Ética en la inteligencia artificial y Auditoria en TI:

- ¿Cuáles son las fases fundamentales que componen una auditoría de Tecnología de la Información?
- ¿Qué tipos de auditorías existen en el ámbito de la ciberseguridad y cómo se aplican en las organizaciones?
- ¿Qué herramientas o metodologías son más utilizadas para realizar una auditoría técnica efectiva?
- ¿Cómo se garantizan los principios de confidencialidad, integridad y disponibilidad durante una auditoría?
- ¿Qué consecuencias puede tener una organización por no realizar auditorías periódicas a sus sistemas de información?
- ¿Qué principios éticos deben considerarse al diseñar e implementar sistemas de inteligencia artificial?
- ¿Qué riesgos éticos presentan los algoritmos de IA cuando no se supervisan adecuadamente?
- ¿Cómo afecta el sesgo algorítmico a la equidad y justicia en sistemas automatizados?

- ¿Qué marcos regulatorios existen o se proponen para garantizar una IA ética y responsable?
- ¿De qué manera se pueden integrar las auditorías tecnológicas con mecanismos de evaluación ética en sistemas inteligentes?

PERTINENCIA E IMPORTANCIA

La pertinencia de este proyecto proviene de la necesidad de integrar de forma urgente la auditoría de sistemas con principios éticos en el desarrollo y el uso de tecnologías emergentes como inteligencia artificial. En un entorno digital cada vez más automatizado, resulta fundamental asegurar la seguridad, la equidad y la transparencia de los sistemas que procesan y deciden sobre información crítica [6].

Este trabajo robustece la cultura de la ciberseguridad al aportar métodos para la evaluación sistemática de riesgos tanto técnicos como éticos, y ofrece herramientas muy prácticas a educadores y profesionales de la industria que están interesados en promover entornos tecnológicos más responsables y confiables.

METODOLOGÍA Y DISEÑO

En este proyecto se empleó, una metodología cualitativa, exploratoria y aplicada, que busca integrar el saber técnico con el saber ético, en lo que respecta a la evaluación y desarrollo de sistemas de tecnología de la información y de inteligencia artificial.

Lo que se hizo en realidad fue el desarrollo de dos módulos formativos con objetivos complementarios: uno, centrado en la auditoría de TI y ciberseguridad, y el otro, en la ética de la inteligencia artificial. En ambos módulos se combinó teoría con análisis crítico y con el estudio de casos reales. Asimismo, en ambos módulos se realizaron simulaciones prácticas y se usaron herramientas especializadas.

Fase 1: Marco Teórico y Conceptual

Cada módulo comenzó con una introducción a los conceptos básicos. En el Módulo 1, se establecieron los fundamentos de la auditoría de TI, se definieron los principios de confidencialidad, de integridad, y de disponibilidad (Modelo CIA), se trataron los distintos tipos de auditoría, se esbozó el rol del auditor, y se debatió la diferencia entre auditoría interna y externa, y se explicó la importancia de cumplir normativas como la ISO 27001 y el GDPR [7].

El Módulo 2 presentó la ética en IA, y se empezó a profundizar en elementos tales como: principios que deben regir a la IA: justicia, transparencia, privacidad, responsabilidad, autonomía [8],[9]. Ciclo de vida de un sistema de IA (no necesariamente ético) que va desde su diseño hasta su monitoreo continuo en la vida real [10].

Fase 2: Análisis de Herramientas y Prácticas

En la Tabla 2 podemos ver cómo se realizaron investigaciones y se aplicaron varias herramientas de auditoría tecnológica como:

- Invgate que es para administración de activos tecnológicos, licencias de software y contratos de servicios [11].
- Segundo tenemos Netwrix Auditor que es para rastrear cambios en el entorno de TI, registrar eventos de seguridad y generar informes [12].
- Por último, tenemos Risk Cloud de LogicGate que emplea para elaborar modelos y realizar evaluaciones y seguimientos de los riesgos operativos y de seguridad informática [13].

