



Propuesta para implementar un modelo preventivo para la detección del fraude en Colpensiones basada en Inteligencia Artificial.

Modalidad: PAT

Innovación aplicada a la gestión pública

Paola Andrea Orjuela Ortiz, Andrea Orjuela Silva y Gleimir Yesenia Tellez Espitia.

Propuesta para implementar un modelo preventivo para la detección del fraude en
Colpensiones basada en Inteligencia Artificial.

Paola Andrea Orjuela Ortiz, Andrea Orjuela Silva y Gleimir Yesenia Tellez Espitia

Trabajo de grado presentado como requisito para optar al título de:
Magister en Gobierno y Gestión Pública

Director:

Santiago A. Roa Ortiz PhD (c)

Modalidad:

Innovación aplicada a la gestión pública

Universidad EAN

Facultad de Administración Finanzas y Ciencias Económicas

Programa Maestría en Gobierno y Gestión Pública

Bogotá, Colombia

13/11/2025

Resumen

Colpensiones enfrenta uno de los mayores desafíos estructurales en la detección y prevención del fraude, situación que compromete la sostenibilidad financiera del régimen de prima media. A pesar de las normativas, instrucciones internas y controles administrativos vigentes, persisten irregularidades derivadas de la falta de interoperabilidad entre entidades públicas y la baja calidad de los datos. En este contexto, la entrada en vigor de la Ley 2381 de 2024 representa una oportunidad para fortalecer la eficiencia institucional mediante el uso de inteligencia artificial, garantizando trazabilidad, transparencia y respeto al debido proceso.

La investigación propone un modelo preventivo basado en inteligencia artificial explicable (XAI) que integra algoritmos de detección de anomalías y herramientas como SHAP, LIME y Anchors, las cuales permiten interpretar los resultados y acompañar cada alerta con evidencia verificable. El modelo se desarrolla en fases progresivas: diagnóstico, diseño de interoperabilidad.

Los resultados proyectan una reducción del 50 % al 70 % de los casos de fraude, con ahorros estimados entre 4.1 y 5.7 billones de pesos en cinco años y un Retorno sobre la Inversión (ROI) del 13.900 %, recuperable en menos de cinco meses. Asimismo, se plantean dos soluciones innovadoras: el cierre de la brecha de interoperabilidad, con ahorros entre 828 mil millones y 1.16 billones de pesos, y la mejora en la calidad de los datos, que reduciría entre 40 % y 60 % las inconsistencias, generando ahorros adicionales entre 2 y 3.6 billones de pesos. La propuesta fortalece la gobernanza digital, la transparencia y la confianza ciudadana.

Palabras Clave: Prevención del Fraude, Interoperabilidad, Gobernanza Digital, Sostenibilidad Fiscal, Inteligencia Artificial, Transparencia Institucional, Modernización del Sector Público, Trazabilidad documental

Introducción

La Administradora Colombiana de Pensiones - Colpensiones, es una empresa industrial y comercial del Estado, organizada como entidad financiera de carácter especial y vinculada al Ministerio del Trabajo. Su estructura actual, establecida mediante el Decreto 309 de 2017, fortaleció la capacidad operativa, los mecanismos de control y la atención al ciudadano bajo los principios del servicio público y la transparencia.

Con base en el artículo 155 de la Ley 1151 de 2007, Colpensiones integra el Sistema General de Pensiones (SGP) y tiene como objeto la administración estatal del Régimen de Prima Media con Prestación Definida (RPM), así como del Sistema de Ahorro de Beneficios Económicos Periódicos (BEPS), garantizando la sostenibilidad financiera y el acceso equitativo a la seguridad social en Colombia.

Entre sus principales responsabilidades se destacan la determinación y reconocimiento de derechos pensionales y prestaciones económicas, la gestión de la historia laboral, la administración de reservas e inversiones, y el recaudo y pago oportuno de los recursos del régimen. De manera complementaria, impulsa procesos de innovación tecnológica, digitalización y mejora continua para optimizar la eficiencia administrativa y la confianza ciudadana.

En síntesis, Colpensiones es el pilar público del sistema pensional colombiano, garante de la solidaridad intergeneracional, la sostenibilidad del régimen y la materialización efectiva del derecho a la seguridad social.

Colpensiones enfrenta uno de sus retos más críticos: prevenir el fraude que amenaza la sostenibilidad financiera del régimen de prima media con prestación definida. A pesar de los controles administrativos y legales vigentes, persisten prácticas fraudulentas que afectan directamente a los beneficiarios legítimos y desvían recursos que deberían garantizar pensiones adquiridas con justo título y de buena fe.

La entrada en vigencia de la Ley 2381 de 2024 abre una oportunidad para fortalecer la eficiencia institucional y avanzar hacia una prevención más efectiva. Sin embargo, la falta de interoperabilidad entre las entidades públicas limita el análisis integral de las solicitudes, impidiendo contar con información veraz y actualizada sobre cada afiliado o beneficiario.

Esta brecha tecnológica interinstitucional no solo dificulta la detección oportuna del fraude, sino que también compromete la capacidad de respuesta del sistema pensional frente a riesgos emergentes. Aunque existen procedimientos normativos y administrativos para la prevención, se requiere una transformación profunda que integre herramientas avanzadas como la inteligencia artificial (IA).

La inteligencia artificial (IA), aplicada de forma ética, legal y transparente, puede convertirse en un agente preventivo clave para Colpensiones. Su implementación permitiría automatizar el análisis de datos, identificar patrones sospechosos, fortalecer la trazabilidad de la información y garantizar la protección de los derechos fundamentales de los afiliados, respetando el debido proceso.

Bajo la metodología *How Might We?*, surgió la siguiente pregunta: ¿Cómo podríamos, desde Colpensiones y en coordinación con entidades públicas, usar inteligencia artificial explicable para interoperar y detectar a tiempo el fraude, asegurando base legal, trazabilidad y revisión humana?

Para lograrlo, es necesario diseñar un plan de acción claro que contemple recursos técnicos, capacitación especializada y colaboración interinstitucional. Solo así será posible construir un sistema más seguro, eficiente y confiable, que proteja la seguridad social de todos los colombianos y asegure la sostenibilidad del régimen pensional a largo plazo.

En este sentido, el presente estudio tiene como propósito diseñar un modelo de gestión pública innovador para Colpensiones que integre inteligencia artificial explicable (XAI) en sus procesos institucionales, promoviendo la interoperabilidad y la transparencia administrativa como mecanismo para prevenir el fraude, optimizar los recursos públicos y fortalecer la confianza ciudadana en el sistema pensional.

Para alcanzar este objetivo general, se plantean los siguientes objetivos específicos:

1. Diagnosticar las falencias tecnológicas, normativas y de gobernanza digital que afectan la interoperabilidad entre Colpensiones y las entidades públicas relacionadas, con el fin de identificar las principales brechas que limitan la detección preventiva del fraude.
2. Analizar e identificar los modelos y herramientas de inteligencia artificial explicable (IAX), adecuados para la detección del fraude en el sistema pensional colombiano.

3. Proponer un plan de implementación del modelo de inteligencia artificial explicable (XAI) que garantice su sostenibilidad técnica, institucional y ética, mediante la integración de recursos tecnológicos, la capacitación del talento humano, la interoperabilidad entre entidades públicas y la adopción de políticas de protección de datos y gobernanza digital en cumplimiento de la Ley 2381 de 2024.

Diagnóstico del Problema

Este diagnóstico revela por qué Colpensiones enfrenta dificultades para detectar y prevenir el fraude pensional: falta de interoperabilidad entre entidades y baja calidad de los datos. Estas fallas limitan la validación de información, el análisis de casos complejos y la detección de inconsistencias. El análisis combina documentos, cifras, procesos y proyecciones para ofrecer una visión completa del problema.

Diagnóstico de las falencias tecnológicas y normativas que limitan la interoperabilidad y el análisis integral de casos de fraude en Colpensiones

Actualmente, Colpensiones se apoya en un conjunto de controles administrativos y legales para la prevención del fraude (la Ley 797 de 2003 Art. 19, Sentencia SU 182 de 2019, circulares internas, Reporte de Ético, Gerencia de Prevención del Fraude). Estos mecanismos, si bien son un punto de partida, son principalmente reactivos, es decir, detectan el fraude una vez que ha ocurrido o durante el proceso de verificación. La principal limitación tecnológica es la falta de interoperabilidad entre las bases de datos de Colpensiones y otras entidades públicas. Esto significa que la información de los afiliados no se cruza de manera eficiente con datos de otras fuentes, como la Registraduría Nacional del Estado Civil, el Ministerio de Salud, la UGPP (Unidad de Gestión Pensional y Parafiscales), u otras entidades, lo que dificulta la identificación de inconsistencias y duplicidades.

De acuerdo con el boletín oficial de Colpensiones (marzo de 2025), la entidad administradora de pensiones Colpensiones cuenta con 6.966.002 afiliados, de los cuales el 52.8% son hombres y el 47.2% mujeres. Durante ese mes, se registraron 24,427 nuevas afiliaciones y 16,460 traslados desde fondos

privados, con un incremento neto del 0.2% frente al mes anterior. Sin embargo, apenas el 40.7% de los afiliados realizaron cotizaciones activas (2'837,647 personas), lo que evidencia una densidad de cotización insuficiente y una fragilidad estructural en la sostenibilidad financiera del sistema.

En este escenario, la entrada en vigencia de la Ley 2381 de 2024 plantea un reto adicional, dado que el volumen de trámites (afiliaciones, traslados, reconocimientos de pensión, verificación de cotizaciones) aumentará de manera significativa, presionando aún más los sistemas tecnológicos y de control.

La principal brecha tecnológica a cerrar es la integración de datos a través de la implementación de un sistema de inteligencia artificial. Este sistema no sólo conectaría las bases de datos internas, sino que también establecería protocolos de comunicación segura con entidades externas para realizar validaciones en tiempo real.

