



Al contestar refiérase a: **ID-131486**

AS-ATIC-0010-2025

11 de febrero de 2025

Máster

Robert Picado Mora, subgerente a.i.

SECRETARIO CONSEJO TECNOLÓGICO INSTITUCIONAL

DIRECCIÓN TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIONES – 1150

Estimado señor:

ASUNTO: Oficio de asesoría respecto al uso responsable de la Inteligencia Artificial en el sector salud y pensiones.

En cumplimiento de las actividades preventivas consignadas en el Plan Anual Operativo para el período 2025 y con fundamento en los artículos N.º 21 y 22 de la Ley General de Control Interno, específicamente en su rol de asesor, esta Auditoría emite el siguiente oficio en torno al uso responsable de la Inteligencia Artificial (IA) en el sector salud, con el objetivo de que, como institución, se considere el aprovechamiento ético y responsable de este tipo de tecnologías.

Lo anterior, concordante a lo definido en el Plan Estratégico Institucional 2023-2033, en el objetivo 3, sobre la importancia del uso de tecnología y aprovechamiento de los datos, que cita:

“Potenciar el bienestar de las personas usuarias, mediante el incremento en el uso de las tecnologías, la innovación e investigación para desarrollar soluciones más eficientes en la prestación de los servicios de salud y pensiones”.

Y específicamente, la línea de acción 3.f, sobre el uso y aprovechamiento de la analítica de datos:

“Incorporación de la analítica de datos (diagnóstica, predictiva, prescriptiva) y modelos de información mediante herramientas como big data, BI, IOT, entre otras, que fomente el pensamiento científico y la mejora de los procesos institucionales, la transparencia, la minimización de los riesgos, fraude y corrupción en la gestión institucional”.

Con la globalización y el avance tecnológico, las instituciones se enfrentan a retos y oportunidades sin precedentes, y uno de los aspectos críticos en este proceso es lo referente a la gestión de los datos y la implementación de tecnologías emergentes.

Es un hecho que las tecnologías están ayudando cada vez más a las organizaciones a interconectar datos e identificar y dar respuesta a riesgos oportunamente, generando valor significativo a la toma de decisiones.

Al respecto, de acuerdo con información recopilada, una de las tendencias que ha transformado los procesos ha sido la Inteligencia Artificial (IA), y la CCSS no se exime del aprovechamiento de estas tecnologías, por lo que se emite una serie de consideraciones del uso de esta tecnología para su valoración.



Consideraciones de la inteligencia artificial en el sector salud y pensiones

La inteligencia artificial (IA) es una tecnología que permite a las máquinas simular habilidades humanas como aprender, razonar, resolver problemas y tomar decisiones. Utiliza algoritmos y datos para automatizar tareas que normalmente requieren inteligencia humana, como el reconocimiento de voz, análisis de datos y conducción autónoma. OpenAI. (2025).

En el sector salud, la IA está transformando la atención médica, ya que ofrece herramientas avanzadas para el diagnóstico temprano de enfermedades, la personalización de tratamientos y la optimización de procesos clínicos; ha apoyado además con el análisis de grandes volúmenes de datos permitiendo identificar patrones complejos que ayudan a los profesionales a tomar decisiones más rápidas, precisas y documentadas, no solo busca mejorar la eficiencia en los procesos, si no que vislumbra un futuro donde la atención médica y servicios sean más accesibles, eficientes y adaptados a las necesidades de cada paciente.

En el sector pensiones, permite crear algoritmos para análisis de datos que colaboren con la gestión de carteras, maximizar rendimientos, minimizar riesgos, identificar patrones inusuales, predicciones de tendencias y automatización de procesos, entre otros que permitan toma de decisiones estratégicas.

En ese sentido, la Organización Mundial de la Salud (2021), reconoce que la IA es muy prometedora para la práctica de la salud pública y la medicina, y en el documento de las Orientaciones “Ética y Gobernanza de la Inteligencia Artificial en el Ámbito de Salud”, propone una serie de principios a fin de aprovechar plenamente sus beneficios, enfocados en abordar los problemas éticos que conlleva para los sistemas de atención de la salud, los profesionales sanitarios y los beneficiarios de la salud pública.

Los principios éticos fundamentales de la OMS, se utilicen como base para que los gobiernos, los desarrolladores de tecnologías, las empresas, la sociedad civil y las organizaciones intergubernamentales adopten enfoques éticos para el uso adecuado de la IA en el ámbito de la salud.

Los seis principios se citan a continuación:

- 1. Proteger la autonomía humana:** La utilización de la IA puede conducir a situaciones en las que la facultad de adoptar decisiones podría transferirse a las máquinas. El principio de autonomía requiere que el empleo de la IA u otros sistemas informáticos no socave la autonomía humana. En el contexto de la atención de la salud, esto significa que las personas deben mantener el control sobre los sistemas de atención de la salud y las decisiones médicas. El respeto de la autonomía humana también conlleva deberes conexos a fin de que los proveedores cuenten con la información necesaria para hacer un uso seguro y eficaz de los sistemas de IA y de que las personas entiendan el papel de dichos sistemas en la atención de la salud. También requiere la protección de la privacidad y la confidencialidad, así como la obtención de un consentimiento informado válido mediante marcos jurídicos adecuados para la protección de datos.
- 2. Promover el bienestar y la seguridad de las personas y el interés público.** Las tecnologías de IA no deben perjudicar a las personas. Los diseñadores de tecnologías de IA deberían



cumplir las prescripciones normativas en materia de seguridad, precisión y eficacia para indicaciones o usos bien definidos. Es preciso establecer medidas de control de la calidad en la práctica y de mejora de la calidad en la utilización de la IA a lo largo del tiempo. La prevención de daños requiere que la IA no ocasione perjuicios mentales o físicos que podrían evitarse mediante el uso de otras prácticas o enfoques alternativos.