Tabla 2
Herramientas en la Auditoria de TI

Herramienta	Funcionalidad	Uso en el módulo	Ventajas
InvGate	Gestión de inventario de activos de TI	Registro y clasifica-ción de dispositivos	Interfaz intuitiva y fácil de usar
Netwrix Auditor	Auditoria de cambios en sistemas de TI	Monitoreo de cambios y actividad	Reportes automá-ticos de eventos
Logic Gate	Gestión de Riesgo	Evalúa de riesgo	Integra-ciones flexibles

Fase 3: Análisis Ético y Estudio de Casos

La reflexión crítica y la evaluación ética de los sistemas de IA fueron el objetivo del segundo módulo. Se tomaron, para ello, casos reales y controvertidos, tales como:

COMPAS, un sistema de justicia predictiva, presenta sesgos. Y sesgos bastante notorios, por cierto.

Con todo, COMPAS no es el único veredicto sesgado, que toma decisiones con base en una economía del bien y del mal repleta de inequidades. A quienes se les ocurre diseñar un sistema de justicia predictiva deben asegurarse de que dicho sistema sea:

- Válido: que lo que mide sea lo que se supone que debe medir.
- No sesgado en sus predicciones: no prediga como muy probable que un negro o una negra va a cometer crímenes en el futuro cuando tiene como base que su economía, en general, está signada por una preocupación baja por el bien y por un precio alto por el mal. Es decir, problemas sociales que afectan comunidades enteras.

Sistema de contratación de Amazon que discriminaba por género. Microsoft Tay Bot es un chatbot que fue controlado para crear discursos ofensivos. Vehículos sin conductor (Uber) fue para examinar problemas de hacernos responsables y asuntos de moralidad cuando se trata de decisiones entre la vida y la muerte.

Estos casos fueron discutidos en actividades grupales. En ellas se llevó a cabo no solo un análisis del problema ético en cuestión, sino también una identificación clara de los afectados. Además, en estas actividades se llevaron a cabo propuestas de solución que, ciertamente, permiten la aplicación de algunos de los principios que hemos aprendido y revisado a lo largo de nuestro curso.

Fase 4: Evaluación y Síntesis

Cada uno de los módulos concluyó con un examen sobre lo que se había aprendido, el cual era de tres tipos: para los que eran principiantes, para los que estaban en un nivel intermedio y para los que ya estaban en un nivel avanzado. Las actividades incluyeron la elaboración de unos informes de auditoría que eran simulados, pero que debían tener una estructura formal (y, por supuesto, con contenido): un resumen ejecutivo, una lista de los riesgos que se podían presentar (y que estaban presentados), una lista de los impactos que esos riesgos podían tener y unas recomendaciones.

RESULTADOS

El desarrollo e implementación de los módulos de auditoría y ética en inteligencia artificial ofrecieron resultados significativos en cuanto a la comprensión y la aplicación de conceptos clave, tanto en el ámbito técnico como en el ético. En el ámbito técnico, los participantes concentraron sus esfuerzos en desglosar y aprehender las fases de la auditoría de TI:

- Planificación
- Evaluación
- Recolección de evidencia
- Informe final

Estas actividades fueron llevadas a cabo por los participantes en ese mismo orden, haciendo uso de algunas herramientas que probablemente son nuevas para ellos.

Las actividades prácticas permitieron simular situaciones del mundo real donde se evaluaron los controles internos y se documentaron las vulnerabilidades. Contribuyeron a desarrollar habilidades para la elaboración de auditorías, formulando recomendaciones correctivas con base en lo que se había encontrado.

Esto en cierto modo fue un ensayo para la etapa de prácticas en el programa.

En el terreno de lo ético, los estudiantes diseccionaron los sistemas de Inteligencia Artificial por el uso de los cuales se ven involucrados en dilemas morales. Y, como es de suponer, trabajaron sobre todo a partir de elucidar situaciones hipotéticas. Pero el panorama que les dibujaron las elucubraciones no pasó de ser apocalíptico. Más bien, la conclusión a la que llegaron fue que no hay

solución de problemas que valga si los algoritmos no son justos, y si, además, como demuestra la experiencia, aceptan el sesgo como parte de su día a día [14]. El reconocimiento de que es un problema de justicia lo convierte, sin duda, en un problema ético.

En resumen, los resultados muestran que la combinación de auditoría técnica y reflexión ética potencia nuestras capacidades de diseñar, implementar y evaluar soluciones digitales que sean responsables, seguras y, sobre todo, alineadas con el bienestar social.

CONCLUSIÓN

Este proyecto demostró que la unión de la auditoría de TI con los principios éticos de la IA es pertinente y necesaria, sobre todo en el contexto digital actual.

