

Artículo de revisión

Uso de la inteligencia artificial en instituciones sanitarias: oportunidades, riesgos y desafíos para la seguridad social

Artificial Intelligence in Healthcare Institutions: Opportunities, Risks, and Challenges for Social Security

Edén Galán-Rodas^{1,2,a}

Filiación

- 1 Gerencia de Oferta Flexible, Gerencia Central de Operaciones, Seguro Social de Salud EsSalud. Lima, Perú
- 2 Escuela de Medicina y Cirugía, Universidad Hispanoamericana. San José, Costa Rica.
- ^a Médico Cirujano, Gestor en innovación en salud.

Correspondencia:

Edén Galán-Rodas
Dirección: Av. Arenales N°1402, Lima, Perú.
Teléfono: (511) 4707000
Correo: roman.galan@essalud.gob.pe; edgarod6@gmail.com

Declaración de autoría

El autor declara que participó en conceptualización, metodología, software, investigación, curación de datos, redacción-borrador original, redacción-revisión y edición.

Declaración de conflicto de interés

El autor declara no tener conflictos de intereses.

Financiamiento

La investigación no recibió financiamiento específico de entidades públicas, comerciales ni organizaciones sin fines de lucro.

Uso de Inteligencia Artificial

Para el diseño y la mejora visual de los gráficos incluidos en el manuscrito se utilizó inteligencia artificial generativa como herramienta de apoyo, a partir de contenidos previamente elaborados. Su uso se limitó a la reorganización didáctica, síntesis visual, propuesta de iconografía y mejora de la presentación gráfica de tablas conceptuales, manteniéndose la revisión, validación y responsabilidad final del contenido científico a cargo del autor.

Proceso editorial

Fecha de envío: 04/06/2026
Fecha de aprobación: 25/06/2026
Fecha de publicación: 30/06/2026

Cómo citar este artículo: Galán-Rodas E. Uso de la inteligencia artificial en instituciones sanitarias: oportunidades, riesgos y desafíos para la seguridad social. Rev CyT Salud Pública. 2026;1(2):139-149.



Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional.
ISSN: 3119-7981 (En Línea)
OJS: <https://rctsp.org.pe/>

RESUMEN

La inteligencia artificial (IA) se ha convertido en una herramienta estratégica para las instituciones sanitarias y de seguridad social, al permitir el análisis de grandes volúmenes de datos, la automatización de procesos, la personalización de servicios, la predicción de riesgos y el apoyo a decisiones clínicas y administrativas. Su adopción global avanza desde chatbots y asistencia digital hacia sistemas basados en aprendizaje automático, procesamiento del lenguaje natural, visión computacional, IA generativa, analítica predictiva y modelos integrados a flujos asistenciales y de gestión.

Esta revisión narrativa analiza experiencias internacionales y regionales de uso de IA en instituciones sanitarias, con énfasis en la seguridad social. Se revisan documentos de la Asociación Internacional de la Seguridad Social, informes sobre innovación institucional, el Índice Latinoamericano de Inteligencia Artificial 2025, la herramienta AI-GUARD de la OPS y el BID, documentos éticos y regulatorios internacionales, y casos clínicos y de salud pública seleccionados.

La evidencia muestra que la IA puede mejorar atención al usuario, gestión de prestaciones, auditoría, rehabilitación, retorno al trabajo, tamizaje, diagnóstico asistido, alertas tempranas y vigilancia clínica. Sin embargo, su valor depende menos de la precisión técnica aislada que de la validación local, calidad y gobernanza de datos, integración al flujo de trabajo, supervisión humana, equidad, ciberresiliencia, evaluación continua y capacidad institucional para aprender y escalar de forma responsable.

Palabras clave: *Inteligencia artificial; seguridad social; sistemas de salud; salud digital; aprendizaje automático; gobernanza de datos; seguridad del paciente; innovación institucional. (Fuente: DeCS-BIREME)*

ABSTRACT

Artificial intelligence (AI) has become a strategic tool for healthcare and social security institutions by enabling large-scale data analysis, process automation, service personalization, risk prediction, and support for clinical and administrative decision-making. Its global adoption has evolved from chatbots and digital assistance to machine learning, natural language processing, computer vision, generative AI, predictive analytics, and models embedded in care and management workflows.

This narrative review analyzes international and regional experiences of AI use in healthcare institutions, with emphasis on social security. It reviews documents from the International Social Security Association, reports on institutional innovation, the 2025 Latin American Artificial Intelligence Index, the PAHO/IDB AI-GUARD tool, international ethical and regulatory documents, and selected clinical and public health use cases.

Evidence suggests that AI can improve user services, benefit management, auditing, rehabilitation, return to work, screening, assisted diagnosis, early warning systems, and clinical surveillance. However, its value depends less on technical performance alone than on local validation, data quality and governance, workflow integration, human oversight, equity, cyber-resilience, continuous evaluation, and institutional capacity to learn and scale responsibly.

Keywords: *artificial intelligence; social security; health systems; digital health; machine learning; data governance; patient safety; institutional innovation (Source: DeCS-BIREME)*

INTRODUCCIÓN

Las instituciones sanitarias y de seguridad social enfrentan presiones crecientes derivadas del envejecimiento poblacional, el aumento de enfermedades crónicas, la escasez de profesionales, la demanda ciudadana de servicios oportunos, la sostenibilidad financiera y la expansión acelerada de datos clínicos, administrativos, laborales y sociales. En este escenario, la Inteligencia Artificial (IA) emerge como una herramienta clave con potencial para apoyar la transformación digital, mejorar la capacidad resolutoria y fortalecer la protección social (1-5).