- 3. Garantizar la transparencia, la claridad y la inteligibilidad.** Las tecnologías de IA deben ser inteligibles o comprensibles para los desarrolladores, profesionales de la salud, pacientes, usuarios y reguladores. Dos enfoques generales para la inteligibilidad consisten en mejorar la transparencia de la tecnología de IA y hacer que esta sea explicable. La transparencia requiere que se publique o documente información suficiente antes del diseño o despliegue de una tecnología de IA, y que dicha información facilite consultas y debates públicos significativos sobre la manera de diseñar la tecnología y sobre cómo debe o no debe utilizarse. Las tecnologías de IA deben ser explicables conforme a la capacidad de las personas a quienes se dirige esa explicación.
- 4. Promover la responsabilidad y la rendición de cuentas.** Las personas necesitan una especificación clara y transparente de las tareas que los sistemas pueden realizar y las condiciones en las que pueden alcanzar el rendimiento deseado. Aunque las tecnologías de IA realizan tareas específicas, incumbe a las partes interesadas velar por que esas tareas puedan llevarse a cabo y porque personas debidamente capacitadas utilicen la IA en condiciones adecuadas. La responsabilidad puede asegurarse mediante la aplicación de la «garantía humana», que implica la evaluación por parte de los pacientes y los médicos en el desarrollo y despliegue de las tecnologías de IA. La garantía humana requiere la aplicación de los principios regulatorios en toda la extensión del algoritmo mediante el establecimiento de puntos de supervisión humana. Si surgen problemas con una tecnología de IA, se precisa la rendición de cuentas. Debe disponerse de mecanismos apropiados para la investigación y la reparación en beneficio de las personas y los grupos que se vean afectados negativamente por decisiones basadas en algoritmos.
- 5. Garantizar la inclusividad y la equidad.** La inclusividad requiere que la IA aplicada a la salud se conciba de manera que aliente la utilización y el acceso equitativos, amplios y apropiados en la mayor medida de lo posible, con independencia de la edad, el sexo, el género, los ingresos, la raza, el origen étnico, la orientación sexual, la capacidad u otras características amparadas por los códigos de derechos humanos. Las tecnologías de IA, como cualquier otra tecnología, deben compartirse en la mayor medida de lo posible, y deben estar disponibles no solo para cubrir las necesidades en los entornos de ingresos altos, sino también teniendo en cuenta los contextos, las capacidades y la diversidad de los países de ingresos medianos y bajos (PIMB). Las tecnologías de IA no deben codificar sesgos en detrimento de grupos identificables, especialmente si ya están marginados. El sesgo es una amenaza para la inclusividad y la equidad, ya que puede dar lugar a una desviación, a menudo arbitraria, en la igualdad de trato. Las tecnologías de IA deben minimizar las inevitables disparidades de poder que se producen entre los proveedores y los pacientes, entre las instancias normativas y las personas, y entre las empresas y los gobiernos que crean e implementan tecnologías de IA y quienes las utilizan o dependen de ellas. Las herramientas y los sistemas de IA deben ser objeto de vigilancia y evaluación para identificar efectos desproporcionados en grupos específicos de personas. Ninguna tecnología, sea de IA o de otro tipo, debe apoyar o empeorar las formas existentes de sesgo y discriminación.



- 6. Promover una IA con capacidad de respuesta y sostenible.** La capacidad de respuesta requiere que los diseñadores, desarrolladores y usuarios evalúen de forma continua, sistemática y transparente las aplicaciones de IA en situación real. Deben determinar si la IA responde de manera adecuada y apropiada y de acuerdo con las expectativas y los requisitos legítimos que se hayan comunicado. La capacidad de respuesta también requiere que las tecnologías de IA sean coherentes con una promoción más amplia de la sostenibilidad de los sistemas, los entornos y los lugares de trabajo en el ámbito de la salud. Los sistemas de IA deben diseñarse de modo que se reduzcan al mínimo los efectos ambientales y aumente la eficiencia energética. Es decir, la utilización de la IA debe ser coherente con las iniciativas mundiales para reducir el impacto humano en el medio ambiente, los ecosistemas y el clima terrestres. La sostenibilidad también requiere que los gobiernos y las empresas hagan frente a las perturbaciones previstas en los lugares de trabajo, en particular mediante la capacitación del personal de salud para que se familiarice con el uso de los sistemas de IA, así como a las posibles pérdidas de empleos a causa de la utilización de sistemas automatizados.

La CCSS no ha sido exenta de las bondades de la IA, por lo que es importante considerar la implementación de estas tecnologías con diseños y políticas basados en principios éticos, promoviendo la adopción de enfoques responsables para el uso adecuado de la IA en el ámbito de la salud y pensiones.

