

Inteligencia artificial: hacia una gobernanza ética

Artificial intelligence: towards an ethical governance

Alfredo Estrada Merino¹

© El autor. Artículo de acceso abierto,
distribuido bajo los términos de la Licencia
Creative Commons Atribución 4.0 Internacional.



DOI: <https://doi.org/10.20453/ah.v68i2.7505>

RESUMEN

El avance acelerado de la inteligencia artificial (IA) en América Latina ha redefinido las dinámicas institucionales y ha expuesto la urgencia de consolidar marcos robustos de gobernanza ética capaces de equilibrar innovación, regulación y protección de derechos. Este artículo analiza los fundamentos conceptuales de la gobernanza ética de la IA, los avances regionales y la posición actual del Perú, integrando evidencia reciente de organismos multilaterales y estudios académicos. A partir de este examen, se identifican brechas estructurales en capacidades institucionales, infraestructura digital y mecanismos de supervisión algorítmica, así como oportunidades para fortalecer la transparencia, la responsabilidad pública y la cooperación regional. El análisis concluye destacando la necesidad de construir ecosistemas de IA confiables que articulen estándares éticos, capacidades técnicas y marcos normativos coherentes, los cuales constituyan la base para una agenda futura de investigación y acción pública.

Palabras claves: gobernanza de IA; ética y tecnología; políticas públicas; América Latina.

ABSTRACT

The accelerated advancement of artificial intelligence (AI) in Latin America has redefined institutional

dynamics and highlighted the urgency of establishing robust ethical governance frameworks capable of balancing innovation, regulation, and the protection of rights. This article analyzes the conceptual foundations of ethical AI governance, regional progress, and the current situation in Peru, integrating recent evidence from multilateral organizations and academic studies. Based on this examination, structural gaps are identified in institutional capacities, digital infrastructure, and algorithmic oversight mechanisms, as well as opportunities to strengthen transparency, public accountability, and regional cooperation. The analysis concludes by emphasizing the need to build trustworthy AI ecosystems that integrate ethical standards, technical capabilities, and coherent regulatory frameworks, which constitute the basis for a future agenda for research and public action.

Keywords: AI governance; ethics and technology; public policy; Latin America.

HACIA UN ABORDAJE CONCEPTUAL DE LA GOBERNANZA DE LA INTELIGENCIA ARTIFICIAL

El avance de la inteligencia artificial (IA) se desarrolla en un contexto marcado por una combinación de marcos regulatorios y éticos, desafíos estructurales e infraestructurales, un creciente impulso de políticas de datos, la persistente brecha digital y una continua

¹ Universidad de Ingeniería y Tecnología (UTEC). ORCID: 0000-0002-6488-8139

redefinición del papel del sector público frente a esta tecnología. Su rápida expansión ha desplazado el debate tecnológico hacia el ámbito de la gobernanza.

En menos de una década, la IA ha pasado de ser un instrumento de apoyo operativo a convertirse en un motor estructural que transforma instituciones, mercados laborales, decisiones públicas y dinámicas sociales. Este cambio obliga a transitar hacia una nueva forma de entender cómo se toman decisiones, quién las supervisa y bajo qué principios éticos se implementan sistemas inteligentes que influyen en la vida ciudadana. Este proceso requiere una etapa cognitiva inicial basada en una mirada conceptual e interdisciplinaria, que articule transparencia, responsabilidad, supervisión humana y mecanismos de validación continua para reducir riesgos y proteger derechos.

La funcionalidad de un abordaje conceptual de la gobernanza ética de la IA consiste en construir un marco de referencia que asegure que las tecnologías

incorpore, además de los elementos mencionados, valores como la explicabilidad y la sostenibilidad a lo largo del ciclo de vida de la IA. Los autores presentan el caso europeo como referencia para construir una taxonomía estructurada en torno a tres niveles de riesgo e introducen nociones relevantes como la gobernanza adaptativa y la ética desde el diseño. Por su parte, desde una perspectiva de gobernanza pública, Criado (2021) define la gobernanza de la IA como el conjunto de principios, estructuras, capacidades institucionales y mecanismos de supervisión necesarios para gestionar sistemas algorítmicos durante su diseño, implementación y uso colectivo, considerando los valores públicos bajo los cuales deben operar.