A diferencia de las herramientas informáticas tradicionales, la IA permite reconocer patrones, clasificar información, interpretar lenguaje, analizar imágenes, predecir eventos, automatizar tareas y generar recomendaciones a partir de grandes volúmenes de datos. En salud, estas capacidades se expresan en diagnóstico asistido, tamizaje, predicción de deterioro, alertas tempranas y apoyo a decisiones clínicas. En seguridad social, se aplican en la atención al usuario, clasificación de solicitudes, gestión de prestaciones, detección de fraude, rehabilitación, evaluación de incapacidad y retorno al trabajo (1-4).

En América Latina y el Caribe, la adopción de IA debe analizarse en un contexto de brechas estructurales. El Índice Latinoamericano de Inteligencia Artificial 2025 (ILIA) señala que la región enfrenta desafíos asociados a baja capacidad de crecimiento, desigualdad, baja movilidad social y capacidades institucionales limitadas. El informe plantea que la IA puede contribuir a democratizar el acceso a la salud y otros servicios públicos, ampliar la protección social, fortalecer gobiernos y reducir brechas, pero advierte limitaciones en inversión, talento especializado, disponibilidad de datos, infraestructura, ciberseguridad y ejecución efectiva de estrategias nacionales (6).

En la Región de las Américas, la herramienta AI-GUARD (Artificial Intelligence Governance & Use Assessment for Responsible Deployment) de la Organización Panamericana de la Salud (OPS) y el Banco Interamericano de Desarrollo (BID) aporta un marco operativo para pasar de principios generales a decisiones institucionales concretas. AI-GUARD propone evaluar cada iniciativa de IA antes de su adquisición, desarrollo, adopción o ampliación, considerando su valor para la salud pública, exposición al riesgo, preparación para la gobernanza, control humano, mitigación de sesgos, evidencia, transparencia y sostenibilidad (7).

Por ello, la IA no debe entenderse como una solución automática a problemas estructurales. Su impacto depende de la calidad y representatividad de los datos, la pertinencia del problema abordado, la validación en poblaciones reales,

la interoperabilidad, la integración con procesos existentes, la supervisión humana y la existencia de marcos de gobernanza ética, legal y operativa (8-12). En instituciones de seguridad social, estas condiciones son críticas porque las decisiones algorítmicas pueden afectar derechos, prestaciones, continuidad de atención, priorización de recursos y confianza ciudadana.

Esta revisión tiene como objetivo analizar el uso de la IA en instituciones sanitarias, con énfasis en la seguridad social, identificando principales casos de uso, beneficios potenciales, riesgos, condiciones de implementación segura y desafíos para América Latina y el Perú.

MÉTODOS

Se realizó una revisión narrativa de documentos técnicos, informes institucionales y literatura científica seleccionada sobre IA en salud, seguridad social e innovación institucional. Se incluyeron informes de la Asociación Internacional de la Seguridad Social (AISS / ISSA) sobre IA en seguridad social, chatbots inteligentes, rehabilitación, tendencias regionales e innovación; el ILIA 2025; la herramienta AI-GUARD de la OPS y el BID para evaluación de gobernanza y uso responsable de IA en salud; guías internacionales sobre ética, regulación y gobernanza; y estudios clínicos representativos de aplicaciones de IA.

El análisis se organizó en seis ejes: aplicaciones institucionales de IA en seguridad social; casos de uso clínico y de salud pública; innovación institucional; gobernanza, ética, derechos y regulación; ciberseguridad y resiliencia; y condiciones para implementación segura en América Latina y Perú. Dada la naturaleza narrativa del artículo, no se realizó metaanálisis ni evaluación formal de riesgo de sesgo.

Inteligencia artificial en instituciones sanitarias y de seguridad social

La IA puede definirse como un conjunto de métodos computacionales que permiten a sistemas digitales realizar tareas asociadas a capacidades cognitivas humanas, como reconocer patrones, interpretar lenguaje, clasificar imágenes, aprender de datos, predecir eventos o apoyar decisiones. En la práctica institucional actual, la mayoría de aplicaciones corresponden a IA estrecha, diseñada para tareas específicas y no para reemplazar integralmente el juicio humano (13,14). En salud y seguridad social, las ramas más utilizadas incluyen aprendizaje automático, aprendizaje profundo, procesamiento del lenguaje natural, visión computacional, IA generativa, analítica predictiva y automatización inteligente. Estas tecnologías operan sobre datos estructurados, como historias clínicas electrónicas, diagnósticos codificados,

cotizaciones y prestaciones; y sobre datos no estructurados, como imágenes, documentos escaneados, notas clínicas, llamadas, correos o conversaciones digitales.

La incorporación de IA en seguridad social responde a una necesidad concreta: prestar servicios más ágiles, personalizados, accesibles y sostenibles sin perder de vista el mandato público de protección social. En este contexto, la IA no solo tiene una dimensión tecnológica, sino también institucional, ética y política.

Principales categorías de uso institucional

Atención al usuario, chatbots y servicios digitales

Uno de los usos más extendidos de la IA en seguridad social ha sido la incorporación de chatbots, asistentes virtuales y sistemas conversacionales para responder consultas, orientar trámites y mejorar la disponibilidad de servicios. La AISS señala que los organismos públicos han optado especialmente por chatbots para atender grandes volúmenes de usuarios, ofrecer servicios personalizados y facilitar respuestas a consultas frecuentes (3).