Uso de la Inteligencia Artificial Generativa (IAG)

Una de las tendencias que ha transformado los procesos, y ha acaparado la atención de los sectores, es la Inteligencia Artificial Generativa (IAG), estas son herramientas capaces de crear contenido nuevo a partir de algoritmos avanzados que se alimentan de datos, entre las herramientas más comunes están ChatGPT (OpenAI), Gemini (Google), y Copilot (Microsoft).

Esta tecnología ha revolucionado la forma de trabajo actual, se utilizan como catalizadores de creatividad, innovación, sin menoscabo de su uso en la automatización de tareas repetitivas, optimizan y mejoran procesos, y brindan apoyo en la toma de decisiones.

La IA generativa tiene el potencial de automatizar entre el 60% y el 70% de las actividades laborales actuales, y plantea un enorme potencial, sin embargo, las compañías y organizaciones centroamericanas que estén considerando implementar esta tecnología deben ser conscientes de una serie de riesgos y desafíos, según estudio de McKinsey & Company. (2024).

Entre los desafíos se destacan:

- 1. Movimientos en la fuerza laboral y cambios en los requerimientos de talento:** demanda en áreas como ingeniería de instrucciones (o “prompts”) o ingeniería de datos, trabajos que requieren nuevas habilidades actualmente escasas.
- 2. Disponibilidad y calidad de los datos y sistemas de información:** Los modelos de Gen AI requieren datos de alta calidad para producir resultados precisos y confiables. El estudio de McKinsey & Company, revela que, en América Central, las empresas aún enfrentan problemas de disponibilidad y calidad de los datos y de capacidad de almacenamiento. La falta de datos o de sistemas adecuados podría colocar a la región en una posición desventajosa y complicar la adopción de la IA generativa. En consecuencia, algunas compañías quizás tengan que realizar fuertes inversiones para acceder a conjuntos de datos diversos y representativos.



- 3. Implicaciones éticas, sesgos de modelos, confiabilidad y precisión, e infracción de la propiedad intelectual:** Los modelos de IA generativa pueden reflejar sesgos implícitos en los datos de entrenamiento, o divulgar inadvertidamente información privada contenida en ellos. El estudio realizado por McKinsey & Company, indica que las organizaciones tendrán que definir pautas para prevenir resultados sesgados que puedan afectar a usuarios y causar daños reputacionales. También deberán dictar normas para proteger la privacidad de las personas. Dado que la mayoría de las organizaciones de Centroamérica aún se encuentra en la fase inicial de sus modelos de gobierno de datos e IA, estas acciones cobran aún más importancia.

El uso de estas herramientas promueve un cambio y adaptación en la institución, y vale la pena que los entes competentes regulen su uso, no solo para maximizar sus beneficios, si no para minimizar posibles riesgos, en cuanto a la protección de los datos, ética y transparencia, supervisión de resultados, entre otros.

Propiedad intelectual

Por otro lado, y en cuanto a la transparencia y propiedad intelectual, no está demás mencionar lo que APA (American Psychological Association) en su 7ª edición ha establecido como pautas para citar herramientas de IA.

De acuerdo con las guías actualizadas, es necesario mencionar el uso de herramientas como ChatGPT o similares, cuando su contenido se incluye en un texto académico o profesional.

Indicando entre los aspectos a considerar:

- **Transparencia:** Mencionar el uso de ChatGPT o similares asegura que se mantiene la integridad académica y se da crédito a la herramienta.
- **Propiedad intelectual:** Aunque ChatGPT o similares no es una "persona", el contenido generado es un producto de OpenAI u otras organizaciones y debe reconocerse como tal.

Consideraciones del uso de IA en la analítica de datos

La institución ha realizado esfuerzos valiosos en la digitalización de sus procesos, atesorando una importante cantidad de datos, generalmente descentralizados y con un crecimiento exponencial en sistemas como EDUS, SICERE, recientemente ERP y otros.

La IA tiene el potencial de convertir los datos en inteligencia de negocio, por lo que es vital e indispensable una fuente depurada y certera para su efectividad, es decir, el rendimiento de la IA depende directamente de la calidad de los datos.

Expertos en el tema coinciden que, los principales retos para la efectividad de la IA es considerar eliminar los silos de datos, democratizar el acceso a estos, y poner en práctica reglas de gobernanza para garantizar que todos los tipos y fuentes de datos actúen al unísono, independientemente de dónde residan, tratándose como una fuente unificada y cohesionada.



Por lo que las estrategias ligadas a la gobernanza de los datos en la institución, son relevantes para garantizar la calidad, seguridad y uso ético de la información.

“Muchas empresas tienen la visión de convertirse en una organización basada en datos, especialmente por la relevancia actual de la IA. Pero, para ello, deben establecer unos fundamentos claros”, afirma Andy Longworth, coordinador de plataformas de datos de los servicios de asesoría profesional de Hewlett Packard Enterprise, en un estudio que demuestra que superar las complejidades de la gestión de datos puede reportar grandes beneficios, indicando en el estudio que, el enfoque ideal comienza con un proceso exhaustivo de detección de datos para averiguar exactamente qué datos se tienen y dónde se encuentran por parte de TI, para continuar posteriormente a facilitar el acceso de las partes interesadas a esos datos, paso fundamental para desbloquear su valor (Hewlett Packard Enterprise, 2023).