Desde una perspectiva sociotécnica, la gobernanza de la inteligencia artificial exige considerar tres dimensiones interrelacionadas que, en conjunto, permiten comprender la complejidad del despliegue de sistemas algorítmicos en contextos institucionales y sociales. La primera es la dimensión política, en

la que los Gobiernos establecen estrategias nacionales, marcos normativos, principios éticos y reglas de responsabilidad. La segunda es la dimensión organizacional, donde empresas e instituciones integran algoritmos en sus modelos de operación con el fin de mejorar servicios, fortalecer procesos internos y aumentar la productividad. La tercera es la dimensión humana,

en la cual las personas interactúan directamente con estos sistemas inteligentes y experimentan sus efectos en términos de discrecionalidad, transparencia, acceso y confianza. Estas tres dimensiones, presentadas de forma resumida en la tabla 1, conforman un marco analítico que permite conceptualizar y entender cómo la IA transforma tanto la toma de decisiones organizacionales como la experiencia ciudadana. Según Criado (2021), solo cuando estos niveles operan de manera estratégica y alineada es posible desarrollar un modelo de gobernanza ética capaz de asegurar integridad, reducir riesgos y ofrecer protección frente a los sesgos algorítmicos, especialmente en entornos donde la adopción tecnológica avanza con rapidez y requiere una supervisión pública y privada sólida.



Esta brecha evidencia **la urgencia de consolidar estándares éticos**, marcos de supervisión algorítmica y principios claros de responsabilidad pública como pilares para el despliegue confiable de la IA en el país.



se desarrollen alineadas con los valores sociales y los principios éticos. Un enfoque conceptual permite comprender y atender la complejidad del fenómeno, adaptándolo a los contextos legales, económicos y culturales de cada comunidad. Los pilares que sustentan esta construcción conceptual son la equidad, la transparencia, la responsabilidad y la privacidad. Articulados con la gestión regulatoria, el relacionamiento con las partes interesadas y la evaluación constante de los impactos, dichos pilares contribuyen a un enfoque de gobernanza más objetivo y práctico.

En esta línea de pensamiento, Ismail y Ahmad (2025) identifican la necesidad de un marco conceptual que

Tabla 1. Dimensiones sociotécnicas de la gobernanza de la inteligencia artificial

Dimensión	Descripción	Actores principales	Función clave
Política	Establece estrategias nacionales, marcos normativos, principios éticos y reglas de responsabilidad para orientar el uso de la IA.	Gobierno central, legisladores y organismos reguladores.	Definir estándares, garantizar derechos, establecer límites y responsabilidades.
Organizacional	Implementa algoritmos para mejorar gestión interna, servicios, productividad y toma de decisiones.	Empresas, instituciones públicas, directivos y áreas tecnológicas.	Integrar la IA en procesos, asegurar eficiencia y gestionar riesgos operativos.
Humana	Interacción de personas con sistemas inteligentes, experimentando efectos en acceso, discrecionalidad y confianza.	Ciudadanos, usuarios finales, población económicamente activa, estudiantes y empresarios.	Asegurar supervisión humana, confianza y accesibilidad equitativa.

Fuente: Elaboración propia basada en Criado (2021).

El concepto implica construir la gestión ética de la inteligencia artificial desde un enfoque colectivo y participativo. Saidin et al. (2025) sostienen que el desarrollo colaborativo con las partes interesadas define el rol verdaderamente responsable de la IA, al conciliar la innovación tecnológica con los valores fundamentales y los desafíos éticos que emergen en su aplicación. Bajo esta lógica, la gobernanza de la IA no se limita a configurar reglas técnicas o marcos normativos, sino que se consolida como un proceso social en el que intervienen múltiples actores que buscan garantizar que la tecnología responda al interés público. En suma, el concepto se sustenta en principios democráticos que deben traducirse en prácticas concretas de no discriminación, equidad, accesibilidad y confiabilidad, aspectos indispensables para mitigar riesgos y fortalecer la legitimidad institucional.

Desde el enfoque de los *stakeholders*, la gobernanza de la IA se convierte en un dispositivo de poder institucional

que articula decisiones públicas y privadas a partir de las necesidades, expectativas y derechos de la ciudadanía. Este enfoque reconoce que la tecnología no es neutral y que su impacto depende de las relaciones sociales y políticas en las que se inserta. Por ello, la participación informada y activa de la ciudadanía, las empresas, la academia y la sociedad civil es esencial para orientar la IA hacia resultados socialmente beneficiosos. Asimismo, este alcance involucra directamente al sector privado, que desempeña un rol determinante tanto por su capacidad de innovación como por su responsabilidad en garantizar que los sistemas algorítmicos operen con integridad y respeto hacia los usuarios.