En seguridad social, los chatbots no deben entenderse solo como herramientas de atención al cliente. Pueden funcionar como puerta de entrada digital al sistema, mecanismo de educación previsional, canal de orientación para prestaciones, herramienta de reducción de carga operativa y componente de inclusión cuando se diseñan con lenguaje claro, accesibilidad, multicanalidad y derivación humana. Sin embargo, requieren bases de conocimiento actualizadas, monitoreo de calidad, mecanismos de escalamiento, protección de datos y evaluación de posibles errores o exclusiones (3,5).

Automatización administrativa y gestión de prestaciones

La IA puede apoyar el procesamiento documental, la

clasificación de solicitudes, el reconocimiento óptico de caracteres, la extracción de información clínica o financiera, la asignación de expedientes y la priorización de casos. Estas aplicaciones reducen tiempos de respuesta y liberan recursos humanos para situaciones complejas, pero deben mantener trazabilidad, revisión humana y mecanismos de apelación cuando afectan derechos o prestaciones (1,2,5).

Detección de fraude, auditoría y sostenibilidad financiera

Los modelos de aprendizaje automático permiten identificar patrones inusuales en solicitudes, reembolsos, cotizaciones o prestaciones. Esta capacidad puede fortalecer la sostenibilidad financiera y la integridad institucional. No obstante, las aplicaciones de auditoría y control son especialmente sensibles porque falsos positivos o sesgos pueden restringir indebidamente el acceso a servicios o beneficios (1,2).

Predicción de riesgos clínicos, administrativos y poblacionales La IA puede anticipar demanda, reingresos, deterioro clínico, eventos adversos, incapacidad prolongada, necesidades de rehabilitación o riesgo de abandono de tratamiento. En seguridad social, estos modelos permiten orientar servicios proactivos y priorizar recursos, siempre que su uso sea proporcional, auditable y conectado con intervenciones efectivas (1,4,8).

Rehabilitación, discapacidad y retorno al trabajo

La IA tiene un campo especialmente relevante en rehabilitación y seguridad social vinculada a enfermedad, discapacidad, invalidez y retorno laboral. Puede apoyar la personalización de planes, el monitoreo de progreso, la predicción de recaídas y la recomendación de trayectorias de retorno al trabajo. Experiencias reportadas por la AISS, como K-COMWEL en Corea y NIHDI en Bélgica, muestran el potencial de Big Data e IA para articular salud, empleo y protección social (4).

Tabla 1. Principales usos institucionales de la IA en salud y seguridad social

Área	Aplicaciones	Potencial beneficio	Riesgo principal
Atención al usuario	Chatbots, asistentes virtuales, IA generativa	Atención 24/7, orientación de trámites, menor carga operativa	Respuestas incorrectas, exclusión digital, falta de derivación humana
Gestión administrativa	OCR, clasificación documental, asignación de casos	Reducción de tiempos y mayor trazabilidad	Errores de clasificación o automatización opaca
Prestaciones y auditoría	Detección de anomalías, pagos indebidos y fraude	Sostenibilidad financiera e integridad institucional	Falsos positivos, sesgos y afectación de derechos
Clínica asistencial	Alertas, predicción de deterioro, apoyo diagnóstico	Intervención temprana y priorización clínica	Fatiga de alertas y baja validación local
Imágenes y tamizaje	Retina, mamografía, radiología, colonoscopia	Mayor capacidad diagnóstica y acceso a especialistas	Sobreconfianza y errores no detectados
Rehabilitación y trabajo	Predicción de incapacidad, planes personalizados	Retorno laboral y continuidad funcional	Reducción de casos complejos a perfiles algorítmicos
Gobernanza y planificación	Análítica predictiva, tableros, simulación	Mejor planificación y gestión de riesgos	Datos incompletos, sesgados o no interoperables

Elaboración propia a partir de AISS/UNU-EGOV, AISS Américas 2026 y OPS/BID AI-GUARD (1,4,5,7).

Tabla 2. Factores críticos para implementar chatbots inteligentes en seguridad social

Dimensión	Criterio operativo	Implicancia institucional
Propósito	Definir dominio, usuarios y trámites cubiertos	Evita respuestas fuera de alcance y expectativas irreales
Contenido	Base de conocimiento oficial, actualizada y validada	Reduce desinformación y mejora confianza ciudadana
Escalamiento	Derivación a persona cuando hay complejidad, riesgo o reclamo	Protege derechos y evita automatización ciega
Accesibilidad	Lenguaje claro, diseño inclusivo y canales alternativos	Evita exclusión de personas mayores, rurales o con baja alfabetización digital
Privacidad	Minimización de datos, registro seguro y límites de uso	Protege información sensible de salud y prestaciones
Evaluación	Indicadores de resolución, satisfacción, errores y equidad	Permite mejora continua y auditoría institucional

Elaboración propia a partir de la AISS sobre chatbots inteligentes en instituciones de seguridad social (3).

