

Enpath

INFORME · INVESTIGACIÓN

Generar confianza institucional en la IA: un marco verificable

Octubre de 2024

Organización independiente dedicada a promover un uso responsable de la inteligencia artificial en las decisiones públicas y organizacionales. Cada publicación incluye su metodología, sus fuentes y sus limitaciones.

Solo tres de cada diez colombianos confían en su gobierno nacional. Sobre ese suelo, cada sistema de inteligencia artificial que el Estado despliega es un préstamo tomado contra un capital escaso. Este documento propone un marco de cinco pilares para que la adopción de IA construya confianza en lugar de gastarla — con una regla única: nada cuenta si un tercero no puede verificarlo.

Actualización.

Actualizado el: 18/07/2026. Esta sección resume acontecimientos posteriores a la fecha original de publicación y no forma parte del análisis histórico original.

- La OCDE aplicó en 2025 su encuesta de confianza específicamente en América Latina y el Caribe, ampliando la base comparada que este informe usa; los resultados regionales confirmaron el patrón descrito.
- El CONPES 4144 (febrero de 2025) incorporó componentes de gobernanza y confianza; los informes de 2025 y 2026 de esta serie evalúan su implementación.

Resumen ejecutivo.

La adopción pública de inteligencia artificial suele tratarse como un problema técnico con una dimensión comunicacional: primero se despliega, después se explica. Este documento invierte el orden: **en democracias con confianza institucional baja —el caso de América Latina—, la confianza no es el resultado esperable de la IA pública; es su condición de posibilidad.** Un sistema técnicamente correcto desplegado por una institución en la que nadie confía será leído como amenaza; y tendrá razón quien lo lea así si la institución no puede demostrar lo contrario.

Los datos de partida son inequívocos. Según la encuesta de confianza de la OCDE publicada en julio de 2024, el 32 % de los colombianos declara confianza alta o moderada en el gobierno nacional — siete puntos por debajo del promedio OCDE (39 %), aunque con mejora respecto de 2021; Chile registra 30 % y México 54 % ([OCDE, 2024](#)). Las mediciones regionales agregan el ángulo tecnológico: la conversación latinoamericana en redes sociales sobre IA es mayoritariamente pesimista (31,3 % negativa frente a 23,2 % positiva, según el ILIA), y menos de tres de cada diez latinoamericanos se declaran satisfechos con el funcionamiento de su democracia ([Latinobarómetro, 2023](#)).

El **Marco Enpath de confianza institucional en la IA** organiza la respuesta en cinco pilares — transparencia verificable, voz y debido proceso, competencia demostrada, salvaguardas independientes y coherencia en el tiempo — contruidos sobre el marco de determinantes de la confianza de la OCDE y

sobre la evidencia de esta serie. Su regla constitutiva lo distingue de las cartas de principios: **cada pilar se define por prácticas que un tercero puede comprobar**. La confianza no se declara ni se pide; se deja verificar.

Hallazgos principales.

1. **Colombia adoptará IA pública sobre un capital de confianza del 32 %.** La cifra no es un dato de contexto: es la restricción de diseño. Con ese punto de partida, el margen de error reputacional de cualquier sistema visible es mínimo, y el costo de un escándalo, máximo. La buena noticia del mismo estudio: Colombia es uno de los países donde la percepción de los determinantes de gobernanza mejoró entre 2021 y 2023 — la confianza responde a la práctica (OCDE, 2024).
2. **La evidencia de esta serie converge en una sola causa raíz de la desconfianza justificada: la opacidad.** En los catorce casos del catálogo de junio, el daño se multiplicó donde el sistema era invisible (sin registro, sin métricas, sin vía de reclamo) y se contuvo donde había escrutinio temprano. La transparencia no es un valor decorativo: es el mecanismo causal de la corrección a tiempo.
3. **La confianza tiene determinantes medidos, no misterios.** El marco de la OCDE los ordena en competencia (capacidad de respuesta, confiabilidad) y valores (integridad, apertura, equidad); los cinco pilares de Enpath los traducen a prácticas específicas de IA pública, del registro de sistemas a la autoridad de suspensión.
4. **El desempeño sin visibilidad no genera confianza; la visibilidad sin desempeño la destruye.** Los dos fracasos simétricos: sistemas que funcionan y nadie puede comprobarlo, y transparencia teatral sobre sistemas que fallan. El marco exige las dos cosas a la vez — por eso los pilares 1 y 3 son inseparables.
5. **La región tiene los verificadores; le faltan las ventanas.** Academia, sociedad civil digital, prensa investigativa y jueces han demostrado capacidad de escrutinio (los casos de Salta, Alerta Niñez y Buenos Aires fueron corregidos por ellos). Lo que no existe son las ventanas institucionales —registros, métricas públicas, informes de incidentes— que harían ese escrutinio sistemático en lugar de heroico.

Por qué importa.