Una perspectiva contemporánea incorpora el concepto de gobernanza ética corporativa, orientado a promover la IA responsable a partir del liderazgo organizacional, con el fin de consolidar culturas internas donde los aspectos éticos estén presentes a lo largo de todo el ciclo de vida de la tecnología, generando confianza

social y fomentando la innovación sostenible. En esta línea, Nangoy y Chan (2025) sostienen que el papel del liderazgo corporativo en la construcción de culturas éticas de IA requiere un análisis más profundo, especialmente porque aún persiste una brecha crítica en la investigación sobre gobernanza ética, confianza y rendimiento organizacional. Aunque áreas como la gobernanza, la equidad y la explicabilidad han avanzado en la literatura, los autores advierten que otras dimensiones esenciales como la privacidad, la seguridad y el florecimiento humano suelen quedar relegadas. Así las cosas, resulta clave ampliar la agenda investigativa académica para comprender de manera integral cómo el liderazgo influye en la creación de ecosistemas de IA que sean verdaderamente responsables, inclusivos y orientados al bienestar colectivo.

AVANCES REGIONALES EN GOBERNANZA ÉTICA DE LA INTELIGENCIA ARTIFICIAL

Los avances internacionales en materia de gobernanza de la inteligencia artificial permiten identificar prioridades comunes en los componentes de las estrategias nacionales de IA. En esa línea, OECD y Corporación Andina de

Fomento (2022) sintetizan los temas críticos que los Gobiernos están incorporando de manera sistemática para asegurar ecosistemas de IA orientados al impacto colectivo, abarcando desde enfoques éticos hasta el desarrollo de capacidades y la colaboración intersectorial. Este panorama permite comprender cómo distintos países de la región están configurando estrategias para el despliegue ético de la IA, construyendo marcos que combinan innovación, regulación y fortalecimiento institucional.

Como se aprecia en la tabla 2, la articulación entre gobernanza, ética, capacidades institucionales y gestión de datos constituye la base de las estrategias modernas de IA. Este marco integrado evidencia que los países que avanzan con mayor solidez son aquellos capaces de combinar lineamientos normativos con infraestructura tecnológica, mecanismos de experimentación y procesos claros de adopción. Un aspecto especialmente relevante es el énfasis en la colaboración intersectorial y la profesionalización de los servicios públicos vinculados, lo que confirma la interdependencia entre la IA, la regulación y la capacidad de gestión pública para asegurar una implementación responsable y efectiva.



En el caso peruano, si bien **los avances en infraestructura digital y gobernanza de datos representan bases relevantes**, estos resultan insuficientes si no se integran con lineamientos explícitos de transparencia, auditoría, trazabilidad y rendición de cuentas en el uso de sistemas inteligentes.



Tabla 2. Principales temas y componentes de las estrategias nacionales de IA

Tema clave	Descripción sintética
Enfoques confiables y éticos	Construcción de confianza pública, adaptación cultural de soluciones y orientación para el uso ético y transparente de la IA.
Gobernanza de IA y datos	Definición de marcos regulatorios, liderazgo institucional y responsabilidades para coordinar iniciativas de IA en el sector público.
Adopción de IA	Impulso al uso de IA en servicios ciudadanos, proyectos de alto impacto y mejora del desempeño gubernamental.
Adquisiciones de IA	Establecimiento de estándares técnicos y funcionales para la contratación de soluciones de IA en el sector público.
Capacidades del servicio civil	Formación y desarrollo de habilidades para el uso responsable de la IA, así como creación de espacios para compartir aprendizajes.
Colaboración intersectorial	Identificación de actores del ecosistema y construcción de alianzas público-privadas y transfronterizas.
Experimentación y pilotos	Fomento de innovación gubernamental mediante laboratorios, pilotos y entornos de prueba.
Infraestructura	Desarrollo de infraestructura avanzada, incluidos centros de datos y capacidades de supercomputación.
Datos abiertos y gobierno digital	Gestión estratégica de datos y apertura controlada para potenciar servicios y soluciones basadas en IA.