Normativa y Marcos de Referencias de la IA en Costa Rica

Actualmente, el país no dispone de legislación que regule el uso de la IA. No obstante, se conocen algunas iniciativas legislativas sobre el tema, que se describen:

- Proyecto de Ley N° 23.771 “Ley de Regulación de la Inteligencia Artificial en Costa Rica”, busca brindar un marco normativo para el desarrollo, implementación y uso ético, seguro y sostenible de la IA en Costa Rica, entre los sectores que aborda está salud, presentado el 30 de mayo 2023.
- Proyecto de Ley N° 23.919 “Ley para la Promoción Responsable de la Inteligencia Artificial en Costa Rica, busca el resguardo de los derechos de las personas ante la IA.
- Proyecto de Ley N° 24.484 “Ley para la Implementación de Sistemas de Inteligencia Artificial”, busca dar las pautas a la administración pública para el manejo y la adopción de tecnologías de IA.

El Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT) por su parte, publicó la Estrategia Nacional de Inteligencia Artificial 2024-2027, documento fechado al 24 de octubre del 2024 y que brinda una visión estratégica, y que busca posicionar a nuestro país como líder en la adopción responsable de la inteligencia artificial y promover el uso, adopción y desarrollo de manera ética, segura y responsable, procurando maximizar los beneficios para la ciudadanía y minimizar cualquier posible daño a las personas.

La estrategia propone un marco de trabajo que se inspira en el trabajo realizado por Estados Unidos, la Unión Europea (UE), la OECD, el HAIP y propone un enfoque adaptado a la realidad del país, con el objetivo de gestionar eficazmente los riesgos de la IA y asegurar su contribución positiva al bienestar social y económico de Costa Rica.

Refiere además que, en Costa Rica se proyecta que la adopción de la IA podría impulsar el crecimiento del Producto Interno Bruto (PIB) para 2030, con una tasa de crecimiento anual de hasta el 7.8% (Microsoft, 2019) y en el sector público, ya que la IA tiene el potencial de optimizar la formulación, ejecución y evaluación de políticas públicas, mejorando la entrega de servicios y agilizando gestiones



internas de las instituciones estatales, como la atención de consultas, correspondencia y emisión de informes. Entre los ejemplos concretos incluyen la automatización de procesos en ministerios y la mejora de la calidad y eficiencia de los servicios de salud mediante plataformas como EDUS de la Caja Costarricense de Seguro Social (CCSS), que utiliza la IA para la prevención y gestión de enfermedades crónicas.

Es relevante conocer que, el gobierno a través del MICITT promueve acciones estratégicas que respalden la adopción de estándares de IA, incluidos principios como el Data Free Flow With Trust (DFFT), es fundamental para garantizar que las tecnologías de IA se desarrollen e implementen de manera ética y transparente, alineadas con las mejores prácticas globales. La participación en AIGO, GPAI y HAIP refuerza el rol de Costa Rica como catalizador en la adopción de lineamientos éticos y técnicos en la IA a nivel global, así como buscar siempre el bienestar de las personas.

El MICITT en su Estrategia Nacional de IA, asume siete principios rectores esenciales para guiar su desarrollo y aplicación de manera responsable, a saber:

1. Paz y dignidad humana

Al integrar la paz como un pilar clave, la estrategia busca asegurar que el desarrollo y la aplicación de la IA se realicen de manera ética y responsable, promoviendo la inclusión y evitando conflictos.

2. Supervisión humana

El principio de supervisión humana se centra en mantener el control humano final sobre los sistemas que utilicen IA, asegurando que la responsabilidad ética y jurídica pueda atribuirse claramente a individuos o entidades legales en cualquier etapa del ciclo de vida de estos sistemas. Aunque los sistemas de IA pueden ser empleados por su eficiencia y capacidad para apoyar la toma de decisiones y la ejecución de tareas, la responsabilidad última y la rendición de cuentas siempre recaen en el ser humano.

3. Transparencia y explicabilidad

Los actores que desarrollen e implementen sistemas que utilicen IA deben proporcionar información comprensible y adecuada al contexto que explique tanto las capacidades como las limitaciones de la IA. Esto es esencial para permitir que las personas usuarias comprendan cómo y cuándo interactúan con una IA y cómo estas influyen en su entorno laboral y personal.

4. Equidad y no discriminación

Los sistemas que utilicen IA deberán diseñarse y operarse de manera que promuevan la inclusión y eviten perpetuar o crear nuevas formas de discriminación.

5. Responsabilidad

El principio de responsabilidad estipula que los actores involucrados en el ciclo de vida de los sistemas que utilicen IA deberán asumir la responsabilidad ética y jurídica de sus operaciones y efectos. Este principio asegura que todas las acciones y decisiones influenciadas por estos sistemas puedan ser atribuidas a personas físicas o jurídicas, garantizando así una rendición de cuentas clara y efectiva.



6. Sostenibilidad y bienestar

El principio de sostenibilidad y bienestar se fundamenta en la promoción y contribución activa al desarrollo sostenible en sus múltiples dimensiones: social, económica, ambiental y cultural.

7. Seguridad, ciberseguridad y protección de la información

La IA debe desarrollarse y aplicarse de manera cibersegura, protegiendo a las personas y al país. Es esencial identificar y mitigar los riesgos que la IA puede generar para garantizar la seguridad y confianza en estas tecnologías. Este principio enfatiza la necesidad de que los sistemas que utilicen IA sean robustos, seguros y confiables durante todo su ciclo de vida, asegurando que operen de manera adecuada tanto bajo condiciones normales de uso como en situaciones previsibles de mal uso o condiciones adversas.