Casos de uso clínico y de salud pública

Los casos clínicos y de salud pública permiten conectar la evidencia técnica con la pregunta institucional: cuándo una tecnología mejora desenlaces reales y cuándo solo mejora métricas de desempeño. A continuación, se presentan cinco casos clínicos trazadores y un caso latinoamericano de salud pública vinculado al diagnóstico de tuberculosis (Figura 1)

Retinopatía diabética con IA autónoma

La retinopatía diabética es una causa prevenible de pérdida visual, pero su tamizaje depende de disponibilidad de cámaras, lectura especializada y rutas de referencia. El ensayo pivotal de Abramoff y colaboradores evaluó un sistema autónomo de IA para detectar retinopatía diabética en atención primaria, mostrando que una herramienta regulada puede acercar capacidad diagnóstica especializada al primer nivel de atención (15). La lección institucional es que la IA puede ampliar acceso si el flujo incluye captura de imagen de calidad, resultado comprensible, derivación oportuna y monitoreo de desempeño.

Mamografía de tamizaje asistida por IA: estudio MASAI

La mamografía poblacional exige doble lectura, radiólogos entrenados y gestión de grandes volúmenes de imágenes. El ensayo MASAI evaluó lectura apoyada por IA frente a doble lectura estándar en tamizaje mamográfico y mostró que la IA puede apoyar programas poblacionales reduciendo carga de lectura sin comprometer seguridad clínica cuando se implementa bajo protocolos controlados (16). La lección no

es reemplazar al radiólogo, sino rediseñar el programa de tamizaje, incluyendo priorización, segunda lectura selectiva, auditoría y evaluación de cánceres de intervalo.

Colonoscopia asistida por IA para detección de adenomas

La colonoscopia es operador-dependiente y la tasa de detección de adenomas es un indicador clave de calidad. Un ensayo clínico aleatorizado evaluó un sistema automático de control de calidad y detección asistida durante colonoscopia rutinaria, mostrando el potencial de apoyo en tiempo real para mejorar la calidad diagnóstica (17). Este caso ilustra una IA integrada al acto clínico, que requiere entrenamiento, aceptación profesional, mantenimiento técnico y evaluación de posibles falsos positivos.

Electrocardiograma con IA para fibrilación auricular oculta

La fibrilación auricular paroxística puede pasar inadvertida y aumentar el riesgo de accidente cerebrovascular. Noseworthy y colaboradores evaluaron un algoritmo de IA aplicado a EKG en ritmo sinusal para identificar personas con mayor probabilidad de fibrilación auricular oculta (18). La lección institucional es que la IA puede transformar una señal rutinaria y de bajo costo en una herramienta de estratificación, pero su implementación exige protocolos de confirmación, comunicación de riesgo y uso racional de recursos.

TREWS: alerta temprana de sepsis integrada al flujo asistencial

La sepsis requiere reconocimiento y tratamiento oportunos. TREWS fue desarrollado como un sistema de alerta temprana

dirigido y evaluado posteriormente en un estudio prospectivo multicéntrico, mostrando que una alerta predictiva puede asociarse con mejores resultados cuando se integra a una cadena clínica de confirmación y respuesta (19,20). Este caso resume el principio central de implementación: una predicción solo genera valor si activa un proceso clínico confiable.

Diagnóstico de tuberculosis con IA y salud móvil en Perú

Un caso especialmente relevante para América Latina es el uso de IA para optimizar el diagnóstico de tuberculosis en el Perú. Curioso y Brunette describen la herramienta eRx, una solución basada en salud móvil, análisis remoto de radiografías de tórax, aprendizaje profundo y redes neuronales convolucionales para detectar anomalías pulmonares y sugerir manifestaciones clínicas preliminares compatibles con tuberculosis (21).

En el proyecto eRx, profesionales de enfermería o personal técnico capacitado capturan imágenes de radiografías de tórax mediante teléfonos inteligentes en establecimientos de salud y las transmiten a través de una aplicación móvil para análisis remoto. El piloto se desarrolló en cinco establecimientos de salud de Huaycán, Ate, Lima, con participación de enfermeras y médicos especialistas en neumología (21). Este caso muestra el potencial de la IA para acercar capacidades diagnósticas a contextos con limitaciones de infraestructura, escasez de especialistas y barreras de acceso.

La experiencia peruana también refuerza que el valor de la IA no reside solo en el algoritmo, sino en el rediseño del proceso diagnóstico y en la consideración de factores sociotécnicos, culturales y organizacionales. Esta perspectiva es particularmente importante para tuberculosis, enfermedades transmisibles, atención primaria y poblaciones vulnerables en países de ingresos medios y bajos (21,22).



Nota: ACV = accidente cerebrovascular; FA = fibrilación auricular; mSalud = salud móvil

Elaboración propia mediante inteligencia artificial generativa, bajo supervisión del autor.

Figura 1. Casos trazadores de IA en salud y lecciones institucionales

Innovación institucional: de proyectos aislados a capacidades sostenibles

El documento “Innovación con propósito” de AISS aporta un marco útil para evitar una visión tecnocéntrica de la IA. Su tesis central es que el reto de las instituciones de seguridad social no es necesariamente la falta de ideas, pilotos o talento, sino la falta de alineamiento entre intención de liderazgo, gobernanza, prácticas operativas y cultura organizacional (23).

Desde esta perspectiva, la IA no debería implementarse como una colección de pilotos desconectados ni como una adquisición tecnológica aislada. Debe entenderse como parte

de una capacidad institucional que se construye mediante portafolios de iniciativas conectadas, donde la coherencia entre gobernanza, operaciones y cultura importa más que la completitud de cualquier proyecto individual (23) (Figura 2).

La AISS identifica patrones frecuentes de fracaso en innovación pública: dependencia de individuos heroicos, innovación de escaparate desconectada de decisiones, prototipos que no salen del laboratorio y pilotos que no escalan. En IA, estos riesgos se expresan en modelos que funcionan en laboratorio pero no se integran al flujo real, chatbots que responden consultas pero no resuelven problemas complejos, o tableros predictivos que no modifican decisiones institucionales (23).



