

REIA(REGULATOR INTELLIGENT IA): Evaluación reproducible y multilingüe del riesgo de cumplimiento legal de IA en Latinoamérica

En colaboración con Apart Research

Hackathon de Seguridad en IA — YAIS Lab LatAm (Bolivia)

Integrantes:

Iver Ceron Condori

Gonzalo Veizaga Veizaga Justiniano

Richard Junior Montaña Flores

Yeisa Isabel Torrico Romero

Resumen (Abstract)

Las organizaciones de Latinoamérica adoptan inteligencia artificial a gran velocidad y, al hacerlo, exponen datos personales y toman decisiones automatizadas que pueden incumplir leyes de IA emergentes en la región. La revisión legal tradicional es lenta y cuesta miles de dólares. Presentamos LegisAI LatAm, un evaluador de riesgo de cumplimiento que responde en segundos. El usuario elige un país (la jurisdicción) y describe su proyecto de IA; el sistema devuelve un porcentaje de riesgo, un semáforo, un índice global de cumplimiento, las normas del país que podría estar infringiendo (con su enlace de fuente), recomendaciones y un veredicto. El diseño separa la búsqueda del análisis: usamos Tavily para buscar en la web las normas vigentes del país y devolver la URL de cada fuente, y DeepSeek V4 Flash para analizar el proyecto contra esas normas, puntuar, redactar y traducir. La clave es de seguridad: el analista solo puede citar normas que la búsqueda devolvió (con URL) o que figuran en un catálogo curado de respaldo, de modo que no puede inventar leyes; el número global de cumplimiento lo calcula código determinista a partir de sub-puntajes; y una caché por huella de contenido garantiza que el mismo proyecto produzca siempre el mismo veredicto (reproducibilidad). La interfaz es responsive y multilingüe: además de castellano, ofrece quechua, aymara, guaraní y plautdietsch, ampliando el acceso a comunidades lingüísticas de Bolivia. Sobre tres casos representativos el sistema produce veredictos consistentes y citas con fuente. La conclusión es que separar una búsqueda con citas

de un análisis restringido a esas citas, sumado a agregación determinista y a una caché de consistencia, convierte la evaluación automatizada de cumplimiento en algo auditable y reproducible, atacando el riesgo de dictámenes legales opacos y propensos a la alucinación.

1. Introducción

La adopción de IA en Latinoamérica avanza más rápido que la capacidad de las organizaciones para verificar si su uso cumple la ley. Alcaldías, fintechs, hospitales y plataformas despliegan sistemas que procesan datos sensibles y automatizan decisiones de alto impacto sin una evaluación previa, porque la vía existente (contratar estudios jurídicos) es cara y lenta.

El modelo de amenaza que abordamos es doble. Primero, el de las decisiones automatizadas que incumplen normas de IA emergentes (registro biométrico, scoring social, manipulación, ausencia de supervisión humana, tratamiento de neurodatos). Segundo, un riesgo propio de este tipo de herramienta: si el dictamen legal lo produjera un modelo de caja negra sin restricciones, tendríamos un asesor opaco, no reproducible y capaz de alucinar artículos de ley. Nuestro diseño está hecho explícitamente para no caer en eso.

Nuestras principales contribuciones son:

- Un pipeline que separa la búsqueda de la decisión: una búsqueda web (Tavily) recupera normas reales con su URL, y un analista que solo puede citar esas normas, eliminando por construcción la alucinación de leyes.
- Una capa de reproducibilidad: agregación determinista del índice global por código, más una caché por huella de contenido, de modo que el mismo proyecto da el mismo veredicto.
- Una interfaz de cumplimiento inclusiva, responsive y multilingüe (castellano más quechua, aymara, guaraní y plautdietsch), con detección del idioma del documento de entrada.

2. Trabajo relacionado

El Reglamento de IA de la Unión Europea (AI Act, 2024) popularizó la clasificación por niveles de riesgo; de allí tomamos la idea de prohibiciones absolutas y categorías de alto riesgo, pero las aterrizamos a las normas latinoamericanas: el proyecto de ley boliviano PL 178/2024-2025 (crea la ARIA y fija prohibiciones), el PL 2338/2023 de Brasil, la Ley 21.383 de Chile (neuroderechos) y el CONPES 4144 de Colombia, entre otras normas nacionales.

Los asistentes legales basados en RAG y los modelos de lenguaje sin anclaje son los antecedentes más cercanos. Nos diferenciamos de ambos: en vez de un RAG con base vectorial, usamos una búsqueda web orientada a agentes que devuelve resultados con URL, y restringimos al modelo a citar únicamente esas fuentes (o un catálogo curado de respaldo). Así ganamos trazabilidad (cada

cita tiene su enlace) y reproducibilidad (caché por contenido), evitando tanto la alucinación de artículos como la opacidad de un dictamen no auditable.