Igualmente, promueve cinco principios transversales, a saber:

1. Enfoque de género

El enfoque de género debe ser un principio transversal en todas las fases del desarrollo y aplicación de la IA. Esto implica asegurar que las tecnologías no perpetúen ni amplifiquen las desigualdades de género existentes.

2. Inclusión y accesibilidad

La IA debe ser desarrollada y aplicada de manera que sea accesible e inclusiva para todas las personas, independientemente de su edad, género, origen étnico, capacidad económica, nivel educativo o discapacidad.

3. Protección de datos, propiedad intelectual y privacidad

La protección de datos, la privacidad y la propiedad intelectual deben ser consideraciones centrales en todas las aplicaciones de IA. Esto implica implementar fuertes medidas de seguridad y robustecer los marcos legales para garantizar que los datos personales sean manejados de manera segura y ética, y que los derechos de privacidad de las personas sean respetados y protegidos.

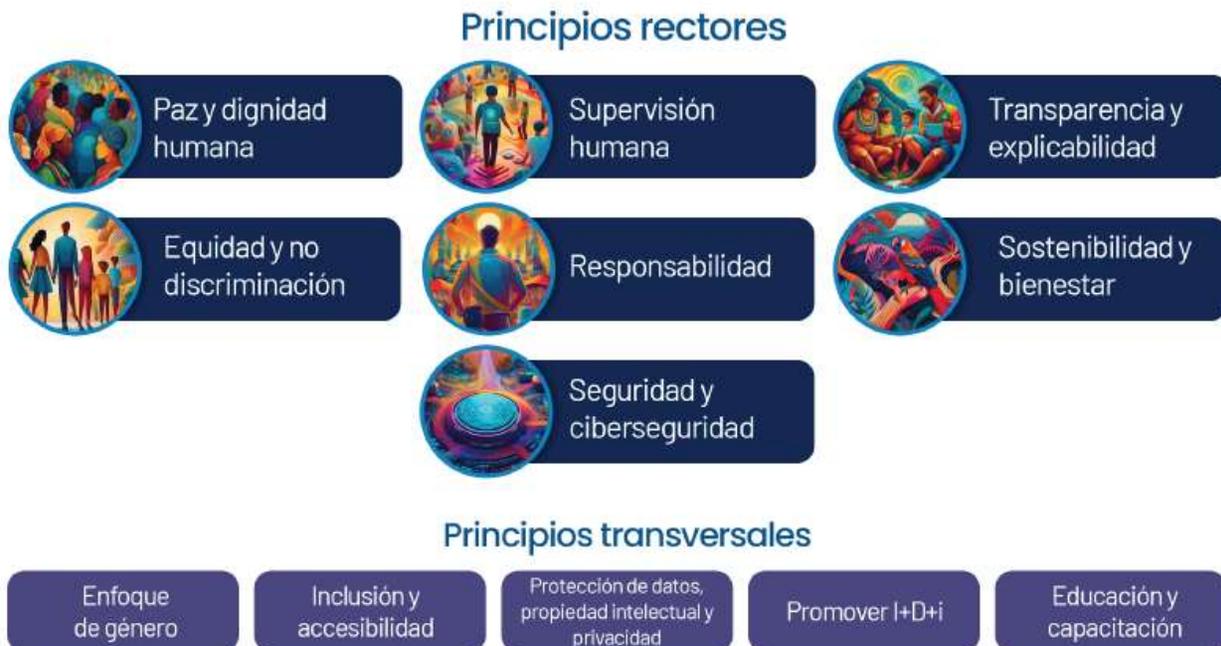
4. Promover la I+D+i

La investigación, el desarrollo y la innovación (I+D+i) en el campo de la IA deben ser fomentados de manera responsable, asegurando que los avances tecnológicos beneficien equitativamente a toda la sociedad, protegiendo siempre la vida humana.

5. Educación y capacitación

La educación y la capacitación son esenciales para garantizar que todas las personas tengan las habilidades necesarias para utilizar y beneficiarse de la IA.