Fuente: Elaboración propia basada en OCDE y Corporación Andina de Fomento (2022).

La gobernanza ética de la inteligencia artificial en América Latina evidencia un progreso positivo, aunque profundamente heterogéneo. Los desafíos estructurales más relevantes tales como la innovación, la regulación y las capacidades institucionales demandan esfuerzos de sincronización mucho más consistentes para lograr una implementación real y efectiva del enfoque ético. En esta línea, Soto et al. (2025) advierten que, aunque nueve países cuentan con estrategias nacionales de IA, solo una

minoría ha avanzado en su actualización, presupuesto y mecanismos de evaluación. Esto genera un escenario regional fragmentado, donde la cantidad de países sin hoja de ruta duplica a los que ya la implementan. A pesar de que algunas naciones como Brasil, Chile o Uruguay han mostrado procesos más inclusivos, la participación ciudadana sigue siendo limitada, por lo que la transparencia, la legitimidad y la confianza pública en los sistemas algorítmicos se ven afectadas.

Esta heterogeneidad se refleja con claridad en el gráfico 1, donde las brechas en gobernanza resultan evidentes. Países como Brasil (92.9), Chile (83.2) y Uruguay (77.7) encabezan los avances regionales, mientras que otros como Guatemala (17.37), Bolivia (14.59) o Venezuela (12.46) exhiben niveles aún por desarrollar. Esta dispersión confirma que la región avanza con velocidades distintas, condicionando la posibilidad de construir ecosistemas de IA robustos y alineados con estándares éticos comunes.

La clasificación regional mostrada en el mismo gráfico evidencia que los países con gobernanza avanzada superan los 60 puntos al tomar decisiones inclusivas y contar con estrategias actualizadas y estructuras institucionales consolidadas. Los países con desempeño intermedio reflejan avances relevantes, aunque aún presentan debilidades persistentes en implementación y evaluación. Finalmente, aquellos ubicados en la categoría inicial muestran estructuras mínimas y amplias áreas por mejorar, lo que confirma la brecha regional en capacidades para sostener una IA verdaderamente ética. En síntesis, estos resultados permiten concluir que la gobernanza de la IA en América Latina depende tanto de la arquitectura institucional como de la capacidad real de implementación, constituyendo este el reto central para los próximos años.

Finalmente, cabe señalar que diversos instrumentos regionales analizan y mejoran constantemente la

composición de métricas para medir la formulación, la taxonomía, la escalabilidad y el impacto de la gobernanza ética de la IA, con el fin de observar la complejidad y amplitud del marco institucional que los países necesitan fortalecer para garantizar un ecosistema inteligente ético. Durante el año 2025, y en línea con lo trabajado por Soto et al. (2025), se han integrado nuevos indicadores de medición, tales como género, ciberseguridad, resiliencia digital y sostenibilidad, lo que evidencia que la lógica de gestión de riesgos se hace cada vez más presente en la medición regional de estos avances.

En el caso del Perú, las recomendaciones planteadas por Yeung (2020), orientadas a promover la innovación basada en sistemas de inteligencia artificial confiables tanto en países miembros de la OECD como en los que no lo son, han sido progresivamente incorporadas en la agenda nacional. Estas orientaciones remarcán la importancia de una administración responsable de la IA para fortalecer la confianza pública, un principio que Perú ha ido adoptando mediante lineamientos que buscan garantizar transparencia, seguridad y supervisión adecuada en el desarrollo tecnológico. De forma complementaria, Zambrano et al. (2025) sostienen que el país cuenta con iniciativas regulatorias emergentes que reflejan un compromiso creciente con la gobernanza ética; sin embargo, advierten la necesidad de consolidar un marco regulatorio coherente que equilibre la innovación con la protección de los derechos fundamentales y la mitigación de riesgos éticos. La