En cuanto a la brecha normativa, la Ley 2381 de 2024 abre la puerta a esta transformación. Es crucial interpretar y aplicar la Ley para crear los marcos normativos que permitan el intercambio de información entre entidades, garantizando la protección de datos personales y el debido proceso.

En términos de impacto financiero, se estima que el fraude en el sistema pensional podría incrementarse en un 10% y un 15% con la implementación de la nueva ley, debido a la posible manipulación de la información para agilizar procesos o el aprovechamiento de lagunas en la interoperabilidad. Finalmente, la implementación de este modelo no solo reducirá el fraude, sino que también mejorará la eficiencia operativa de Colpensiones, disminuyendo los tiempos de respuesta a las peticiones presentadas por los beneficiarios legítimos.

Del análisis realizado, se observa que dos de las grandes falencias que enfrenta actualmente Colpensiones: La primera, es la ineficiente interoperabilidad interinstitucional y la segunda, la baja calidad en el suministro de datos.

La ineficiente interoperabilidad interinstitucional en Colpensiones constituye una de los riesgos más críticos en la gestión de la seguridad social en Colombia. Este problema se origina en la fragmentación de los sistemas de información de entidades como la Unidad de Gestión Pensional y Parafiscales (UGPP), la Dirección de Impuestos y Aduanas Nacionales (DIAN), la Registraduría Nacional,

las empresas promotoras de salud (EPS) y los empleadores, lo que genera retrasos en la validación de datos, duplicidad de trámites y ausencia de controles oportunos. Entre los factores que intensifican el riesgo se encuentran la carencia de estándares de intercambio de información, la dependencia de procesos manuales en la verificación de documentos, la baja calidad de los registros suministrados por terceros y las limitaciones tecnológicas derivadas de plataformas heredadas que no permiten integraciones en tiempo real. Estas condiciones facilitan que se presenten fraudes relacionados como por ejemplo la falsificación de semanas cotizadas, la suplantación de identidad y el reconocimiento indebido de prestaciones, lo que conlleva a un detrimento patrimonial.

Desde una perspectiva estadística y económica, resulta evidente que esta falencia representa un riesgo financiero considerable. Con base en la estimación de Colpensiones de pérdidas anuales por fraude de 1.2 billones de pesos (Colpensiones 2019), es posible proyectar que entre un 10% y un 40% de ese monto podría ser atribuible a problemas de interoperabilidad, lo que equivale a pérdidas anuales de entre 120 mil millones y 480 mil millones de pesos. Bajo un escenario medio del 25%, esta cifra alcanzaría aproximadamente 300 mil millones de pesos anuales. Si se proyecta dicho valor a cinco años, asumiendo un crecimiento conservador del 5% anual, las pérdidas acumuladas ascenderían a 1.65 billones de pesos; y bajo un escenario más pesimista, con un crecimiento del 10% anual, superarían los 1.83 billones de pesos. Estas cifras ilustran la magnitud del impacto económico de la falencia y la urgencia de implementar mecanismos preventivos que mitiguen los riesgos asociados.

Ante este panorama, se evidencia la necesidad de desarrollar un modelo preventivo basado en inteligencia artificial que permita cerrar la brecha de interoperabilidad. Este modelo debería fundamentarse en la integración de fuentes de datos mediante Interfases de Programación de Aplicaciones (APIs, en inglés) seguras, la creación de un repositorio centralizado con estándares de calidad y la implementación de algoritmos de detección de anomalías que identifiquen patrones sospechosos en tiempo real. El uso combinado de técnicas de aprendizaje supervisado y no supervisado permitiría no solo detectar fraudes conocidos, sino también descubrir nuevas modalidades de irregularidades a través del análisis de grandes volúmenes de información. Asimismo, la incorporación de análisis de grafos facilitaría el rastreo de redes de suplantación y conexiones entre empleadores, afiliados y cotizaciones anómalas.

La aplicación de este enfoque generaría beneficios económicos sustanciales. Bajo el escenario medio proyectado, un modelo preventivo con inteligencia artificial tendría la capacidad de reducir entre

un 50% y un 70% de las pérdidas atribuibles a la falta de interoperabilidad, lo que representaría un ahorro estimado de entre 828 mil millones y 1.16 billones de pesos en un periodo de cinco años. Más allá del impacto financiero, este modelo fortalecería la gobernanza institucional, aumentaría la transparencia en el reconocimiento de derechos pensionales y consolidaría la confianza de los ciudadanos en la capacidad del Estado para proteger los recursos públicos.

La segunda falencia que se encontró es la baja calidad en el suministro de datos, la cual constituye un obstáculo estructural para la gestión eficiente de Colpensiones, pues afecta directamente la confiabilidad de la información utilizada para el reconocimiento de derechos pensionales. Entre los factores que intensifican este riesgo se encuentran los registros incompletos de semanas cotizadas, errores en las fechas de vinculación y retiro laboral, duplicación de aportes y la falta de homologación en los formatos de reporte entre empleadores, administradoras de fondos y entidades estatales. Así mismo la baja calidad en el suministro de datos se evidencia en documentación e información reportada en las oficinas y páginas oficiales de las entidades públicas, Registros civiles incompletos, errores en fechas de cotización, fechas de nacimiento, declaraciones extrajuicio, duplicidad, falta de información de afiliación en salud de los afiliados y núcleo familiar, actualización de datos, entre otros.

A lo anterior se suma la existencia de historiales laborales fragmentados, producto de la informalidad en el mercado laboral y de la insuficiente capacidad de control sobre la calidad de los datos reportados por terceros. Esta situación incrementa la probabilidad de errores en la liquidación de pensiones, genera litigios costosos para la entidad y abre espacios para fraudes a través de la manipulación o falsificación de registros.

El impacto económico de esta falencia puede estimarse a partir del volumen de inconsistencias detectadas en los procesos de verificación de semanas y cotizaciones. Considerando que Colpensiones gestiona más de 1.7 millones de pensionados y alrededor de 6.6 millones de afiliados activos, incluso un margen de error del 3% en la calidad de datos podría afectar a más de 200,000 registros. Si se asume que, en promedio, un error en los historiales laborales puede derivar en litigios o reconocimientos indebidos de prestaciones con un costo estimado desde los 5 millones de pesos por caso, las pérdidas potenciales anuales superarían el 1 billón de pesos. Bajo un escenario de persistencia del problema y con un crecimiento proyectado del 5% anual en el número de afiliados y pensionados, las pérdidas acumuladas a cinco años podrían ubicarse entre 5.5 y 6 billones de pesos, dependiendo de la magnitud de los errores no corregidos.

La magnitud del riesgo descrito hace evidente la necesidad de un modelo preventivo soportado en inteligencia artificial, orientado a mejorar la calidad de los datos y garantizar la trazabilidad de la información. Este modelo debería incluir sistemas de validación automática de registros en el momento de su ingreso, algoritmos de detección de inconsistencias y duplicados, y mecanismos de reconciliación inteligente que contrastan datos entre entidades mediante reglas de negocio predefinidas. El uso de técnicas de *machine learning* permitiría identificar patrones de error recurrentes y anticipar fallas en la carga de información, mientras que los modelos de procesamiento de lenguaje natural (NLP) podrían ser aplicados a documentos no estructurados para extraer y estandarizar información relevante.

En términos de beneficios, un modelo de este tipo podría reducir entre un 40% y un 60% las inconsistencias actuales, lo que representaría un ahorro de entre 2 y 3.6 billones de pesos en cinco años. Adicionalmente, la reducción en el número de litigios y procesos judiciales no solo generaría eficiencia económica, sino también mayor confianza en el sistema pensional y en la capacidad del Estado de administrar de forma transparente y justa los aportes de los ciudadanos.

El efecto directo de estas falencias se refleja en el incremento sostenido de los casos de fraude dentro de Colpensiones, afectando la sostenibilidad financiera del régimen de prima media y la confianza ciudadana en la administración de los recursos públicos. De acuerdo con estimaciones institucionales, en el año 2019 las pérdidas asociadas a prácticas fraudulentas ascendían a aproximadamente 1.2 billones de pesos anuales; sin embargo, para el año 2025 esta cifra se proyecta entre los 1.32 y 1.38 billones de pesos, lo cual constituye una alerta institucional que justifica la adopción urgente de estrategias basadas en inteligencia artificial explicable, trazabilidad digital y revisión humana supervisada, en el marco de la Ley 2381 de 2024 y los principios de transparencia y responsabilidad fiscal.

Metodología

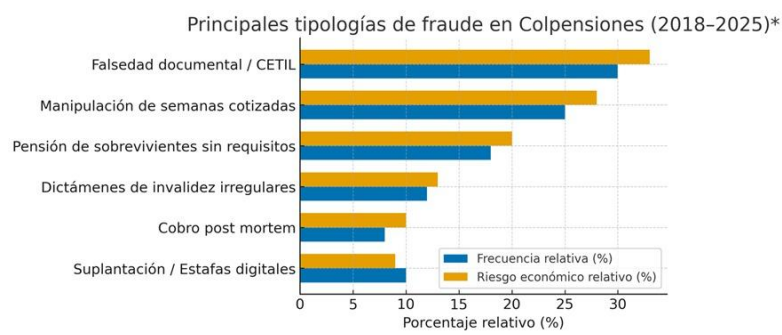
Ante las falencias detectadas en la gestión del fraude (estudio de caso) y, mediante una metodología deductiva se identificó que Colpensiones requiere un modelo que combine innovación tecnológica y eficiencia institucional. La inteligencia artificial explicable (XAI) surge como una herramienta clave para fortalecer la transparencia, la trazabilidad y la toma de decisiones basadas en evidencia.

Identificación de modelos y herramientas de inteligencia artificial explicable aplicables a la detección y prevención de fraude en el sector pensional.

Para el año 2018, la Administradora Colombiana de Pensiones, Colpensiones, informó que se habían interpuesto 691 denuncias por presuntos casos de fraude al régimen pensional de prima media, en los que están indiciadas 886 personas. Presuntamente, las acciones fraudulentas de los denunciados habrían afectado a más de 2,690 ciudadanos, de igual forma se debe dejar claro que del año 2019 a 2025, Colpensiones no publica estadística consolidada y abierta exclusivamente sobre fraude (conteo anual de denuncias, revocatorias por fraude y montos evitados/recuperados). Sin embargo, existen repositorios oficiales donde suelen quedar los insumos para verificar y completar la información: rendición de cuentas anual, informe de PQR y denuncias, alertas recientes, toda esta información que se puede observar en la página oficial de la entidad.