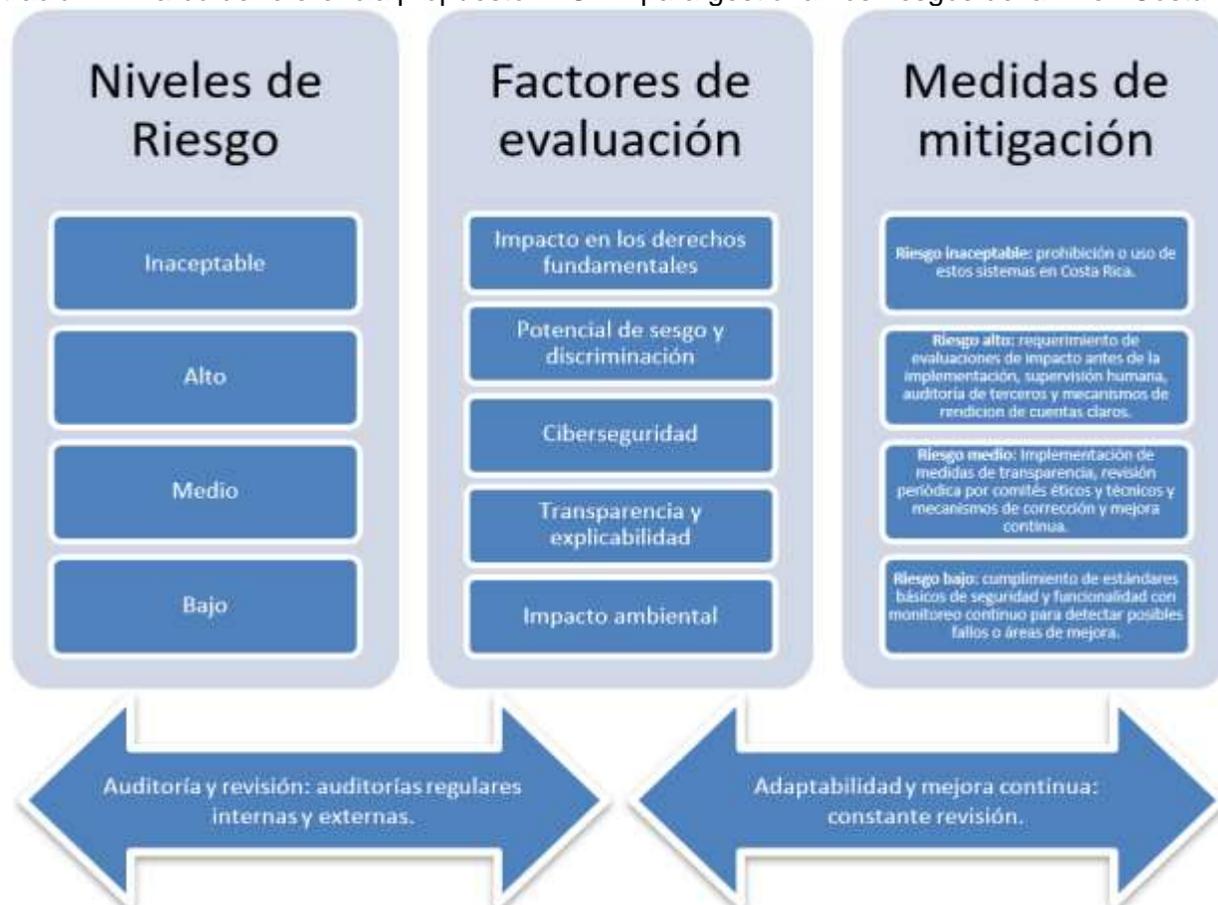
Ilustración 1. Principios rectores y transversales de la IA - MICITT



Fuente: Principios rectores y transversales de la Estrategia Nacional de Inteligencia Artificial Costa Rica 2024-2027 MICITT.

Finalmente, propone un marco de trabajo que permita identificar, evaluar y clasificar los riesgos asociados con los sistemas de IA, garantizando que su diseño, desarrollo y uso no comprometan los principios fundamentales de la sociedad costarricense, el cual se inspira en el trabajo realizado por Estados Unidos, la Unión Europea (UE), la OECD, el HAIP y propone un enfoque adaptado a la realidad del país, con el objetivo de gestionar eficazmente los riesgos de la IA y asegurar su contribución positiva al bienestar social y económico de Costa Rica, como se muestra en la ilustración 2 “Marco de referencia propuesto MICITT para gestionar los riesgos de la IA en Costa Rica”.

Ilustración 2. Marco de referencia propuesto MICITT para gestionar los riesgos de la IA en Costa Rica



Fuente: Elaboración propia con la información propuesta del marco de referencia de la Estrategia Nacional de Inteligencia Artificial Costa Rica 2024-2027 MICITT

Por su parte, las Normas Técnicas para la Gestión y el Control de las Tecnologías de Información del MICITT, en el apartado II, “Gestión de TI”, menciona lo siguiente:

“La institución debe implementar y mantener prácticas de gestión de las TI, que defina formalmente los siguientes componentes para la entrega de servicios al nivel de tecnologías de información en alineación con el marco estratégico y el modelo de arquitectura empresarial:

(...) 4. Investigación sobre tecnologías emergentes que permitan a través de su eventual incorporación, la innovación y mejora continua al nivel institucional para el logro de los objetivos y la entrega de valor público”.

Además, en el apartado VI “Calidad de los Procesos Tecnológicos”, señala:

“La institución debe implementar prácticas que permitan controlar los procesos organizacionales, posibilitando la mejora continua de productos y servicios, buscando



asegurar la satisfacción de las necesidades institucionales, manteniendo estándares de documentación de los lineamientos requeridos, esquemas para la medición del desempeño y control sobre la vigencia de las prácticas aplicables a los procesos.

Igualmente, debe generar servicios de TI de conformidad con los requerimientos de los usuarios con base en un enfoque de eficiencia y mejoramiento continuo de los procesos que habilitan la gestión de las tecnologías de información”.