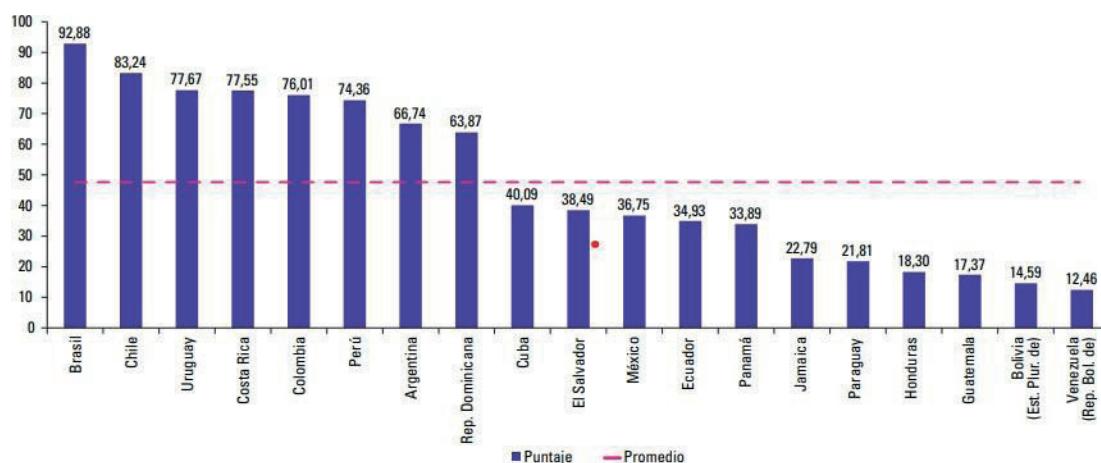


Figura 1. Puntaje de la dimensión de gobernanza según el Índice Latinoamericano de Inteligencia Artificial 2025 (Soto et al., 2025, p. 156).

ausencia de regulaciones regionales vinculantes genera fragmentación y dificulta la articulación normativa, por lo que fortalecer la cooperación regional y aprender de modelos exitosos resulta clave para avanzar hacia prácticas de IA verdaderamente responsables.

Por otro lado, estudios como el de Miranda-Hospital et al. (2024) evidencian el potencial de la IA para mejorar la transparencia y la eficiencia en la gestión de obras públicas, lo que se vincula indirectamente con la gobernanza ética al promover la rendición de cuentas y reducir espacios para la corrupción. Su trabajo pone de relieve que el despliegue efectivo de estas tecnologías depende de la infraestructura disponible, la claridad de las políticas gubernamentales y la asignación de recursos financieros adecuados. Asimismo, subraya que las consideraciones éticas deben integrarse de forma transversal en estos marcos habilitadores para asegurar que el uso de la IA en el sector público se desarrolle bajo principios de integridad, equidad y responsabilidad institucional.

El ritmo desigual del avance en gobernanza ética de la inteligencia artificial también se refleja en el caso peruano. Según Oxford Insights (2024), Perú obtuvo 47.3 puntos en el Índice de Preparación Gubernamental para la IA, superando el promedio regional de 41.1, pero aún distante de líderes globales como Singapur (85.4) y el Reino Unido (82.2). Esta brecha evidencia la urgencia de consolidar estándares éticos, marcos de supervisión algorítmica y principios claros de responsabilidad pública como pilares para el despliegue confiable de la IA en el país.

Asimismo, el índice señala que solo el 27 % de los países evaluados en la región cuentan con mecanismos formales para evaluar riesgos algorítmicos, lo que dificulta la adopción de prácticas éticas sólidas. En el caso peruano, si bien los avances en infraestructura digital y gobernanza de datos representan bases relevantes, estos resultan insuficientes si no se integran con lineamientos explícitos de transparencia, auditoría, trazabilidad y rendición de cuentas en el uso de sistemas inteligentes. En conjunto, los datos de 2024 confirmán que la consolidación de una gobernanza ética de la IA en América Latina —y particularmente en Perú— dependerá de la capacidad de

los Estados para equilibrar marcos regulatorios coherentes con mecanismos reales de implementación, cerrando finalmente la brecha entre la intención normativa y los resultados efectivos.

A MANERA DE CONCLUSIÓN

La revisión conceptual demuestra que la gobernanza ética de la inteligencia artificial requiere una arquitectura integrada que articule valores, principios y capacidades institucionales. La combinación de enfoques sociotécnicos, marcos éticos y participación de actores públicos y privados confirma que la IA no puede gestionarse únicamente desde la regulación, sino desde un proceso continuo de alineamiento normativo y operativo.