Marco orientado a la implementación responsable de la IA en salud y seguridad social

Figura 2. Lógica de portafolio para implementar IA en instituciones sanitarias y de seguridad social
Elaboración propia mediante inteligencia artificial generativa, bajo supervisión del autor.

Gobernanza clínica, ética, derechos y regulación

La expansión de la IA en salud y seguridad social plantea riesgos de sesgo algorítmico, baja representatividad de datos, sobreajuste, opacidad, automatización excesiva, errores no detectados, inequidad en acceso, dependencia de proveedores, problemas de interoperabilidad, ciberataques y uso secundario no autorizado de datos. En instituciones de seguridad social, estos riesgos pueden afectar prestaciones, auditoría, priorización, incapacidad laboral y confianza pública (8-12,24,25).

El Manual de Buenas Prácticas para la IA en Medicina de la Organización Médica Colegial resume una idea rectora: medicina inteligente, decisiones humanas. La IA debe entenderse como inteligencia aumentada: una herramienta que amplía la capacidad diagnóstica, analítica y organizativa del profesional, sin reemplazar su juicio clínico ni su responsabilidad deontológica (8).

AI-GUARD: evaluación proporcional del riesgo antes de implementar IA

La herramienta AI-GUARD, desarrollada por la OPS y el BID, ofrece un marco práctico para evaluar la gobernanza y el uso responsable de iniciativas de IA en salud antes de su adopción, adquisición, desarrollo o escalamiento. Su aporte central es traducir principios como transparencia, equidad, rendición de cuentas y supervisión humana en una ruta

de decisión institucional aplicable a ministerios de salud, hospitales, programas nacionales, unidades de salud digital, organismos reguladores y entidades de supervisión (7).

AI-GUARD propone una evaluación escalonada según el nivel de impacto y riesgo (Figura 3). El Nivel 1 corresponde al uso responsable de IA en tareas de bajo riesgo, como apoyo administrativo, redacción, resumen, traducción o automatización interna. El Nivel 2 aplica a herramientas a nivel de programa que apoyan o influyen en decisiones, como puntuación de riesgos, sistemas de apoyo clínico, aplicaciones públicas o tableros de vigilancia. El Nivel 3 corresponde a IA de alto impacto, como diagnóstico asistido, modelos de elegibilidad, asignación de recursos públicos o vigilancia nacional, donde se exigen salvaguardias más estrictas (7).

El instrumento organiza la evaluación en cuatro pilares: gobernanza y responsabilidad; uso y control humano; salvaguardias contra riesgos y sesgos; y preparación para la implementación. Además, utiliza índices como el valor para la salud pública, preparación para la gobernanza, riesgo de sesgo y equidad, evidencia y transparencia, y control humano. Esta lógica es particularmente útil para instituciones de seguridad social, porque permite diferenciar entre herramientas administrativas de bajo riesgo y sistemas que pueden afectar diagnóstico, elegibilidad, prestaciones, asignación de recursos o poblaciones vulnerables (7).

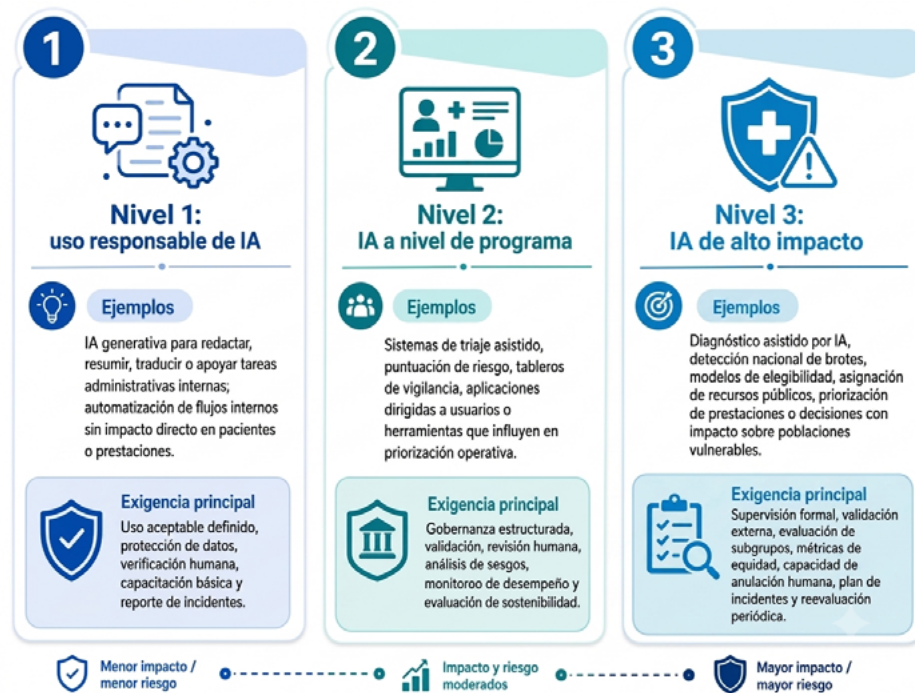


Figura 3. Aplicación institucional de AI-GUARD para iniciativas de IA en salud y seguridad social
 Elaboración propia mediante inteligencia artificial generativa, bajo supervisión del autor.