Principales riesgos de la IA en Costa Rica según el MICITT y OMS

La Estrategia Nacional de IA Costra Rica 2024-2027 del MICITT, ofrece un listado de los riesgos críticos basados en investigaciones efectuadas, los que incluyen la propagación de desinformación y amenazas a los derechos humanos, y que en este oficio se enlistan como una base de conocimiento sobre el tema:

- **Impacto en el empleo:** La automatización impulsada por la IA puede desplazar significativamente la fuerza laboral en diversos sectores, lo que plantea la necesidad de políticas públicas precisas.
- **Impacto en el comportamiento humano:** La IA puede reconfigurar normas sociales y autonomía individual, especialmente cuando se utiliza para manipular comportamientos a través de sistemas de recomendación o publicidad dirigida.
- **Sesgo y discriminación:** Los datos de entrenamiento de la IA pueden contener prejuicios históricos o culturales, lo que puede llevar a la IA a automatizar y amplificar estos sesgos, resultando en discriminaciones en áreas críticas como contratación laboral, préstamos bancarios y servicios policiales.
- **Ciberataques:** Los sistemas de IA son susceptibles a ataques cibernéticos que pueden alterar su funcionamiento, con consecuencias potencialmente devastadoras en sistemas críticos como infraestructura de red eléctrica, sistemas financieros y redes de salud.
- **Datos personales y privacidad:** La IA requiere vastas cantidades de datos para su funcionamiento, lo que plantea serias preocupaciones sobre la privacidad y la seguridad de la información personal. El uso indebido o el acceso no autorizado a estos datos puede resultar en violaciones de la privacidad que afectan a individuos y grupos.
- **Toma de decisiones autónomas:** La capacidad de la IA para tomar decisiones sin intervención humana presenta dilemas éticos, especialmente en contextos donde estas decisiones pueden tener implicaciones significativas en la vida de las personas, como en la atención médica.
- **Uso militar y vigilancia:** El desarrollo de tecnologías de IA para uso militar y vigilancia genera debates éticos sobre la militarización de la IA y el potencial de abuso en la vigilancia masiva, planteando preguntas sobre la moralidad de la automatización del conflicto armado y las implicaciones para los derechos civiles.



- **Generación de contenido falso:** La IA puede generar contenido falso convincente, como videos *deepfake* y noticias falsas, utilizados para desinformar y manipular la opinión pública, erosionando la confianza en los medios y las instituciones democráticas.
- **Incidencia en decisiones de ciudadanía:** La IA puede influir significativamente en las decisiones políticas y sociales, desde la asignación de recursos en la administración pública hasta el diseño de políticas basadas en datos. Estas decisiones pueden estar motivadas por eficiencia algorítmica en lugar de consideraciones humanísticas, afectando negativamente la equidad y justicia social.
- **Dependencia tecnológica:** La creciente dependencia de sistemas automatizados y soluciones de IA puede llevar a una “des-skilling” en la fuerza laboral, donde las habilidades humanas se deterioran porque las máquinas realizan tareas complejas, aumentando la vulnerabilidad en caso de fallo de la IA.
- **Desigualdad en el acceso a la tecnología:** La implementación desigual de soluciones de IA puede ampliar la brecha digital, exacerbando las desigualdades existentes en educación, salud y oportunidades económicas entre diferentes poblaciones y países.
- **Impacto en el cambio climático:** El entrenamiento de modelos de IA es altamente intensivo en energía. Por ejemplo, un solo modelo de IA puede emitir más de 284 toneladas de CO₂, equivalente a cinco veces las emisiones de un automóvil promedio a lo largo de toda su vida útil. Esto subraya la urgencia de desarrollar IA de manera sostenible, priorizando la eficiencia energética y la adopción de energías renovables.
- **Impacto en la creatividad y la innovación:** Aunque la IA puede mejorar la eficiencia en varios procesos, existe el temor de que pueda limitar el espacio para la creatividad humana y la innovación al priorizar soluciones basadas en patrones y datos históricos.
- **Problemas de interpretación y contextos inapropiados:** Los sistemas de IA pueden fallar en interpretar correctamente el contexto social o cultural de sus aplicaciones, lo que puede llevar a errores o decisiones inapropiadas, como en el caso de los sistemas de reconocimiento facial con tasas de error más altas para ciertos grupos demográficos.
- **Obsolescencia y mantenimiento de sistemas de IA:** Mantener y actualizar sistemas de IA puede ser costoso y técnicamente desafiante, especialmente cuando los modelos necesitan adaptarse a nuevos datos o circunstancias, planteando riesgos significativos en términos de seguridad y funcionalidad.

Por su parte la OMS, ha publicado inquietudes que se desprenden de análisis y que considera necesario contemplar para que el uso de estas tecnologías de manera segura, eficaz y ética:

- **Datos sesgados:** Los datos utilizados para entrenar la IA pueden estar sesgados, lo que generaría información engañosa o inexacta que podría plantear riesgos para la salud, la equidad y la inclusividad.
- **Datos incorrectos:** Los grandes modelos lingüísticos generan respuestas que pueden parecer autorizadas y plausibles para el usuario final; pero cabe la posibilidad de que las respuestas



sean totalmente incorrectas o que contengan errores graves, especialmente las vinculadas con la salud.

- **Confidencialidad de los datos:** Puede que los grandes modelos lingüísticos estén entrenados con datos para cuyo uso no se haya obtenido consentimiento previo, y que no protejan los datos confidenciales (en particular los datos de salud) que un usuario introduzca en una aplicación para generar una respuesta.
- **Desinformación o información engañosa:** Puede que tales modelos se utilicen indebidamente para generar y difundir desinformación muy convincente en forma de contenido textual o audiovisual que el público no puede distinguir fácilmente de contenidos de salud fidedignos.
- **Seguridad y protección del paciente:** Si bien la OMS tiene el compromiso de aprovechar las nuevas tecnologías, como la IA y la salud digital, para mejorar la salud humana, recomienda a las instancias normativas que garanticen la seguridad y protección del paciente, a la par que las empresas tecnológicas se afanan por comercializar los grandes modelos lingüísticos

Consideraciones finales

La implementación de Inteligencia Artificial en los diversos procesos de la CCSS puede transformar los servicios, la automatización, herramientas avanzadas para el diagnóstico temprano de enfermedades, análisis de grandes volúmenes para toma de decisiones, identificación de patrones o tendencias, la personalización de tratamientos, alertas tempranas ante fraudes, entre otros, son beneficios de esta tecnología que impactan directamente en la optimización, eficiencia y mejora continua.