En el sector pensional, Colpensiones se enfrenta a distintas modalidades de fraude. Algunas casuísticas identificadas son: Declaraciones extrajuicio para acreditar convivencia, la suplantación de identidad, el cobro de auxilio funerario, incremento injustificado de semanas, incremento injustificado de pérdida de capacidad laboral, dictámenes no emitidos por la entidad competente, certificación electrónica de tiempos laborados (CETIL), Calculo actuarial, registro civil de nacimiento para acreditar parentesco, duplicidad o incompatibilidad de prestaciones entre otros.

Figura 1. Principales Tipología de fraude



Fuente: Elaboración propia con base en datos de Colpensiones (2018–2025), DNP y UGPP (2023). *Estimaciones realizadas a partir de reportes institucionales y bases de casos de fraude detectados por Colpensiones, UGPP y DNP (2018–2025).*

Tabla 1. Posibles tipologías de fraude y su descripción

Tipología de fraude	Descripción general	Ejemplo práctico	Riesgo institucional asociado
Falsedad documental / CETIL	Uso de documentos falsos o alterados (registros civiles, certificados laborales, historias clínicas) para obtener pensiones o prestaciones.	Presentar un registro civil adulterado para acreditar semanas o parentesco.	Reconocimiento indebido de pensión y pérdida de recursos del fondo común.
Manipulación de semanas cotizadas	Alteración de los datos laborales o de aportes en la historia laboral con el fin de cumplir requisitos de pensión.	Incluir certificaciones de empresas inexistentes o modificar periodos de cotización.	Afectación de la trazabilidad del sistema y fraude al RPM.
Pensión de sobrevivientes sin requisitos	Solicitud del beneficio por parte de personas que no cumplen los requisitos legales de dependencia económica o convivencia.	Declarar falsamente convivencia o dependencia para recibir sustitución pensional.	Pago indebido y vulneración del derecho de los beneficiarios reales.
Dictámenes de invalidez irregulares	Emisión de dictámenes médicos fraudulentos o sin soporte técnico que determinan invalidez no real.	Certificar una discapacidad inexistente para obtener pensión anticipada.	Aumento del gasto pensional y pérdida de confianza en los procesos médicos.
Cobro post mortem	Continuación del cobro de pensiones o auxilios tras el fallecimiento del titular.	Familiares o terceros que no reportan la muerte y siguen recibiendo el pago.	Detrimento patrimonial y fallas en interoperabilidad con la Registraduría.
Suplantación / Estafas digitales	Acceso no autorizado a información o suplantación de identidad mediante medios digitales o phishing.	Suplantar a un pensionado para redirigir pagos o alterar datos bancarios.	Riesgo cibernético, fuga de información y pérdida de confianza ciudadana.

Fuente: Elaboración propia con base en datos de Colpensiones (UGPP & DNP, 2023–2024)

Para cada tipo, se propuso un modelo utilizando técnicas de explicación algorítmica de la comunidad de Inteligencia Artificial Explicable (XAI). Estos son herramientas matemáticas o modelos auxiliares que no toman decisiones, ni actúan automáticamente para prevenir fraudes; su trabajo se basa en operar con el modelo de detección de fraude preventivo.

Para lograr esto, podemos usar estas tres herramientas algorítmicas como: SHAP, LIME, Anchors.

En este sentido SHAP es el encargado de justificar por qué tomó esa decisión; LIME puede demostrar qué hizo su modelo en ese caso; Anchors puede crear reglas transparentes: si A y B, entonces el modelo predecirá fraude.

Por lo cual, podemos decir que SHAP, LIME y Anchors son los "intérpretes" o "explicadores" en un agente de inteligencia artificial (IA); estos pueden integrarse en agentes inteligentes o sistemas de inteligencia artificial (IA), de manera que estos agentes se vuelvan más transparentes, auditables y comprensibles. Un ejemplo aplicado a Colpensiones:

El agente de inteligencia artificial (IA) antifraude procesa miles de solicitudes de pensiones. Este agente ha usado modelos predictivos (aprendizaje automático) para la detección de anomalías, en este caso de que el agente marque un caso como sospechoso, un módulo Artificial Explicable XAI (SHAP, LIME o Anchors) explica su razón al funcionario/auditor humano. Por lo tanto, el sistema no solo decide, sino que puede rendir cuentas, y para verificación y trazabilidad, en cumplimiento con los principios de gobernanza algorítmica y transparencia pública.

La inteligencia artificial no puede ser una "caja negra" para los ciudadanos, y a que nos referimos cuando hablamos de cajas negras, nos referimos a modelos de Inteligencia Artificial que generan resultados sin que sea posible entender fácilmente cómo llegaron a esa conclusión, por lo tanto, toda alerta de fraude debe estar sustentada en razones claras y comprensibles, de manera que el ciudadano en caso de solicitar una revisión pueda entender por qué se generó una alerta de fraude. Aquí es donde entran en juego las herramientas de Inteligencia Artificial explicable (XAI), como SHAP, LIME o Anchors, que permiten mostrar y evidenciar cuáles variables fueron determinantes en la decisión. Estas herramientas generan explicaciones que luego se documentan en un "paquete de evidencia" para acompañar cualquier actuación administrativa.

De esta forma, se garantiza el respeto al debido proceso (Ley 1437 de 2011), la protección de datos personales (Ley 1581 de 2012) y la revocatoria directa establecida en el artículo 19 de la Ley 797 de 2003. Esto no solo protege a los beneficiarios legítimos, sino que también fortalece la sostenibilidad del régimen pensional al cerrar la brecha de fraude.

Cada alerta generada por el sistema debe producir un conjunto de documentos estandarizados. Este paquete de evidencia incluye: una ficha del caso con metadatos básicos; la explicación de las variables clave que influyeron en la alerta; la comparación con casos similares; un contrafactual que muestre qué condiciones hubieran evitado la alerta; y las trazas de interoperabilidad, es decir, de dónde provino la información y en qué momento fue consultada. Todo esto sirve para sustentar el acto administrativo y garantizar transparencia, ayudando a la Gerencia de prevención del fraude para emitir los autos de decisión que posteriormente serán remitidos a la Dirección de prestaciones económicas para su decisión final.

Se debe diseñar plantillas con todo el lleno legal y jurídico que permita a Colpensiones comunicar al ciudadano, de forma clara y garantista, las decisiones preliminares. Primero, se prevé una comunicación de apertura de verificación preventiva, donde se informa al solicitante que su caso presenta señales de riesgo y que se requiere información adicional y la apertura a etapa probatoria para validar la sospecha existente sin que esto implique una decisión de fondo. En segundo lugar, está el requerimiento de información, dirigido al ciudadano o a otras entidades, con base en las facultades legales de interoperabilidad y por esto la necesidad urgente de una cooperación interadministrativa que nos permita en tiempo real la verificación de los datos de cada afiliado. Tercero, se emita un acto administrativo motivado de alerta preventiva, en el que se detalle los hechos, las pruebas y las explicaciones legales, dejando constancia de que hubo valoración humana antes de adoptar la medida. Finalmente, se establece una plantilla para el cierre del caso, ya sea por un auto de decisión o un auto de archivo.

Aunque la inteligencia artificial ayuda a priorizar y detectar casos, las decisiones no son automáticas. Siempre se requiere la intervención de un analista antes de comunicar cualquier decisión. Este enfoque “humano” garantiza que los resultados técnicos sean interpretados a la luz de la norma y que los ciudadanos conserven su derecho de defensa. Asimismo, cada paso queda documentado en una investigación administrativa especial, lo que permite auditorías futuras.

Al aplicar este modelo, Colpensiones no solo detecta fraudes consumados, sino que actúa de manera preventiva, reduciendo pérdidas antes de que se materialicen. La institución gana en eficiencia y transparencia, los ciudadanos tienen mayores garantías jurídicas y el sistema pensional en su conjunto

se fortalece. Además, la adopción de este enfoque es coherente con la Ley 2381 de 2024, que promueve la modernización y sostenibilidad del sistema de protección social en Colombia.

La inteligencia artificial explicable no es un fin en sí mismo, sino un medio para reforzar la legitimidad de las decisiones administrativas, blindar los recursos del sistema pensional y garantizar los derechos de los afiliados.

Tabla 2. Comparativa de modelos y herramientas de Inteligencia Artificial explicable (XAI) aplicadas a la detección de fraude en Companions.

Elemento	Explicación jurídica-técnica	Ejemplos aplicados a Colpensiones
¿Qué busca este objetivo?	Reconocer qué modelos de IA pueden detectar fraudes y, al mismo tiempo, explicar claramente sus resultados para garantizar el debido proceso.	Modelos que no sean “cajas negras”, sino que permitan mostrar pruebas y razones al ciudadano.
Modelos recomendados	- Árboles de decisión / Reglas: fáciles de interpretar. - EBM (<i>Explainable Boosting Machine</i>): precisión con explicaciones claras. - <i>XGBoost</i> / <i>LightGBM</i> : alta detección, con explicaciones vía SHAP. - Modelos de anomalías: detectan conductas inusuales. - Análisis de grafos: para redes de colusión.	Ej.: detectar aportes ficticios en empresas fachada, o identificar vínculos entre tramitadores y afiliados.
Herramientas XAI (explicabilidad)	- SHAP / LIME / Anchors: muestran variables clave en cada caso. - <i>InterpretML</i> (EBM): <i>dashboard</i> con explicaciones globales y locales. - AIX360 / Alibi: contrafactuales y reglas para argumentar alertas. -	Permiten entregar al ciudadano un informe claro: “su caso fue marcado porque las semanas reportadas en 2023 son inconsistentes con su historial anterior”.