3. Métodos

El sistema no es un chatbot. Separa la búsqueda (Tavily) del análisis (DeepSeek V4 Flash), y el código aporta el catálogo de respaldo, la agregación determinista y la consistencia. El flujo es: país + idioma + PDF/texto -> detección de idioma -> [caché?] -> búsqueda de las leyes del país -> análisis contra esas leyes -> índice global por código -> tablero/informe traducido.

3.1 Investigador (Tavily)

Dado un resumen del proyecto y el país elegido, el investigador consulta la API de Tavily (un buscador orientado a agentes de IA) para recuperar las normas del país que podrían aplicar, y devuelve cada una con su URL de fuente. Las consultas se siembran con un catálogo curado de normas conocidas por país y con la señal o el sector detectados en el proyecto. El plan gratuito de Tavily (1.000 búsquedas por mes) cubre el volumen de un hackathon y de la demo.

3.2 Analista/Redactor (DeepSeek V4 Flash)

El analista recibe el proyecto y la lista de normas encontradas (con URL). Detecta el idioma y el sector, calcula todas las dimensiones fijas del informe (núcleo de seguridad y ética, y calidad del sistema) y los sub-puntajes, lista los riesgos con su nivel, propone recomendaciones y acciones, y emite un veredicto. Las etiquetas de las dimensiones son fijas en la interfaz; el análisis solo produce los porcentajes. Reglas duras: citar solo las normas provistas (nunca inventar), y bloquear (estado BLOQUEADO) ante neurodatos, scoring social o manipulación. Se usa temperatura cero para máxima estabilidad.

3.3 Índice global determinista

El índice global de cumplimiento lo calcula código a partir de los sub-puntajes del analista, con el cumplimiento legal ponderado más alto por tratarse de una herramienta de cumplimiento. Así, el número final lo decide el código, no el modelo.

3.4 Consistencia por caché

Una caché por huella (hash del texto normalizado más el país) congela el primer análisis de cada proyecto. Reanalizar el mismo proyecto en el mismo país devuelve el resultado guardado, idéntico. Esto garantiza 'mismo proyecto, mismo porcentaje' sin depender de que el modelo sea determinista, y hace la demo instantánea y sin gastar búsquedas ni tokens.

3.5 Multi-país, multilingüe (i18n) y responsive

Hay dos selectores independientes: país (jurisdicción, que define qué leyes se buscan) e idioma de visualización. La interfaz se traduce con catálogos JSON congelados por idioma (castellano, quechua, aymara, guaraní, plautdietsch) y un ayudante t(clave, idioma); no se usa gettext. El contenido generado (explicación, título) lo redacta el modelo en el idioma elegido, aunque el PDF venga en otro. La interfaz usa un layout flexible para ser responsive en web.

4. Resultados

La Tabla 1 resume el comportamiento esperado sobre tres casos representativos. Los porcentajes y estados son salidas del sistema; las citas se restringen a lo que devolvió la búsqueda. Nota: las cifras finales deben confirmarse ejecutando el sistema con las API; lo que sigue es el diseño de la demostración.

Caso	País	Estado	Semáforo	Citas (con fuente)
1. Reconocimiento facial para subsidios, sin humano	Bolivia	ALTO / BLOQUEADO	Rojo	PL 178 (biometría) + privacidad (CPE / Ley 164)
2. Detección de emociones del cliente por cámara	Chile	BLOQUEADO	Rojo	Ley 21.383 (neuroderechos, Art. 19 N°1 CPE)
3. Chatbot de atención al cliente de una tienda	Bolivia	BAJO	Verde	Recordatorio de transparencia (sin bloqueo)

Tabla 1. Veredictos de diseño del sistema sobre los tres casos de prueba.

El resultado clave es cualitativo pero robusto: los casos que disparan prohibiciones absolutas (neurodatos, scoring social, manipulación) quedan bloqueados, mientras que los de bajo impacto pasan en verde. Gracias a la caché, reanalizar el mismo proyecto devuelve exactamente el mismo veredicto, propiedad imprescindible en una herramienta de cumplimiento.

5. Discusión y limitaciones

La implicancia para la seguridad en IA es directa. A medida que los dictámenes de cumplimiento se automatizan, el peligro no es solo que un sistema falle, sino que falle de forma opaca e irreproducible. Al recuperar normas reales con su URL, al restringir el modelo a citar solo esas fuentes, al calcular el número global por código y al congelar resultados en caché, cada veredicto es auditable y estable. Además, el soporte de lenguas originarias amplía el acceso a comunidades habitualmente excluidas de estas herramientas, un aporte de inclusión pertinente para Bolivia.