Los participantes recibieron formación técnica en procesos de auditoría y lograron adquirir competencias fundamentales para: Evaluar infraestructuras, tecnológicas, identificar vulnerabilidades, Proponer medidas correctivas, cumplir con estándares que aseguran la ciberseguridad de la organización.

Al mismo tiempo, analizar los riesgos éticos vinculados a la utilización de sistemas inteligentes hace visible lo que puede ocurrir (y ya está ocurriendo) cuando decisiones automatizadas mandan un mensaje de que esta o aquella conducta es la correcta, sin que alguien haya pensado competentemente y de antemano en qué es lo justo [15]. Y cuando uno se para a pensar al respecto, surgen casos con observación directa, que son el mejor argumento, para decir que la falta de supervisión ética puede comprometer la confianza pública en la tecnología.

Los aprendizajes más significativos destacan que el desarrollo de la tecnología no puede imaginarse solamente desde un prisma de eficiencia técnica, sino que debe estar orientado al bien común [16]. Capacitar a los futuros profesionales en la evaluación de los sistemas digitales desde dos perspectivas, la técnica y la ética, es una prioridad a fin de garantizar entornos digitales más seguros, equitativos y sostenibles. Como trabajo futuro, se propone fortalecer la conexión entre teoría y práctica mediante experiencias aplicadas y extender el alcance de los módulos con certificaciones, evaluaciones de impacto y una plataforma digital que garantice la actualización continua del aprendizaje.

Para concluir, la convergencia de auditoría y ética tecnológica es esencial para afrontar los retos actuales y anticiparse a los que planteará el futuro digital, y esto desde una perspectiva que no sólo sea crítica y preventiva, sino también humanista.

REFERENCIAS

[1] A. Jobin, M. Ienca y E. Vayena, "The global landscape of AI ethics guidelines," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 389–399, 2019.

[2] UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence* [En línea]. Disponible: https://unesdoc.unesco.org/ark:/48223/pf0000380455.

[3] Comisión Europea, *Directrices éticas para una IA fiable*, 2019. Disponible en: https://digital-strategy.ec.europa.eu/es/library/ethics-guidelines-trustworthy-ai.

[4] S. Russell y P. Norvig, *Artificial Intelligence: A Modern Approach*, 4ª ed., Pearson, 2021.

[5] OECD. (2019). *Artificial intelligence* [En línea]. Disponible: https://www.oecd.org/en/topics/artificial-intelligence.html.

[6] European Union Agency for Cybersecurity (ENISA). (2020). *Artificial Intelligence Cybersecurity Challenges* [En línea]. Disponible en: https://www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges.

[7] CISA. (2023). *Cybersecurity and Infrastructure Security Agency* [En línea]. Disponible en: https://www.cisa.gov/.

[8] L. Floridi y J. Cowls, "A Unified Framework of Five Principles for AI in Society," en *Harvard Data Science Review*, vol. 1, no. 1, 2019. DOI: https://doi.org/10.1162/99608f92.8cd550d1.

[9] IEEE Standards Association, *Ethically aligned design: A vision for prioritizing human wellbeing with artificial intelligence and autonomous systems*, IEEE, 2020.

[10] National Institute of Standards and Technology (NIST). (2022). *AI Risk Management Framework* [En línea].

Disponible: https://www.nist.gov/itl/ai-risk-management-framework.

[11] InvGate. (s. f.). *InvGate: Assest Management* [En línea]. Disponible: https://invgate.com/es/asset-management.

[12] Netwrix. (s. f.). *Netwrix*, [En línea]. Disponible: https://www.netwrix.com.

[13] LogicGate. (s. f.). *LogicGate* [En línea]. Disponible: https://www.logicgate.com/.

[14] B. Mittelstadt, “Principles alone cannot guarantee ethical AI,” en *Nature Machine Intelligence*, vol. 1, no. 11, pp. 501–507, 2019.

[15] B. Goodman y S. Flaxman, “European Union regulations on algorithmic decision-making and a 'right to explanation',” en *AI Magazine*, vol. 38, no. 3, pp. 50–57, 2017.

[16] Center for Humane Technology. (2021). *The AI Dilemma Steering towards a more humane Future* [En línea]. Disponible: https://www.humanetech.com/.