Supervisión humana significativa

La supervisión humana no debe ser simbólica. Debe permitir que el profesional comprenda, revise, contraste, modifique o rechace una recomendación antes de que produzca efectos clínicos o administrativos relevantes (8,11). En la práctica, implica conocer los datos de entrada, el propósito del modelo, sus límites, el nivel de confianza, el contexto de validación, la trazabilidad y los procedimientos para reportar incidentes.

Validación local y adaptación al entorno asistencial

Un modelo entrenado en una población, sistema de codificación, tecnología o flujo asistencial puede no funcionar igual en otro contexto. Por ello, la validación externa y local debe ser recurrente, no solo inicial, e incluir calibración, análisis por subgrupos, evaluación de drift, comparación con práctica habitual y medición de impacto clínico o institucional (8,26,27). Las guías de reporte y evaluación temprana, como TRIPOD+AI, DECIDE-AI, SPIRIT-AI y CONSORT-AI, refuerzan la necesidad de documentar datos, desempeño, intervención humana y evaluación clínica antes de escalar soluciones de IA [28-30].

Datos, derechos y decisiones automatizadas

La IA en salud y en la seguridad social funciona con datos particularmente sensibles. Su uso exige base legal, minimización, seguridad, trazabilidad, control de accesos,

información comprensible y límites al uso secundario. Las decisiones exclusivamente automatizadas con efectos significativos sobre salud, prestaciones o derechos deben evitarse o estar sujetas a intervención humana efectiva, explicación y posibilidad de impugnación (8-11).

Ciberseguridad y resiliencia

La IA amplía la superficie de riesgo institucional al introducir nuevos flujos de datos, modelos, interfaces, proveedores y dependencias tecnológicas. En América Latina, el ILIA 2025 advierte avances legales en ciberseguridad y protección de datos, pero también brechas en capacidades técnicas y organizativas (6). La ciberresiliencia debe incluir recuperación, continuidad operativa, control de proveedores, monitoreo de modelos, detección de drift, gestión de incidentes y protección frente a usos indebidos (31).

Antes de implementar una solución de IA en instituciones sanitarias o de seguridad social, es necesario verificar un conjunto mínimo de condiciones que permitan pasar de la intención tecnológica a una adopción segura, útil y sostenible. El Figura 4 sintetiza ocho dominios clave que deberían revisarse de manera sistemática: necesidad y valor, evidencia, datos, riesgo, flujo de trabajo, supervisión humana, monitoreo y capacidades institucionales.



Figura 4. Criterios mínimos antes de implementar una herramienta de IA clínica o institucional. Elaboración propia mediante inteligencia artificial generativa, bajo supervisión del autor.

América Latina, Perú y seguridad social

EILIA2025 ofrece un marco útil para analizar la implementación de IA en instituciones sanitarias y de seguridad social de América Latina y el Caribe. La región presenta avances en conectividad, adopción de IA generativa, programas de formación y estrategias nacionales, pero persisten brechas importantes en datos abiertos, infraestructura de cómputo, talento especializado, investigación, inversión, ciberseguridad y ejecución efectiva de políticas (6).

El informe agrupa los ecosistemas nacionales en pioneros, adoptantes y exploradores, y ubica a Perú dentro del grupo de adoptantes. Además, destaca que Perú muestra intensidad de uso de IA, pero enfrenta desafíos en infraestructura, gobernanza de datos, formación avanzada y articulación entre adopción tecnológica e innovación institucional (6). Para el sector salud y seguridad social, esto implica que la oportunidad no consiste únicamente en incorporar herramientas, sino en fortalecer capacidades habilitantes y gobernanza operativa.

En el Perú, la IA podría priorizar problemas de alto valor público: listas de espera, referencia y contrarreferencia, tamizaje, enfermedades crónicas, tuberculosis, continuidad del cuidado, auditoría de prestaciones, incapacidad temporal, rehabilitación, retorno al trabajo y comunicación con poblaciones vulnerables. El caso eRx en tuberculosis

demuestra que es posible diseñar soluciones contextualizadas que integren IA, salud móvil, telesalud y factores sociotécnicos (21,22).

Para la seguridad social, la IA debe evaluarse desde una perspectiva de protección social. No basta con introducir tecnología si no mejora acceso, oportunidad, calidad, continuidad, equidad y confianza. En poblaciones rurales, indígenas, personas mayores, personas con discapacidad o usuarios con baja alfabetización digital, la implementación debe mantener canales alternativos y evitar que la digitalización se convierta en una nueva barrera de acceso.

DISCUSIÓN

La evidencia revisada muestra que la IA se está incorporando progresivamente en instituciones sanitarias y de seguridad social como parte de procesos más amplios de transformación digital, innovación organizacional y modernización de servicios. Su aplicación no se limita al ámbito clínico, sino que abarca atención al usuario, automatización administrativa, gestión de prestaciones, detección de fraude, predicción de riesgos, rehabilitación, investigación biomédica y soporte a decisiones institucionales (1-5,23).

En seguridad social, los chatbots inteligentes constituyen uno de los casos de uso más difundidos. La AISS destaca que los servicios públicos han priorizado chatbots para responder grandes volúmenes de consultas y mejorar la atención digital (3). Sin embargo, su utilidad depende de condiciones institucionales: bases de conocimiento oficiales, actualización continua, monitoreo de desempeño, escalamiento humano y evaluación de errores. En ausencia de estas condiciones, un chatbot puede aumentar desinformación, frustración o exclusión digital.