No obstante, es importante considerar una implementación con diseños y políticas basados en principios éticos, promoviendo que se adopten enfoques responsables para el uso adecuado de la IA en la CCSS, los cuales han sido descritos en este oficio desde la visión de la OMS, así como en el contexto país, suscrito por el MICITT.

La regulación en la CCSS de la IA y las herramientas asociadas es de gran interés para esta Auditoría, promoviendo que se permita maximizar los beneficios de manera segura, y minimizar los riesgos asociados y descritos en esta misiva.

Por lo anterior, es menester de este órgano de fiscalización recomendar al ente rector, establecer la regulación interna necesaria para las tecnologías asociadas a la inteligencia artificial, garantizando su aplicación ética y responsable, y basarse en principios de transparencia, equidad y protección de datos, asegurando que la IA se utilice para mejorar los servicios de la CCSS sin comprometer la privacidad ni generar discriminación.

Además, es esencial implementar mecanismos de supervisión y gestión para minimizar los riesgos que se han mencionado como la desinformación, los sesgos algorítmicos y la dependencia tecnológica. La capacitación del personal y la evaluación continua del impacto de la IA permitirán una integración equilibrada, promoviendo la innovación sin perder el control humano en la toma de decisiones clave.



CAJA COSTARRICENSE DE SEGURO SOCIAL

Auditoría Interna

Teléfono: 2539-0821 ext. 2000-7468

Correo electrónico: coincss@ccss.sa.cr

Se somete a su consideración el establecimiento de estrategias que regulen la IA en la CCSS, contemplando la relevancia de disponer de un equipo para el gobierno de las estrategias de IA, habilitar y desarrollar el talento humano requerido, la gestión de los riesgos y lo la formalización requeridas en cuanto alianzas estratégicas.

En virtud de lo expuesto, se da conocer la información descrita, con el propósito de ser sometidas a valoración y análisis por los integrantes del Consejo Tecnológico, de tal forma que desde sus ámbitos de acción se consideren las medidas pertinentes que se puedan optar en esta temática, verificando la capacidad institucional, la normativa vigente, así como disponibilidad real de recursos, y así coadyuvar a la mejora de los procesos institucionales, y el cumplimiento de las líneas de acción definidas en el Plan Estratégico Institucional 2023 – 2033, lo anterior acorde a lo definido en el Manual Funcional del Consejo Tecnológico referente a sus funciones sustantivas, que entre otras establece la de velar por la alineación de la estrategia tecnológica con la estrategia institucional, de forma que las acciones desarrolladas por la DTIC y otras instancias institucionales se orienten al logro de esas estrategias.

Atentamente,

AUDITORÍA INTERNA



M. Sc. Olger Sánchez Carrillo
Auditor

OSC/RAHM/JPS/ayms

C. Gerencia General –1100.

Máster Mónica Taylor Hernández, presidente, Presidencia Ejecutiva - 1102

Doctor Alexánder Sánchez Cabo, gerente a.i., Gerencia Médica - 2901.

Máster Gabriela Artavia Monge, gerente a.i., Gerencia Administrativa -1104.

Licenciado Gustavo Picado Chacón, gerente, Gerencia Financiera - 1103.

Ingeniero Jorge Granados Soto, gerente, Gerencia de Infraestructura y Tecnologías - 1107.

Doctor Esteban Vega De La O, gerente, Gerencia de Logística -1106.

Licenciado Jaime Barrantes Espinoza, gerente, Gerencia de Pensiones - 9108.

Máster Susan Peraza Solano, directora, Dirección de Planificación Institucional - 2902.

Auditoría Interna 1111

Referencia: ID-131486

Referencias utilizadas:

Hewlett Packard Enterprise. (2023). *Un estudio demuestra que superar las complejidades de la gestión de datos puede reportar grandes beneficios*. Hewlett Packard Enterprise. <https://www.hpe.com/us/en/solutions/data.html>

McKinsey & Company. (2024, enero 16). *Transformando la fuerza laboral y la productividad en Centroamérica con Gen AI*. McKinsey & Company. <https://www.mckinsey.com/locations/south-america/latam/hispanoamerica-en-potencia/transformando-la-fuerza-laboral-y-la-productividad-en-centroamerica-con-gen-ai/es-CL>

OpenAI. (2025). ChatGPT (enero de 2025) [Modelo de lenguaje AI]. OpenAI. <https://chat.openai.com>.

Organización Mundial de la Salud (OMS). (2021). *Ética y gobernanza de la inteligencia artificial en el ámbito de la salud*. Organización Mundial de la Salud. [Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models](https://www.who.int/publications/m/item/ethics-and-governance-of-artificial-intelligence-for-health-guidance-on-large-multi-modal-models).