	<i>Captum</i> (PyTorch): para redes neuronales.	
Fundamento jurídico	Ley 1437/2011: obligación de motivar las decisiones administrativas. - Ley 1581/2012: protección de datos personales. - Ley 2381/2024: modernización y sostenibilidad del sistema pensional. Ley 797/2003: Revocatoria directas pensiones con indicios de fraude	Garantizar que las alertas de fraude respeten el debido proceso, los derechos de defensa y la protección de la información personal.
Valor agregado	- Transparencia: cada alerta se justifica con evidencia. - Prevención: el sistema actúa antes de que se materialice el fraude. - Seguridad jurídica: decisiones motivadas y auditables. - Confianza ciudadana: el afiliado entiende por qué fue requerido.	Colpensiones fortalece la sostenibilidad del régimen pensional y protege a los verdaderos beneficiarios.

Fuente: Elaboración propia con base en la Ley 2381 de 2024, Ley 1581 de 2012, y análisis comparativo de herramientas XAI (BID, 2023; OCDE, 2023; DNP, 2024).

Resultados y discusión

Teniendo en cuenta que para el año 2019 las pérdidas por fraude ascendieron a 1.2 billones de pesos, y considerando que la entrada en vigencia de la Ley 2381 de 2024, podría incrementar de un 20% a un 30% el volumen de trámites relevantes, para lo que se proyecta un aumento proporcional del riesgo de fraude entre 20% y 40%.

Esto se traduciría en pérdidas anuales estimadas entre 1.44 billones (20% de incremento) y 1,68 billones (40% de incremento) de pesos. Dichas cifras podrían ser aún mayores al incorporar factores como inflación, incremento salarial, complejidad normativa y nuevas modalidades de fraude.

Los factores que intensifican el riesgo son: (i) mayor volumen de afiliados y traslados: incremento de la superficie de riesgo, (ii) cambios normativos: nuevas reglas pueden ser aprovechadas por actores fraudulentos, y (iii) falta de validaciones automatizadas: la ausencia de Inteligencia Artificial impide detectar patrones sospechosos en tiempo real.

La falencia tecnológica y de control genera un escenario donde las pérdidas anuales podrían superar los 1.7 billones de pesos si no se adoptan mecanismos preventivos avanzados.

El costo económico no solo se mide en fraudes directos, sino también en:

- Costos administrativos derivados de investigaciones y trámites de corrección. (Gerencia de prevención del fraude / Cosinte, Mac Gregor).
- Pérdida de confianza de los afiliados hacia Colpensiones.
- Riesgo reputacional y político frente a la viabilidad de la reforma pensional.

Ante este panorama, se justifica la implementación de un modelo preventivo basado en inteligencia artificial (IA), que permita:

1. Detectar fraudes de manera temprana mediante algoritmos de aprendizaje supervisado y no supervisado.
2. Mejorar la calidad de datos aplicando técnicas de limpieza, validación y estandarización automática.
3. Asegurar la interoperabilidad a través de soluciones de integración inteligente con otras entidades públicas.
4. Optimizar los recursos tecnológicos incorporando sistemas escalables y adaptados al nuevo esquema normativo.

Este modelo contribuiría a la sostenibilidad financiera del sistema pensional, así como a la eficiencia administrativa y transparencia institucional, elementos indispensables para el éxito de la Ley 2381 de 2024.

Riesgos institucionales y estrategias de mitigación

El análisis institucional realizado permitió reconocer diversos riesgos que podrían incidir en la implementación del modelo preventivo de detección de fraude basado en inteligencia artificial explicable (XAI). Estos riesgos no solo están asociados a los desafíos tecnológicos y normativos, sino también a aspectos éticos, organizacionales y de gestión del cambio que acompañan toda transformación digital en el sector público. Identificarlos oportunamente es fundamental para anticipar posibles barreras y fortalecer las capacidades institucionales de Colpensiones. A continuación, se presentan los principales riesgos detectados y las estrategias planteadas para mitigarlos, asegurando que el modelo mantenga su sostenibilidad técnica, legal y humana en el tiempo.

Tabla 3. *Identificación, descripción y mitigación del riesgo*

Tipo de riesgo	Descripción	Estrategia de mitigación
Riesgo tecnológico	Limitaciones en la infraestructura digital o fallas de interoperabilidad entre sistemas institucionales.	Implementar APIs seguras y estándares del Marco de Interoperabilidad del Estado (MinTIC, 2023), junto con planes de mantenimiento preventivo y respaldo de datos.
Riesgo normativo	Ambigüedades legales frente al uso de inteligencia artificial y tratamiento de datos personales.	Ajustar el modelo a la Ley 2381 de 2024 y a los lineamientos del DNP y MinTIC, garantizando revisión jurídica y trazabilidad en cada fase.
Riesgo ético	Decisiones automatizadas sin explicación clara o con sesgos algorítmicos.	Aplicar principios de IA explicable (XAI) con revisión humana obligatoria y auditorías periódicas del sistema.
Riesgo de gobernanza	Falta de coordinación entre las áreas internas de Colpensiones y las entidades externas aliadas.	Crear un Comité Interinstitucional de Gobernanza Digital, que supervise interoperabilidad, seguridad y cumplimiento ético.
Riesgo de cambio organizacional	Resistencia del personal a adoptar nuevas tecnologías o procesos automatizados.	Implementar un plan de capacitación continua y gestión del cambio, basado en liderazgo participativo y comunicación institucional.
Riesgo de ciberseguridad	Posibles ataques o filtraciones de datos personales.	Fortalecer la seguridad mediante protocolos de cifrado, autenticación multifactor y cumplimiento de la Ley 1581 de 2012 sobre protección de datos.

Fuente: Elaboración propia con base en el Plan Estratégico de Tecnologías de la Información – PETI (Colpensiones, 2023), la Política de Gobierno Digital (MinTIC, 2023) y los lineamientos de gestión del riesgo en el sector público (DNP, 2024).

Solución innovadora: modelo de inteligencia artificial que integre recursos técnicos, capacitación y protocolos de protección de datos, conforme a la Ley 2381 de 2024 y otras normativas vigentes

Integración de la interoperabilidad de la inteligencia artificial en Colpensiones para la detección del fraude

En el marco de la modernización del Estado colombiano y la ejecución de la Ley 2381 de 2024, que transforma de manera estructural el sistema pensional, resulta imperativo que las entidades públicas incorporen tecnologías avanzadas que garanticen procesos eficientes, seguros y transparentes. En este contexto, Colpensiones, como administradora del régimen de prima media, enfrenta el reto más relevante del sector: fortalecer sus mecanismos de verificación, control y prevención del fraude. Para ello, es necesario que el Plan Estratégico de Tecnologías de la Información (PETI 2023-2026) incluya la inteligencia artificial (IA) como un componente estratégico que permita automatizar tareas, generar interoperabilidad entre entidades públicas y consolidar una infraestructura tecnológica sustentada en principios de legalidad, ética y protección de datos personales.

La inteligencia artificial adquiere un papel esencial en la prevención del fraude y la validación de información en los procesos de reconocimiento pensional. Las modalidades de fraude como la suplantación de identidad, la falsificación de documentos o la manipulación de semanas cotizadas comprometen la sostenibilidad financiera del sistema, erosionan la confianza social y generan pérdidas significativas al Estado. Frente a ello, la adopción de un modelo de Inteligencia Artificial explicable (XAI) capaz de analizar patrones, identificar riesgos y justificar cada decisión ofrece una alternativa innovadora y responsable para apoyar la labor de los analistas humanos. Este enfoque no pretende reemplazar al funcionario público, sino asistirlo en la toma de decisiones informadas, trazables y auditables, reforzando la transparencia institucional y la rendición de cuentas.

El desarrollo del modelo “Inteligencia Artificial Preventiva (XAI)” se sustenta en un marco jurídico y político robusto. La Ley 2381 de 2024 establece la obligación de garantizar la sostenibilidad y el control del gasto público mediante la adopción de tecnologías digitales seguras. La Ley 1581 de 2012 y el Decreto 1377 de 2013 definen las condiciones para el tratamiento legítimo de datos personales; el Modelo de Seguridad y Privacidad de la Información (MSPI) regula la gestión y trazabilidad de los datos institucionales; y el Marco de Interoperabilidad del Estado Colombiano (2023) establece los lineamientos técnicos, semánticos y organizacionales para el intercambio seguro de información entre entidades. En conjunto, estas normas conforman un ecosistema legal habilitante para el uso ético y

responsable de la inteligencia artificial en la administración pública, especialmente en procesos de detección y mitigación del fraude.

En los últimos años, Colombia ha dado pasos firmes hacia una regulación responsable y ética del uso de la inteligencia artificial. El Documento CONPES 3975 de 2019, que define la *Política Nacional para el Desarrollo y la Adopción de la Inteligencia Artificial*, marcó un punto de partida al promover una IA centrada en las personas, basada en la transparencia, la responsabilidad y la seguridad. Posteriormente, el Marco Ético de la Inteligencia Artificial (MinTIC, 2022) reforzó estos principios al establecer lineamientos de equidad, trazabilidad y supervisión humana para el uso de sistemas automatizados en el sector público. En esta misma línea, el país avanza actualmente con la Estrategia Nacional de Inteligencia Artificial 2024–2030, liderada por el MinTIC y el DNP, que introduce el concepto de una “IA confiable”, en armonía con las recomendaciones de la OCDE y la UNESCO. Este conjunto de políticas y marcos normativos ofrece un entorno sólido para que entidades como Colpensiones puedan incorporar la inteligencia artificial explicable (XAI) de forma responsable, asegurando el cumplimiento de la ley, la protección de los datos personales y el respeto por los derechos de los ciudadanos en cada proceso.