Los casos clínicos revisados muestran distintos niveles de madurez de la IA en medicina: diagnóstico autónomo regulado, apoyo a lectura de imágenes, asistencia en tiempo real durante procedimientos, estratificación predictiva y alerta temprana integrada al flujo asistencial (15-20). La lección transversal es que la IA no genera valor únicamente por predecir o clasificar, sino cuando modifica decisiones, activa intervenciones efectivas y mejora resultados relevantes. Esta distinción es crítica porque muchas soluciones muestran buen desempeño en métricas internas, pero fracasan por falta de interoperabilidad, mala integración, fatiga de alertas, baja aceptación profesional, ausencia de mantenimiento o validación local insuficiente (26,27).

El caso de tuberculosis en Perú permite ampliar la discusión más allá de experiencias de países de altos ingresos. eRx ilustra cómo una solución basada en IA puede integrarse con teléfonos inteligentes, telesalud y flujos de atención primaria para responder a un problema de salud pública en contextos con recursos limitados (21). Su valor no reside solo en el algoritmo, sino en el rediseño del proceso diagnóstico y en la consideración de factores sociotécnicos, culturales y organizacionales. Esta experiencia es especialmente relevante para sistemas de seguridad social que atienden poblaciones vulnerables y enfrentan brechas territoriales de acceso.

La discusión ética y regulatoria es igualmente central. La OMS ha advertido que la IA en salud debe orientarse por transparencia, inclusión, responsabilidad, protección de datos, seguridad y beneficio público (9,10). La Ley Europea de IA introduce un enfoque basado en riesgo y obligaciones para proveedores y desplegados, mientras que el enfoque de buenas prácticas médicas exige supervisión humana significativa, validación local, trazabilidad, formación profesional y protección de derechos de pacientes y médicos (8,11).

AI-GUARD complementa estos marcos al ofrecer una metodología operativa y proporcional al riesgo para evaluar iniciativas concretas antes de su implementación. Su utilidad para instituciones sanitarias y de seguridad social radica en

que no trata todas las soluciones de IA por igual: exige mayor evidencia, validación externa, evaluación de subgrupos, métricas de equidad, trazabilidad, capacidad de anulación humana y protocolos de incidentes cuando el sistema puede afectar diagnóstico, elegibilidad, asignación de recursos o poblaciones vulnerables (7).

El ILIA 2025 refuerza que el despliegue de IA en América Latina no puede separarse de las brechas de infraestructura, datos, talento y gobernanza. La región produce grandes volúmenes de datos, pero su disponibilidad y estandarización siguen siendo limitadas; la alfabetización en IA crece más rápido que la especialización; la investigación está concentrada en pocos países; y muchas estrategias nacionales carecen de presupuesto, indicadores o planes efectivos de implementación (6). Estas brechas condicionan la capacidad de instituciones sanitarias y de seguridad social para adoptar IA de manera segura y equitativa.

Por ello, la adopción de IA debe entenderse como una capacidad institucional y no como una sucesión de pilotos tecnológicos. La AISS propone gobernar la innovación mediante portafolios de iniciativas conectadas, alineadas con prioridades estratégicas y sostenidas por liderazgo, operaciones y cultura organizacional (23). Este enfoque resulta especialmente pertinente para salud y seguridad social, donde los casos de uso deben conectarse con problemas de valor público: acceso, oportunidad, calidad, sostenibilidad financiera, rehabilitación, protección de derechos y reducción de inequidades.

En síntesis, la pregunta central no es si las instituciones sanitarias deben utilizar IA, sino para resolver qué problema, con qué datos, bajo qué gobernanza, con qué supervisión humana, con qué evaluación de impacto y con qué garantías para pacientes, profesionales y ciudadanos. La IA puede fortalecer instituciones más proactivas, resolutivas e inclusivas, pero solo si se implementa con prudencia, evidencia, transparencia y responsabilidad pública.

Conclusiones

La IA representa una oportunidad estratégica para transformar instituciones sanitarias y de seguridad social, al mejorar la prestación de servicios, apoyar decisiones clínicas y administrativas, optimizar recursos, anticipar riesgos y ampliar capacidades diagnósticas y operativas. Las experiencias revisadas muestran avances en chatbots, automatización, auditoría, rehabilitación, retorno al trabajo, tamizaje, imágenes médicas, alertas tempranas y salud pública.

No obstante, el valor de la IA depende de una implementación responsable y centrada en las personas. La precisión técnica aislada es insuficiente. Se requieren datos de calidad, validación local, integración al flujo de trabajo, supervisión humana significativa, interoperabilidad, ciberresiliencia, monitoreo continuo, auditoría, protección de derechos y evaluación de impacto en equidad y resultados. En esa línea, herramientas como AI-GUARD pueden ayudar a institucionalizar una evaluación proporcional al riesgo antes de adquirir, desarrollar o escalar soluciones de IA.

Para América Latina y el Perú, la IA debe orientarse a problemas prioritarios de salud pública y protección social, evitando reproducir brechas digitales, territoriales o culturales. La experiencia regional muestra que la adopción tecnológica debe acompañarse de capacidades institucionales, talento especializado, gobernanza de datos, cooperación y estrategias de innovación sostenibles.

La IA no debe reemplazar el juicio clínico ni la responsabilidad institucional. Su mayor aporte está en amplificar la capacidad humana, mejorar la oportunidad de las decisiones y fortalecer sistemas de salud y seguridad social más seguros, equitativos, resilientes y sostenibles.