La gobernanza ética de la IA en América Latina mantiene brechas estructurales profundas, lo cual demuestra que la región aún no converge hacia un estándar mínimo común de supervisión algorítmica, un desafío crítico que condiciona la coherencia regulatoria y limita la capacidad de respuesta ante riesgos emergentes.

La región avanza hacia un modelo de gobernanza basado en gestión de riesgos, incorporando indicadores nuevos como género, ciberseguridad, resiliencia digital y sostenibilidad. Esta ampliación metodológica señala una transición hacia métricas más integrales que permitirán monitorear el impacto real de la IA y abrir un espacio para análisis comparativos en futuras investigaciones.

Pocos países de la región cuentan con mecanismos formales de evaluación algorítmica, lo que constituye la principal debilidad para consolidar prácticas éticas sólidas. En Perú, esta ausencia representa un vacío crítico que demanda lineamientos de transparencia, auditoría y trazabilidad para orientar la supervisión responsable de sistemas inteligentes. Aunque el país registra avances moderados, estos resultan aún insuficientes frente a los desafíos emergentes. De cara al futuro, será indispensable acelerar la adopción de marcos de evaluación de riesgos y fortalecer los mecanismos de responsabilidad pública para sostener un ecosistema de IA realmente ético.

REFERENCIAS

- Criado, J. I. (2021). Inteligencia artificial: Madurez tecnológica, adopción e innovaciones en la gestión pública. En Centro Latinoamericano de Administración para el Desarrollo (CLAD) (ed.), *Inteligencia artificial y ética en la gestión pública* (pp. 11-44). CLAD. https://issuu.com/clad_org/docs/libro_7_inteligencia_artificial_y_tica_en_la_gest
- Ismail, O. H. y Ahmad, N. (2025). Ethical and Governance Frameworks for Artificial Intelligence: A Systematic Literature Review. *International Journal of Interactive Mobile Technologies (ijIM)*, 19(14), 121-136. <https://doi.org/10.3991/ijim.v19i14.56981>
- Miranda-Hospital, O., Villodre, J., Valle-Cruz, D. y Angulo-Sánchez, K. (2024). Exploring the impact of artificial intelligence on the transparency and rationality of Peruvian public works: perceptions, expectations, challenges and opportunities. *Transforming Government: People, Process and Policy*. <https://doi.org/10.1108/tg-07-2024-0167>
- Nangoy, A. A. y Chan, P. (2025). Toward Ethical AI: Strategies for Responsible AI Governance. *Journal of Business and Management Studies*, 7(5), 145-155. <https://doi.org/10.32996/jbms.2025.7.5.13>
- OECD y Corporación Andina de Fomento (CAF) (2022). *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*. OECD Public Governance Reviews; OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/03/the-strategic-and-responsible-use-of-artificial-intelligence-in-the-public-sector-of-latin-america-and-the-caribbean_17c90e5e/1f334543-en.pdf
- Oxford Insights (2024). *Government AI Readiness Index 2024*. Oxford Insights. <https://oxfordinsights.com/ai-readiness/ai-readiness-index/?#download-reports>
- Saidin, S. S., Ahmad, N. y Aziz, S. (2025). Ethical Governance of Artificial Intelligence in Organizations: Challenges and Strategies. En E. Ahmed, A. Babar, A. Samad, R. Ahmed y G. Beydoun (eds.), *Strengthening Human Relations in Organizations With AI* (pp. 59-82). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-6507-6.ch003>
- Soto, Á., Durán, R., Moreno, A., Adasme, S., Rovira, S., Jordán, V. y Poveda, L. (coords.) (2025). Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025. *Documentos de Proyectos* (LC/TS.2025/68). Comisión Económica para América Latina y el Caribe y Centro Nacional de Inteligencia Artificial. https://indicelatam.cl/wp-content/uploads/2025/10/Documento_ILIA_2025.pdf
- Yeung, K. (2020). Recommendation of the Council on Artificial Intelligence (OECD). *International Legal Materials*, 59(1), 27-34. <https://doi.org/10.1017/ILM.2020.5>
- Zambrano, O., Talavera-Ruz, M. y Niño Peñalosa, J. A. (2025). Desafíos regulatorios de la inteligencia artificial en la comunidad andina y oportunidades para el emprendimiento digital. *Ciencia UNEMI*, 18(49), 122-136. <https://doi.org/10.29076/issn.2528-7737vol18iss49.2025pp122-136p>