Aunque Colpensiones ya forma parte del Marco de Interoperabilidad del Estado Colombiano (2023), este instrumento no está diseñado para detectar fraudes por sí mismo, sino para facilitar el intercambio seguro de datos entre instituciones públicas mediante estándares, APIs y gobernanza de la información. Sin embargo, esta infraestructura puede potenciarse con herramientas de IA preventiva capaces de procesar los datos interoperables y generar alertas sobre inconsistencias o patrones de riesgo. Por ejemplo:

- Verificación de identidad: comparación automática con bases de la Registraduría Nacional a través del servicio de Autenticación Digital.
- Cruces de elegibilidad: contraste de información con registros de ADRES, UGPP o DANE para detectar duplicidades o inconsistencias en semanas cotizadas y beneficiarios.

Para implementar estas capacidades antifraude entre entidades, deben cumplirse buenas prácticas basadas en la normatividad vigente: (1) definir la base jurídica de cada cruce mediante convenios y finalidades explícitas; (2) aplicar la minimización de datos, solicitando solo la información necesaria para validar un derecho; (3) diseñar servicios interoperables con contratos de API, catálogos

semánticos, niveles de servicio y registros de auditoría, en concordancia con el Modelo de Seguridad y Privacidad de la Información (MSPI); y (4) evaluar riesgos y planes de seguridad antes de cualquier intercambio de datos. De esta forma, la detección del fraude no se realiza dentro del Marco mismo, sino mediante servicios analíticos e inteligencia artificial que operan sobre su infraestructura, cumpliendo con la Ley 1581 de 2012, el Modelo de Seguridad y Privacidad de la Información (MSPI) y la Plataforma de Interoperabilidad del Estado (PDI) y Servicios Ciudadanos Digitales (SCD), como infraestructura oficial.

El modelo de implementación de IA en Colpensiones debe alinearse con un enfoque de gobernanza digital que articule tres dimensiones fundamentales:

Recursos técnicos, que aseguren la interoperabilidad, trazabilidad y seguridad de la información.

Capacitación institucional, que permita al talento humano comprender, interpretar y supervisar los resultados generados por la inteligencia artificial (IA).

Protocolos de protección de datos, en cumplimiento con las leyes y marcos regulatorios nacionales.

La integración de la inteligencia artificial en Colpensiones no debe concebirse como un proyecto aislado, sino como un proceso de innovación institucional que transforme la gestión pública hacia una cultura de análisis basado en evidencia, donde la transparencia y la ética algorítmica sean ejes transversales. Bajo el enfoque de Inteligencia Artificial explicable (XAI), cada decisión automatizada podrá ser entendida, auditada y revisada por un funcionario humano, lo que refuerza la confianza ciudadana y la legitimidad institucional.

Incorporar la inteligencia artificial dentro del Plan Estratégico de Tecnologías de la Información (PETI 2023–2026) permitirá a Colpensiones avanzar hacia una gestión pública basada en evidencia, interoperabilidad y responsabilidad digital, consolidando su papel como entidad líder en innovación pública. Esta integración fortalecerá la capacidad institucional para optimizar procesos, proteger los recursos del sistema pensional y garantizar la transparencia administrativa, contribuyendo de manera directa a la sostenibilidad financiera y a la generación de confianza ciudadana en el sistema pensional Colombiano.

Finalmente, esta iniciativa se encuentra en coherencia con la Política de Gobierno Digital (MinTIC, 2022) y el Marco de Interoperabilidad 2023, los cuales promueven la adopción de tecnologías emergentes, como la IA y el análisis de datos, para fortalecer la gobernanza, la participación ciudadana y la eficiencia institucional. El modelo “Inteligencia Artificial Preventiva” se plantea como un prototipo conceptual y funcional que demuestra cómo una entidad pública puede integrar algoritmos explicables, trazabilidad digital y revisión humana bajo principios de ética, transparencia y responsabilidad algorítmica, contribuyendo a la construcción de un Estado digital confiable, preventivo y centrado en el ciudadano.

Figura 2. Marco de Interoperabilidad del Estado Colombiano para el Gobierno Digital



Fuente. Plataforma Única de Interoperabilidad — Adaptado del Marco de Interoperabilidad del Estado Colombiano para el Gobierno Digital (MinTIC, 2023).

Protocolo de Demostración — Inteligencia Artificial Preventiva (XAI) e Interoperabilidad

El agente sigue un flujo lógico y auditable, como si fuera un funcionario digital que revisa documentos con rigor y transparencia:

- **Ingreso y consentimiento:** Antes de analizar cualquier dato, la Inteligencia Artificial verifica que exista una base legal o consentimiento válido. No usa la información si no tiene autorización o un fin legítimo.
- **Identificación confiable:** El sistema cruza automáticamente la información del ciudadano con bases públicas como la Registraduría, Administrador de los Recursos del Sistema General de Seguridad Social en Salud ADRES, Departamento Nacional de Planeación SISBÉN o Departamento Administrativo Nacional de Estadística) DANE, simulando la interoperabilidad estatal.

Si los datos coinciden, asigna un puntaje de confianza (1.00 es el máximo). Si detecta inconsistencias, alerta al área antifraude.

- Validación documental: Los documentos son revisados digitalmente para garantizar su autenticidad mediante un código único o “huella digital” (hash SHA-256). (*Secure Hash Algorithm 256 bits*)

Así, el sistema puede confirmar si un archivo fue manipulado o falsificado

1. Evaluación de riesgo (IA + XAI):

La IA analiza variables como:

- Valor de la prestación,
- Semanas cotizadas, y
- Canal de solicitud (riesgo operativo)

Si el valor o las inconsistencias son altas, el riesgo aumenta.

Pero y aquí está la innovación el sistema explica por qué lo considera riesgoso. Usa técnicas como SHAP, LIME o Anchors, que muestran la contribución de cada variable (por ejemplo: “El valor de la prestación pesa un 55 % en el riesgo total”).

2. Decisión y revisión humana: Si el riesgo es bajo, el caso pasa a aprobación preliminar.

Si es medio o alto, el sistema etiqueta la decisión como “en revisión”, generando una alerta para que un funcionario humano verifique.

Este principio, llamado “*Human in the loop*”, garantiza que ningún algoritmo decida por sí solo sobre los derechos de una persona.

3. Auditoría total: Todo el proceso desde la verificación legal hasta la decisión final se registra en un archivo de auditoría JSON (trazabilidad técnica y jurídica). Esto permite rastrear quién hizo qué, cuándo y con qué resultado, fortaleciendo la transparencia institucional.

Protocolo de Demostración — Inteligencia Artificial Preventiva (XAI) e Interoperabilidad

Resultado de demostración del simulador

La matriz de poder e interés es un modelo útil para diferenciar a las partes interesadas en función de su influencia e interés en el desarrollo del modelo de inteligencia artificial preventiva (XAI). Con Colpensiones, las áreas de TI, prevención de fraudes y legal, ejercen un poder e interés muy alto, ya que son las que dirigen la ejecución, supervisión y control del sistema. Mientras tanto, el Ministerio de Trabajo, el Ministerio de Hacienda y la UGPP ocupan un papel estratégico para coordinar la definición de políticas, normativas regulatorias, así como en términos de compartir datos entre esos organismos. Complementariamente, están entidades como la Registraduría Nacional del Estado Civil, Administradora de los Recursos del Sistema General de Seguridad Social en Salud (ADRES) y la Superintendencia Financiera que se convierten en aliados técnicos, validando y permitiendo un intercambio seguro de información con respecto a la ley.

Desde el punto de vista social, los afiliados, jubilados y el control ciudadano se vuelven fundamentales porque son los beneficiarios directos de un sistema de pensiones más ético, transparente y preventivo. Su participación fortalece la confianza pública y valida el uso responsable de la inteligencia artificial preventiva en la gestión de derechos. Por último, las agencias de control y los organismos jurisdiccionales como la Fiscalía General de La Nación, la Contraloría General y la Corte Constitucional vigilan la legalidad y protección de los derechos fundamentales en procedimientos donde se utilizan tecnologías automatizadas.

Esta matriz es una contribución a la organización y fortalecimiento de la gestión estratégica de relaciones institucionales, dando prioridad a soluciones de comunicación, actividades de formación y prácticas de cooperación interinstitucional. Su objetivo es fortalecer un ecosistema para la gobernanza digital responsable, en la que la IA se convierta en un aliado del servicio público que favorezca la transparencia y que contribuya a proteger los recursos del sistema de pensiones colombiano en aplicación de la Ley 2381 de 2024 y las políticas nacionales de transformación digital.

Tabla 4. *stakeholders en el marco institucional del modelo de la Inteligencia Artificial Explicable (XAI)*

Stakeholder	Tipo	Nivel de Poder	Nivel de Interés
Colpensiones	Institucional	Muy alto	Muy alto

Ministerio de Trabajo	Gubernamental	Alto	Alto
Ministerio de Hacienda y Crédito Público	Gubernamental	Alto	Medio
Unidad de Gestión Pensional y Parafiscales (UGPP)	Institucional	Medio	Alto
Registraduría Nacional del Estado Civil	Institucional	Alto	Medio
ADRES / Ministerio de Salud	Institucional	Medio	Medio
Superintendencia Financiera de Colombia	Regulador	Alto	Medio
Área de TI de Colpensiones	Operativo	Muy alto	Muy alto
Gerencia de Prevención del Fraude	Operativo	Muy alto	Muy alto
Gerencia Jurídica y de Cumplimiento	Operativo	Alto	Alto
Científicos de datos y analistas de IA	Técnico	Medio	Alto
Afiliados y pensionados	Social	Medio	Muy alto
Proveedores tecnológicos / consultores externos	Técnico	Medio	Medio
Superintendencia de Industria y Comercio (SIC)	Regulador	Alto	Medio
Procuraduría / Contraloría	Control	Alto	Medio
Corte Constitucional	Judicial	Medio	Medio
Consejo de Estado	Judicial	Medio	Medio
Corte Suprema de Justicia	Judicial	Medio	Medio
Tribunales y Juzgados	Judicial	Medio	Medio
SENA / Universidades aliadas	Académico	Medio	Medio
Veedurías ciudadanas / Organizaciones de pensionados	Social	Bajo	Medio

Fuente. elaborada con base en Colpensiones, 2024, y DNP, 2022.