REFERENCIAS BIBLIOGRÁFICAS

- 1 International Social Security Association; United Nations University Operating Unit on Policy-Driven Electronic Governance. Artificial intelligence in social security organizations. Geneva: International Social Security Association; 2024.
- 2 Asociación Internacional de la Seguridad Social. Inteligencia artificial en la seguridad social: antecedentes y experiencias [Internet]. Ginebra: AISS; 2020 [citado 2026 Jun 13]. Disponible en: <https://www.issa.int/es/analysis/artificial-intelligence-social-security-background-and-experiences>
- 3 Asociación Internacional de la Seguridad Social. La inteligencia artificial en las instituciones de seguridad social: el caso de los chatbots inteligentes [Internet]. Ginebra: AISS; 2022 [citado 2026 Jun 13]. Disponible en: <https://www.issa.int/es/analysis/artificial-intelligence-social-security-institutions-case-intelligent-chatbots>
- 4 Park J, Brunois T. Uso de la inteligencia artificial y la big data en la rehabilitación en el marco de la seguridad social. Ginebra: Asociación Internacional de la Seguridad Social; 2025.
- 5 Asociación Internacional de la Seguridad Social. Desarrollos y tendencias en la seguridad social: Américas 2026. Ginebra: AISS; 2026.
- 6 Soto Á, Durán R, Moreno A, Adasme S, Rovira S, Jordán V, et al., coordinadores. Índice Latinoamericano de Inteligencia Artificial 2025. Santiago: Comisión Económica para América Latina y el Caribe; Centro Nacional de Inteligencia Artificial; 2025.
- 7 Organización Panamericana de la Salud; Banco Interamericano de Desarrollo. Herramienta AI-GUARD: inteligencia artificial. Evaluación de la gobernanza y el uso para un despliegue responsable. Washington, D.C.: OPS; 2026.
- 8 Organización Médica Colegial de España. Manual de buenas prácticas para la inteligencia artificial en medicina. Madrid: Organización Médica Colegial de España; 2026.
- 9 World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.
- 10 World Health Organization. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Geneva: World Health Organization; 2024.
- 11 European Parliament; Council of the European Union. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Official Journal of the European Union. 2024.
- 12 Organisation for Economic Co-operation and Development. OECD AI principles. Paris: OECD; 2024.
- 13 Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7.
- 14 Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med.* 2019;380(14):1347-1358. doi:10.1056/NEJMra1814259.
- 15 Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med.* 2018;1:39. doi:10.1038/s41746-018-0040-6.
- 16 Lång K, Josefsson V, Larsson AM, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial: a clinical safety analysis of a randomised, controlled, non-inferiority,

- single-blinded, screening accuracy study. *Lancet Oncol.* 2023;24(8):936-944. doi:10.1016/S1470-2045(23)00298-X.
- 17 Liu J, Zhou R, Liu C, Liu H, Cui Z, Guo Z, et al. Automatic quality control system and adenoma detection rates during routine colonoscopy: a randomized clinical trial. *JAMA Netw Open.* 2025;8(1):e2457241. doi:10.1001/jamanetworkopen.2024.57241.
 - 18 Noseworthy PA, Attia ZI, Behnken EM, Giblon RE, Bews KA, Liu S, et al. Artificial intelligence-guided screening for atrial fibrillation using electrocardiogram during sinus rhythm: a prospective non-randomised interventional trial. *Lancet.* 2022;400(10359):1206-1212. doi:10.1016/S0140-6736(22)01637-3.
 - 19 Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. *Sci Transl Med.* 2015;7(299):299ra122. doi:10.1126/scitranslmed.aab3719.
 - 20 Adams R, Henry KE, Sridharan A, Soleimani H, Zhan A, Rawat N, et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat Med.* 2022;28(7):1455-1460. doi:10.1038/s41591-022-01894-0.
 - 21 Curioso WH, Brunette MJ. Inteligencia artificial e innovación para optimizar el proceso de diagnóstico de la tuberculosis. *Rev Peru Med Exp Salud Publica.* 2020;37(3):554-558. doi:10.17843/rpmpesp.2020.373.5585.
 - 22 Alcantara MF, Cao Y, Liu B, Liu C, Zhang N, Zhang P, et al. eRx: a technological advance to speed-up TB diagnostics. *Smart Health.* 2020;16:100117. doi:10.1016/j.smhl.2020.100117.
 - 23 Scott N, Frota L. Innovation on purpose: leadership and the art of world building. Geneva: International Social Security Association; 2026.
 - 24 Char DS, Shah NH, Magnus D. Implementing machine learning in health care: addressing ethical challenges. *N Engl J Med.* 2018;378(11):981-983. doi:10.1056/NEJMp1714229.
 - 25 Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med.* 2019;25(9):1337-1340. doi:10.1038/s41591-019-0548-6.
 - 26 Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* 2019;17(1):195. doi:10.1186/s12916-019-1426-2.
 - 27 Youssef A, Pencina M, Thakur A, Zhu T, Clifton D, Shah NH. External validation of AI models in health should be replaced with recurring local validation. *Nat Med.* 2023;29(11):2686-2687. doi:10.1038/s41591-023-02540-z.
 - 28 Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ.* 2024;385:e078378. doi:10.1136/bmj-2023-078378.
 - 29 Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al.; DECIDE-AI expert group. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med.* 2022;28(5):924-933. doi:10.1038/s41591-022-01772-9.
 - 30 Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med.* 2020;26(9):1351-1363. doi:10.1038/s41591-020-1037-7.
 - 31 Veeam. Data trust and resilience report 2026. Veeam Insights; 2026.