Limitaciones

- La calidad de las citas depende de la búsqueda web y de las fuentes disponibles en línea; para países sin catálogo curado, conviene marcar las citas como 'a verificar'.
- Varias normas citadas son proyectos de ley aún no vigentes (Bolivia, Brasil); debe comunicarse con claridad y no presentarse como derecho vigente.

- La traducción automática a quechua, aymara, guaraní y especialmente plautdietsch (idiomas de bajos recursos) puede ser imperfecta; requiere revisión de hablantes nativos antes de un uso real.
- Los porcentajes por dimensión los produce el modelo; aunque se cachean para consistencia, su calibración debería validarse con expertos. La herramienta no sustituye asesoría legal.

Trabajo futuro

Pasos naturales: ampliar el catálogo y la cobertura a más países de Latinoamérica, acotar la búsqueda a fuentes oficiales (gacetas y congresos), validar el catálogo y la calibración con profesionales del derecho, y revisar las traducciones con hablantes nativos.

6. Conclusión

LegisAI LatAm muestra que se puede construir un evaluador de riesgo de cumplimiento legal útil y rápido sin entrenar modelos, sin base vectorial y sin delegar ciegamente el dictamen en un modelo de lenguaje. La búsqueda recupera normas reales con su fuente, el modelo redacta y puntúa pero restringido a esas normas, y el código agrega el número global y garantiza la consistencia. Esa combinación, más una interfaz multilingüe e inclusiva, hace la herramienta confiable para un dominio donde la trazabilidad y la reproducibilidad no son negociables.

Código y datos

- Repositorio de código: [enlace a GitHub/GitLab]
- Datos/datasets: no se utilizan conjuntos de datos; las citas provienen de la búsqueda web y de un catálogo curado a mano.
- Otros artefactos: [enlace a demo / video del flujo]

Contribuciones de los autores

[Completar: quién diseñó la arquitectura, quién implementó el código, quién trabajó la cobertura legal y el formato de informe. Todos los autores revisaron el manuscrito final.]

Referencias

- Estado Plurinacional de Bolivia. Proyecto de Ley 178/2024-2025 (crea la ARIA). [verificar]
- República Federativa de Brasil. PL 2338/2023, marco regulatorio de IA. [verificar]
- República de Chile. Ley 21.383 (neuroderechos; Art. 19 N°1 de la Constitución). [verificar]
- República de Colombia. CONPES 4144 (2025); Ley 1581 de 2012 de protección de datos. [verificar]
- Bolivia. Ley 164 (Telecomunicaciones y TIC); Ley 393 de Servicios Financieros y normativa ASFI. [verificar]
- Unión Europea. Reglamento (UE) 2024/1689 (AI Act). [clasificación por niveles de riesgo]

- Tavily. API de búsqueda web para agentes de IA. DeepSeek. API del modelo deepseek-v4-flash.

Apéndice

A. Dimensiones fijas del informe (solo cambian los porcentajes)

Núcleo de Seguridad y Ética: Privacidad de Datos, Seguridad Técnica, Transparencia, Explicabilidad, Sesgo y Equidad, Supervisión Humana, Cumplimiento Normativo, Impacto Social y Derechos Humanos, Gobernanza de Datos, Confiabilidad del Sistema. Calidad del Sistema: Calidad del Sistema, Gestión de Riesgos, Uso Ético de IA, Accesibilidad e Inclusión. Índice Global: Cumplimiento Técnico, Cumplimiento Legal, Cumplimiento Ético, Madurez de Gobernanza.

B. Reglas de bloqueo duro (capa de seguridad determinista opcional)

Condición	Estado	Motivo
neurodatos	BLOQUEADO	Chile, Art. 19 N°1 CPE (Ley 21.383)
social_scoring o manipulación	BLOQUEADO	Brasil riesgo excesivo / Bolivia PL 178
biometría y decisión_automática y sin_humano	ALTO (mitigable)	Exige humano + evaluación de impacto

C. Declaración de uso de modelos de lenguaje

Usamos Cursor para el desarrollo del proyecto (Opus para la planificación y la arquitectura, Sonnet para la implementación del código), Tavily para la búsqueda web de leyes y DeepSeek para el análisis, la redacción y la traducción dentro del producto. Todas las afirmaciones, cifras y citas legales fueron verificadas por el equipo. La versión final de este informe fue redactada principalmente por los autores.