Proyecciones financieras y Retorno de la Inversión (ROI) de innovación

El presente análisis financiero proyecta los costos, beneficios y retorno sobre la inversión (ROI) del modelo Inteligencia Artificial preventiva (XAI) para la detección temprana de fraudes en el reconocimiento de prestaciones sociales de Colpensiones.

El enfoque integra variables técnicas, económicas y éticas, resaltando tanto la rentabilidad financiera directa como los beneficios intangibles asociados a la transparencia institucional, la eficiencia administrativa y la protección de los recursos públicos.

De acuerdo con estimaciones internas y auditorías de control, las pérdidas anuales por fraude en Colpensiones (2025), oscilan entre \$1,32 y \$1,38 billones derivadas principalmente de:

- falsificación documental y suplantación de identidad.
- errores en la validación de semanas cotizadas.
- reconocimientos indebidos de pensiones o auxilios funerarios.

Con la entrada en vigencia de la ley 2381 de 2024, que amplía coberturas y digitaliza trámites, se prevé un aumento del volumen de solicitudes de entre 25% y 35%, lo que podría incrementar el riesgo de fraude entre 20% y 40%, si no se implementan mecanismos tecnológicos de detección temprana.

Frente a este panorama, el modelo Inteligencia Artificial preventiva (XAI) se plantea como una inversión estratégica de sostenibilidad institucional, orientada a blindar el sistema frente a pérdidas multimillonarias y a fortalecer la confianza ciudadana.

Tabla 5. Estimación de inversión inicial (millones de COP)

Concepto	Descripción	Estimación (COP millones)
Infraestructura tecnológica	Adquisición y configuración de entornos híbridos (nube + local) para procesamiento seguro	9.500

Desarrollo de modelos IA	Diseño, entrenamiento y validación de modelos XAI antifraude (árboles, EBM, SHAP)	6.800
Integración e interoperabilidad	Conectores con UGPP, Registraduría, ADRES, MinTrabajo y MinHacienda mediante la PDI	4.200
Ciberseguridad y cumplimiento	Auditorías, cifrado de datos, control de accesos y alineación con el MSPI	2.700
Capacitación y gestión del cambio	Formación de equipos técnicos, jurídicos y analistas antifraude en ética e IA explicable	1.100
Total inversión inicial		24.300

Fuente. Elaboración propia con base en Colpensiones (2023), DNP (2023) y MinTIC (2023).

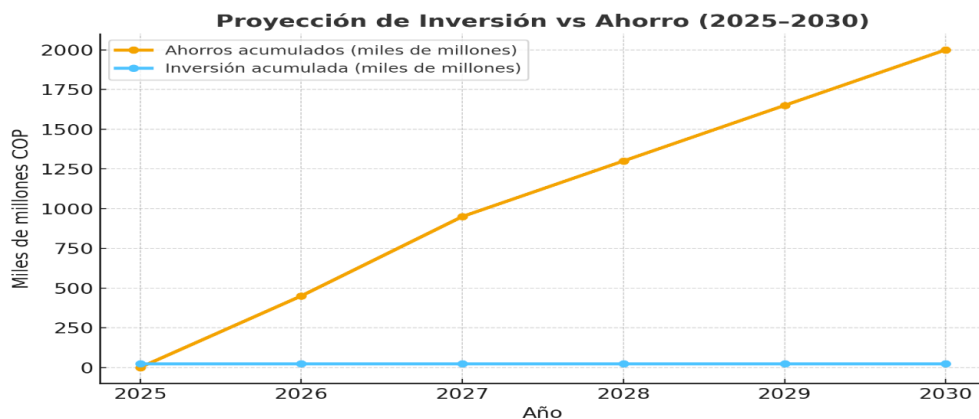
La inversión incluye tanto la infraestructura tecnológica y de datos, como los componentes de gobernanza y capacitación, asegurando que el sistema sea sostenible, seguro y ético desde su fase de implementación.

Proyecciones de Ahorro y Beneficio (5 años)

La puesta en marcha del modelo de Inteligencia Artificial explicable antifraude permitiría reducir entre 50% y 70% los fraudes detectados durante los primeros cinco años, lo que se traduce en ahorros acumulados estimados entre 4.1 y 5.7 billones de pesos.

Estos beneficios derivan de la disminución de pagos indebidos, el control automatizado de inconsistencias y la detección temprana de redes fraudulentas entre afiliados y terceros.

Figura 5. *Proyección de Inversión vs Ahorro (5 años)*



*La proyección incluye tanto los beneficios financieros directos (recursos no pagados por fraude) como los efectos secundarios positivos en eficiencia y reputación institucional.

Fuente. Proyección de Inversión vs Ahorro (2025–2030). Elaboración propia con base en DNP (2023), Colpensiones (2024) y BID (2023).

Tabla 6. Análisis e Interpretación de la Cálculo del ROI de innovación

Inversión total (5 años)	35.000 millones
Ahorros acumulados (escenario medio)	4.9 billones
ROI estimado	≈ 13.900%
Período de recuperación (Payback)	4.5 meses
Valor presente neto (VPN)	3.2 billones

Fuente. Elaboración propia con base en Colpensiones (2024), DNP (2023) y BID (2023).

Esta tabla resume el análisis costo-beneficio del proyecto de Interoperabilidad e Inteligencia Artificial Preventiva (XAI) para Colpensiones, diseñado bajo los lineamientos del Departamento Nacional de Planeación (DNP) y la Metodología General Ajustada (MGA).

- **Inversión total (5 años – 35.000 millones COP):**
Incluye el desarrollo e implementación del sistema Inteligencia Artificial explicable (XAI), infraestructura tecnológica, interoperabilidad con entidades aliadas (UGPP, Registraduría, ADRES, MinHacienda), fortalecimiento de ciberseguridad, formación de talento y mantenimiento operativo.

- Ahorros acumulados (escenario medio – 4.9 billones COP):
Corresponde al valor proyectado de pérdidas evitadas por fraudes y errores en el reconocimiento de pensiones durante el horizonte 2025-2030. Se calcula considerando una reducción conservadora del 25 % - 30% de fraudes, derivada de la automatización, trazabilidad y control de datos interinstitucionales.

ROI estimado ($\approx 13,900\%$):

El ROI se calcula con la fórmula estándar:

$$ROI = \frac{(\text{Beneficio Total} - \text{Costo de Inversión})}{\text{Costo de Inversión}} \times 100$$

Aplicando los datos de tu tabla:

$$ROI = \frac{(4,9 \text{ billones} - 0,035 \text{ billones})}{0,035 \text{ billones}} \times 100 = 13.900\%$$

Es decir:

- Inversión: \$35.000 millones COP
- Beneficio o ahorro acumulado: \$4,9 billones COP
- Resultado: retorno 139 veces superior al costo inicial.

Representa el retorno sobre la inversión, calculado como la relación entre el beneficio neto generado y el costo total del proyecto.

En este caso, por cada peso invertido, Colpensiones recupera aproximadamente 139 pesos en valor público, lo que demuestra una alta eficiencia fiscal y una rentabilidad institucional sobresaliente dentro del marco de innovación pública.

- Período de recuperación (*Payback* – 4.5 meses): Indica el tiempo estimado necesario para recuperar la inversión inicial. El corto periodo de retorno se explica por la magnitud de los ahorros que el sistema genera desde su fase piloto, especialmente en la reducción de pagos indebidos y fraudes detectados tempranamente.
- Valor Presente Neto (VPN – 3.2 billones): Expresa el valor actual de los beneficios futuros descontando la inversión inicial. Un VPN positivo y elevado demuestra que el proyecto genera valor agregado neto para el Estado, incluso considerando costos de oportunidad del dinero público.

Esta tabla se construye para demostrar cuantitativamente la viabilidad financiera y estratégica del proyecto de IA preventiva en Colpensiones.

Su propósito no es solo mostrar rentabilidad económica, sino evidenciar que la inversión en innovación tecnológica y gobernanza digital produce valor público tangible y medible.

En términos de gestión pública, esta tabla cumple tres funciones clave:

1. Soporte para la toma de decisiones institucionales: Proporciona a la alta dirección de Colpensiones y a los organismos de control (Departamento Nacional de Planeación, Ministerio de Hacienda, Contraloría General de la República) una evidencia clara de que la inversión cumple con los criterios de eficiencia, eficacia y sostenibilidad fiscal establecidos por la Ley 819 de 2003 y la Ley 2381 de 2024.
2. Medición del valor público: El ROI de innovación no solo mide dinero recuperado, sino también legitimidad institucional, confianza ciudadana y transparencia. Cada ahorro representa una pensión legítima protegida y un derecho ciudadano garantizado.
3. Motivación institucional y ciudadana: La tabla transmite un mensaje poderoso: invertir en inteligencia artificial explicable no es un gasto, sino una apuesta estratégica por la sostenibilidad del sistema pensional. Muestra que la tecnología, cuando se aplica con ética, gobernanza y transparencia, genera beneficios concretos, rápidos y sostenibles para la sociedad.

El Retorno de la Inversión (ROI) del 13,900 % y un VPN positivo de 3.2 billones de pesos consolidan a este proyecto como una inversión de alto impacto fiscal y social, perfectamente alineada con los objetivos de la Política de Gobierno Digital (MinTIC, 2022) y el Plan de Transformación Institucional de Colpensiones (2025-2030).

En síntesis, la tabla evidencia que el proyecto no solo se paga solo, sino que multiplica exponencialmente el valor público generado, justificando plenamente su adopción como modelo replicable de innovación preventiva en el Estado colombiano.

La incorporación del modelo de Inteligencia Artificial explicable (XAI) en la detección preventiva del fraude pensional en Colpensiones no solo se traduce en ahorros financieros cuantificables, sino también en beneficios intangibles de alto valor institucional, ético y estratégico; Estos beneficios, aunque no siempre medibles en pesos, representan el núcleo del valor público generado por la innovación y son los que garantizan la sostenibilidad a largo plazo del sistema.

Tabla 7. Beneficios intangibles e impacto institucional del modelo de inteligencia artificial (IA) explicable (XAI) en Colpensiones

Área	Beneficio	Impacto
Transparencia institucional	Mayor legitimidad y trazabilidad del proceso pensional	Alto
Cumplimiento normativo	Conformidad con Ley 2381/2024, Ley 1581/2012 y Ley 797 de 2003	Alto
Eficiencia administrativa	Disminución del 50% en tiempos de validación documental	Medio
Capital humano	Formación de equipos en Inteligencia Artificial ética y antifraude	Alto

Fuente. Elaboración propia con base en Colpensiones (2024), MinTIC (2023) y DNP (2023).

El uso de modelos de Inteligencia Artificial explicable (XAI) proporciona trazabilidad completa de cada decisión o alerta generada por el sistema.

Esto significa que toda acción puede ser auditada y justificada, lo que refuerza la legitimidad institucional de Colpensiones ante los organismos de control (Contraloría, Procuraduría, DNP) y ante la ciudadanía.

La transparencia generada reduce el riesgo de discrecionalidad, aumenta la rendición de cuentas y fortalece la confianza social en la entidad, uno de los principales activos intangibles en la administración pública.

El proyecto está alineado con las disposiciones de:

- Ley 2381 de 2024, que impulsa la digitalización y automatización de los trámites pensionales.
- Ley 1581 de 2012, sobre protección de datos personales, garantizando la privacidad y el tratamiento ético de la información.
- Ley 797 de 2003, que fortalece los mecanismos de control y transparencia en el reconocimiento de pensiones.

El cumplimiento de estas normas no solo evita sanciones o riesgos jurídicos, sino que convierte a Colpensiones en referente de gobernanza digital ética, demostrando que la tecnología puede ser compatible con los principios de legalidad, seguridad y respeto a los derechos ciudadanos.

La interoperabilidad entre Colpensiones, Unidad de Gestión Pensional y Parafiscales (UGPP), Registraduría, Administradora de los Recursos del Sistema General de Seguridad Social en Salud (ADRES) y Ministerio de Trabajo reduce significativamente los tiempos de validación de identidad, convivencia y certificación de semanas cotizadas.

En promedio, se estima una reducción del 50 % en los tiempos de validación documental, lo que impacta positivamente en la experiencia del ciudadano y en la capacidad operativa de la entidad. Esto permite reasignar talento humano a tareas de mayor valor analítico y fortalecer la gestión del riesgo en tiempo real, optimizando recursos y procesos.

El proyecto contempla la formación de equipos técnicos y jurídicos especializados en Inteligencia Artificial ética, explicable y antifraude, fortaleciendo las capacidades internas de Colpensiones. Este componente de capital humano genera aprendizaje organizacional sostenible, reduce la dependencia de terceros y promueve una cultura institucional basada en la innovación, la ética y la responsabilidad algorítmica.

Además, el fortalecimiento del talento interno consolida la sostenibilidad del sistema más allá del horizonte 2030.

La implementación del modelo de inteligencia artificial IA explicable para la detección preventiva del fraude demuestra una viabilidad técnica, financiera e institucional sobresaliente.

Con un ROI superior al 13,900 % y un periodo de recuperación menor a cinco meses, el proyecto no solo garantiza un retorno financiero sólido, sino que crea valor público sostenible, expresado en transparencia, confianza y legitimidad.

En términos de política pública, esta inversión consolida a Colpensiones como una entidad líder en innovación digital ética, alineada con los ejes de la Política de Gobierno Digital (MinTIC, 2022) y el Plan Nacional de Desarrollo 2022–2026: Colombia Potencia Mundial de la Vida.

Aplicación Del Enfoque *Lean Government* Para La Optimización De Procesos En Colpensiones

La sostenibilidad del modelo de Inteligencia Artificial Preventiva (XAI) en Colpensiones requiere no solo innovación tecnológica, sino también una transformación profunda de los procesos administrativos. En este sentido, la metodología *Lean Government*, adaptada del enfoque Lean Management al sector público, permite reducir desperdicios, optimizar tiempos y fortalecer el valor público mediante la eficiencia y la transparencia institucional.

Este enfoque resulta esencial para consolidar los resultados del proyecto, pues garantiza que las mejoras tecnológicas se acompañen de procesos más ágiles, interoperables y centrados en el ciudadano. Su aplicación contribuye directamente a los objetivos de la Ley 2381 de 2024 y la Política de Gobierno Digital (MinTIC, 2023), al promover una gestión pública moderna, digital y sostenible.

Tabla 8. *Aplicación de Lean Government a las falencias identificadas en Colpensiones*

Falencia identificada	Pérdida de valor público	Acción Lean Government	Impacto esperado
Incremento del fraude pensional	Pérdida de recursos por reconocimiento indebido de pensiones, falsificación de semanas y suplantación de identidad (\approx \$1.35 billones anuales).	- Mapear el flujo de valor del reconocimiento de pensiones.	Reducción progresiva del fraude, ahorro fiscal y fortalecimiento de la confianza ciudadana.
		- Implementar tableros visuales de control (dashboards antifraude).	
		- Integrar IA explicable (XAI) con bases de datos interoperables (UGPP, Registraduría, ADRES).	
Interoperabilidad interinstitucional limitada	Retrasos por trámites manuales y validación de datos.	- Crear mapa de flujo de datos interinstitucional.	Procesos más ágiles, menor carga administrativa y mejor experiencia ciudadana.
		- Diseñar APIs seguras.	
		- Realizar <i>kaizen digital</i> para unificar formatos y eliminar documentos físicos.	

Baja calidad y trazabilidad de datos	Reprocesos y errores en historiales laborales.	- Incorporar <i>poka-yokes digitales</i> (validaciones automáticas).	Reducción de reclamaciones y mejora en la confiabilidad de los datos.
		- Asignar <i>Data Stewards</i> .	
		- Implementar panel de indicadores Lean (exactitud, completitud, oportunidad).	
Limitaciones en infraestructura tecnológica	Caídas del sistema, lentitud y altos costos operativos.	- Priorizar inversiones de alto valor público.	Mayor estabilidad, ahorro y escalabilidad tecnológica.
		- Migrar gradualmente a infraestructura en la nube.	
		- Adoptar prácticas DevOps para despliegues continuos.	

Fuente. Elaboración propia con base en Colpensiones (2024), MinTIC (2023) y DNP (2023).

La aplicación de *Lean Government* en Colpensiones complementa la implementación del modelo de IA Preventiva, alineando la innovación tecnológica con la eficiencia administrativa. Este enfoque fortalece la interoperabilidad, reduce reprocesos, y mejora la calidad del servicio público, consolidando una cultura institucional basada en la mejora continua y el valor ciudadano.

Conclusiones

La implementación de un modelo preventivo de detección de fraude basado en inteligencia artificial explicable (XAI) representa una oportunidad estratégica para modernizar la gestión pensional en Colombia y fortalecer la gobernanza digital del Estado. A través de esta propuesta, Colpensiones puede evolucionar de un esquema reactivo de control hacia un enfoque predictivo, ético y colaborativo, en el que la interoperabilidad entre entidades públicas se convierte en el pilar central de la transparencia y la eficiencia institucional.

El modelo planteado combina técnicas de aprendizaje automático con principios de explicabilidad, trazabilidad y participación humana, garantizando el cumplimiento del debido proceso y la protección de datos personales conforme a la Ley 2381 de 2024, la Ley 1581 de 2012 y el CPACA (Ley 1437 de 2011). Esta integración tecnológica y jurídica no solo mejora la detección oportuna de inconsistencias y duplicidades, sino que también refuerza la legitimidad de las decisiones administrativas.

Desde la perspectiva financiera, la propuesta demuestra una viabilidad económica excepcional, con un retorno de inversión estimado superior al 13,900 % y un periodo de recuperación inferior a cinco meses. Dicho resultado refleja no solo el ahorro proyectado por la reducción de fraudes, sino también la eficiencia generada en tiempos de respuesta y reducción de litigios.

Además de los beneficios económicos, este modelo contribuye a restaurar la confianza ciudadana, al ofrecer un sistema de pensiones más transparente, ágil y seguro. El uso responsable de la inteligencia artificial fortalece la sostenibilidad fiscal del régimen de prima media, impulsa la cooperación interinstitucional y posiciona a Colpensiones como un referente en innovación pública digital dentro del sistema de protección social colombiano.

En el marco de los resultados obtenidos, se evidencia que la viabilidad institucional de Colpensiones constituye uno de los principales hallazgos del estudio. La entidad cuenta con un marco jurídico habilitante, una infraestructura tecnológica en expansión y un equipo humano con competencias técnicas y compromiso con el servicio público. Estas condiciones facilitan la adopción progresiva del modelo de inteligencia artificial explicable (XAI) sin afectar la operación diaria ni la

atención a los ciudadanos, lo que demuestra la capacidad institucional para asumir procesos de transformación digital con responsabilidad y continuidad operativa.

De igual manera, la sostenibilidad del modelo se consolida como un elemento clave para garantizar su permanencia en el tiempo. Esta dependerá de una estrategia integral que combine la actualización tecnológica, la formación continua del talento humano y la cooperación interinstitucional con entidades como el Departamento Nacional de Planeación (DNP) y el Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia (MinTIC). Dicho enfoque permitirá que la innovación trascienda el ámbito del proyecto y se convierta en una práctica permanente y replicable en otras entidades del sistema de protección social, fortaleciendo la transparencia, la eficiencia y la confianza ciudadana en la gestión pública.

En síntesis, esta propuesta consolida una visión de Estado más inteligente, ético y confiable, donde la tecnología no reemplaza el juicio humano, sino que lo complementa, asegurando que cada decisión administrativa esté sustentada en evidencia verificable y orientada al valor público.

Referencias

Colpensiones. (2023). *Informe anual de gestión 2023*. Administradora Colombiana de Pensiones.

Colpensiones. (2025, marzo). *Boletín oficial de Colpensiones*. Administradora Colombiana de Pensiones.

Congreso de la República de Colombia. (2024). *Ley 2381 de 2024. Por medio de la cual se adoptan medidas para la eficiencia institucional en el sistema pensional*. Diario Oficial No. 52.345.

Corte Constitucional de Colombia. (2019). *Sentencia SU-182/19*. Corte Constitucional.

Congreso de la República de Colombia. (2003). *Ley 797 de 2003. Por la cual se reforman algunas disposiciones del Sistema General de Pensiones*. Diario Oficial de la República de Colombia.

Congreso de la República de Colombia. (2011). *Ley 1437 de 2011. Código de Procedimiento Administrativo y de lo Contencioso Administrativo (CPACA)*. Diario Oficial de la República de Colombia.

Congreso de la República de Colombia. (2012). *Ley 1581 de 2012. Por la cual se dictan disposiciones generales para la protección de datos personales*. Diario Oficial de la República de Colombia.

Portafolio. (2018, septiembre 19). *Colpensiones detectó 886 presuntos responsables de fraude al régimen público pensional*. <https://www.portafolio.co/mis-finanzas/jubilacion/colpensiones-detecto-886-presuntos-responsables-de-fraude-al-regimen-publico-pensional-515046>

ANEXOS TÉCNICOS

ANEXO A. Capturas prototipo Phyton – windows PowerShell

```

"""
Prototipo lineal (demo) – Colpensiones IA Preventiva (XAI)
-----
Flujo secuencial y simple (sin microservicios) para validación de concepto.
No requiere librerías externas. Pensado para explicar el proceso a áreas
no técnicas y ejecutar una demo rápida en un único archivo.

Pasos del flujo:
1) Consentimiento/base legal (OPA simulado)
2) Verificación de identidad (multi-fuente simulada)
3) Validación documental (tipos + hash)
4) Scoring de fraude + explicación (XAI simulada)
5) Decisión y escalamiento (human-in-the-loop)
6) Auditoría básica (archivo JSON)

Ejecutar (ejemplo):
    python linear_prototipo_colpensiones.py

Salida: imprime el resultado y escribe un audit_log.json en el mismo
directorio.
"""
from __future__ import annotations
import json, hashlib, datetime, sys
from typing import Dict, Any, List, Tuple

# ----- Configuración -----
ALLOWED_DOC_TYPES = {"CC", "RC", "CE", "PAZ_SALVO", "HISTORIA_CLINICA"}
MODEL_VERSION = "demo-linear-0.1"
AUDIT_FILE = "audit_log.json"

# ----- Utilidades -----
def now_iso() -> str:
    return datetime.datetime.utcnow().replace(microsecond=0).isoformat() +
    "Z"

def write_audit(event: str, actor: str, subject_id: str, payload: Dict[str,
Any]) -> None:
    record = {
        "timestamp": now_iso(),
        "event": event,
        "actor": actor,
        "subject_id": subject_id,
        "payload": payload,
    }
    try:
        try:
            with open(AUDIT_FILE, "r", encoding="utf-8") as f:

```

```

        data = json.load(f)
    except FileNotFoundError:
        data = []
    data.append(record)
    with open(AUDIT_FILE, "w", encoding="utf-8") as f:
        json.dump(data, f, ensure_ascii=False, indent=2)
except Exception as e:
    print(f"[WARN] No se pudo escribir auditoría: {e}", file=sys.stderr)

# ----- 1) Consentimiento / Base legal -----
def check_consent(purpose: str, consent_granted: bool, base_legal: str =
"interes_publico") -> bool:
    """
    Política simplificada (sustituye OPA en este demo):
    - Permitir si hay consentimiento explícito y el propósito es
'prevencion_fraude'
    - O permitir por base legal de interés público para 'prevencion_fraude'
(denegar por defecto si no)
    """
    if purpose == "prevencion_fraude" and (consent_granted or base_legal ==
"interes_publico"):
        return True
    return False

# ----- 2) Verificación de identidad -----
OFFICIAL_SOURCES = ["Registraduria", "RUAUF", "ADRES", "SISBEN", "MinTrabajo"]

def verify_identity(doc_type: str, doc_number: str, full_name: str,
birth_date: str) -> Dict[str, Any]:
    """
    Simula una verificación: puntaje alto si el número aparenta ser válido y
si el nombre y fecha no están vacíos.
    """
    score = 0.0
    reasons: List[str] = []
    if doc_type in {"CC", "CE", "RC"}:
        score += 0.4
        reasons.append("Tipo documental válido")
    if doc_number.isdigit() and 6 <= len(doc_number) <= 10:
        score += 0.35
        reasons.append("Formato de número plausible en Registraduría")
    if full_name.strip() and birth_date.strip():
        score += 0.25
        reasons.append("Nombre y fecha de nacimiento presentes")
    matched = score >= 0.8
    return {
        "matched": matched,
        "score": round(score, 2),
        "sources_checked": OFFICIAL_SOURCES,
        "reasons": reasons
    }

# ----- 3) Validación documental -----
def validate_document(doc_type: str, file_bytes: bytes | None, file_sha256:
str | None) -> Dict[str, Any]:
    issues: List[str] = []

```

```

if doc_type not in ALLOWED_DOC_TYPES:
    issues.append("Tipo de documento no permitido")
if file_sha256:
    valid_hash = len(file_sha256) == 64 and all(c in "0123456789abcdef"
for c in file_sha256.lower())
    if not valid_hash:
        issues.append("Hash SHA256 inválido")
elif file_bytes is not None:
    # Si se pasan bytes, calculamos el hash
    file_sha256 = hashlib.sha256(file_bytes).hexdigest()
else:
    issues.append("No se suministró hash ni bytes del archivo")

return {
    "valid": len(issues) == 0,
    "issues": issues,
    "sha256": file_sha256 or None
}

# ----- 4) Scoring de fraude + "XAI" -----
def score_claim(features: Dict[str, float]) -> Tuple[float, str, Dict[str,
float]]:
    """
    Regla simple con explicación estilo importancias:
    - valor_prestacion muy alto aumenta riesgo
    - pocas semanas_cotizadas aumenta riesgo
    - canal_riesgo alto aumenta riesgo
    """
    s = 0.1
    expl = {"valor_prestacion": 0.1, "semanas_cotizadas": 0.1,
"canal_riesgo": 0.1}

    if features.get("valor_prestacion", 0) > 5e7:
        s += 0.6; expl["valor_prestacion"] = 0.55
    if features.get("semanas_cotizadas", 0) < 500:
        s += 0.2; expl["semanas_cotizadas"] = 0.25
    if features.get("canal_riesgo", 0) > 0.7:
        s += 0.1; expl["canal_riesgo"] = 0.15

    s = max(0.0, min(1.0, s))
    if s >= 0.75: label = "HIGH"
    elif s >= 0.5: label = "MEDIUM"
    else: label = "LOW"
    return s, label, expl

# ----- 5) Decisión final -----
def decide(identity_ok: bool, doc_ok: bool, risk_label: str) -> Dict[str,
Any]:
    """
    Política simple:
    - Si identidad falla -> Rechazar y escalar a revisión
    - Si documentos fallan -> Solicitar subsanación
    - Si riesgo HIGH/MEDIUM -> Aprobación condicionada a revisión humana
    - Si todo bien y riesgo LOW -> Aprobación preliminar
    """
    if not identity ok:

```

```

        return {"status": "RECHAZADO", "reason": "Identidad no verificada",
"human_review": True}
        if not doc_ok:
            return {"status": "SUBSANAR", "reason": "Documento inválido o
incompleto", "human_review": False}
            if risk_label in {"HIGH", "MEDIUM"}:
                return {"status": "EN_REVISION", "reason": f"Riesgo {risk_label}",
"human_review": True}
                return {"status": "APROBACION_PRELIMINAR", "reason": "Riesgo bajo",
"human_review": False}

# ----- 6) Pipeline lineal -----
def run_pipeline(payload: Dict[str, Any], actor: str = "api-demo") ->
Dict[str, Any]:
    # 1) Consentimiento/Base legal
    consent_ok = check_consent("prevencion_fraude",
payload.get("consent_granted", True), payload.get("base_legal",
"interes_publico"))
    if not consent_ok:
        result = {"ok": False, "error": "Consentimiento/Base legal no
válido"}
        write_audit("consent.denied", actor, payload.get("citizen_id",
"N/A"), result)
        return result

    # 2) Identidad
    idr = verify_identity(payload["doc_type"], payload["doc_number"],
payload["full_name"], payload["birth_date"])
    write_audit("identity.verify", actor, payload.get("doc_number", "N/A"),
idr)

    # 3) Documento
    docr = validate_document(payload["doc_to_check"]["type"], None,
payload["doc_to_check"]["sha256"])
    write_audit("document.validate", actor, payload.get("claim_id", "N/A"),
docr)

    # 4) Scoring
    score, label, expl = score_claim(payload["features"])
    scoring = {"risk_score": score, "label": label, "explanation": expl,
"model_version": MODEL_VERSION}
    write_audit("claims.score", actor, payload.get("claim_id", "N/A"),
scoring)

    # 5) Decisión
    decision = decide(idr["matched"], docr["valid"], label)
    write_audit("decision", actor, payload.get("claim_id", "N/A"), decision)

    return {
        "ok": True,
        "identity": idr,
        "document": docr,
        "scoring": scoring,
        "decision": decision
    }

```

```
# ----- Demo local -----  
if __name__ == "__main__":  
    sample = {  
        "claim_id": "SOL-2025-0001",  
        "citizen_id": "CC-123456789",  
        "doc_type": "CC",  
        "doc_number": "123456789",  
        "full_name": "MARIA PEREZ",  
        "birth_date": "1970-05-04",  
        "consent_granted": True,  
        "base_legal": "interes_publico",  
        "doc_to_check": {"type": "PAZ_SALVO", "sha256": "0"*64},  
        "features": {"valor_prestacion": 80000000, "semanas_cotizadas": 420,  
"canal_riesgo": 0.8}  
    }  
  
    result = run_pipeline(sample, actor="demo-user")  
    print(json.dumps(result, ensure_ascii=False, indent=2))  
    print(f"\n[OK] Auditoría escrita en {AUDIT_FILE}")
```