Este marco apoya una decisión que toda entidad enfrenta al desplegar IA: **qué prácticas de confianza instalar desde el día uno — y cuáles no pueden aplazarse "para cuando el sistema madure"**. La respuesta importa doblemente en 2024–2025: Colombia formula su política nacional de IA en un clima de confianza frágil, y la primera cosecha de reglas exigibles (T-323, circular de la SIC sobre datos personales e IA) llegó por la vía del control, no de la promoción. La lección de esa secuencia es el corazón de este documento: las instituciones que se dejan verificar definen los términos de su escrutinio; las que no, los reciben.

Antecedentes.

Dos tradiciones convergen aquí. La primera: la medición sistemática de la confianza institucional — el marco de la OCDE sobre sus determinantes (competencia y valores), aplicado en encuestas comparadas desde 2021, con la ola de 2024 cubriendo 30 países incluida Colombia (OCDE, 2024). La segunda: la evidencia acumulada de esta serie sobre qué destruye la confianza en sistemas públicos automatizados — el catálogo de fallas (junio), las condiciones de supervisión (febrero, septiembre) y las cláusulas de compra (agosto). Este documento las integra: los determinantes generales de la confianza, aplicados al caso específico de la IA pública, con las prácticas concretas que esta serie ya especificó.

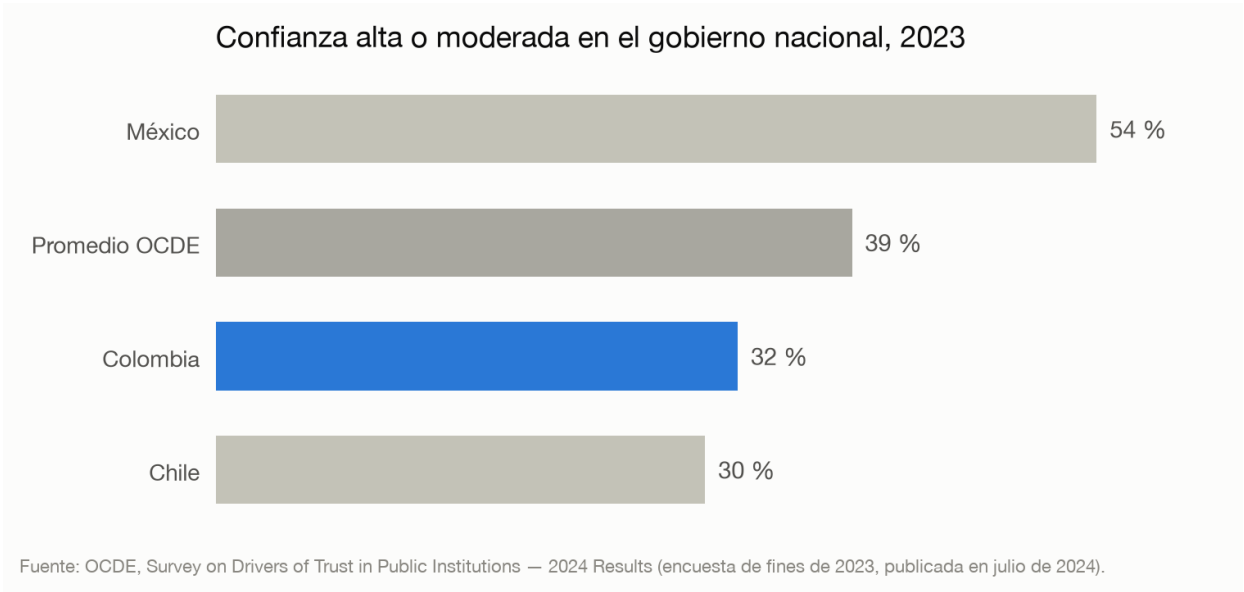


Figura 1. Confianza alta o moderada en el gobierno nacional (2023). Fuente: OCDE (2024).

El marco.



Figura 2. Marco Enpath de confianza institucional en la IA. Fuente: Enpath.

1 · Transparencia verificable. La entidad publica qué sistemas usa (registro), cómo funcionan (documentación), y cómo se desempeñan (métricas trimestrales, incluidas las de supervisión). *Verificación:* un periodista puede escribir un artículo preciso sobre el sistema sin pedir una entrevista.

2 · Voz y debido proceso. Las personas afectadas saben que un sistema participó en su caso, reciben explicación comprensible, pueden impugnar ante una persona con autoridad y obtienen respuesta en plazo definido. *Verificación:* el canal de impugnación existe, está publicado, y sus tiempos de respuesta se reportan.

3 · Competencia demostrada. El sistema mejora algo medible —tiempos, cobertura, errores— y la entidad publica esa medición, incluidos los resultados incómodos y los incidentes gestionados. *Verificación:* existe una línea de base previa al despliegue y reportes posteriores comparables.

4 · Salvaguardas independientes. Alguien externo puede examinar el sistema: auditoría técnica pactada por contrato (cláusula 3 de agosto), órganos de control con capacidad, espacio protegido para el escrutinio académico y de sociedad civil. *Verificación:* la última revisión externa tiene fecha, autor y hallazgos publicados.

5 · Coherencia en el tiempo. Las reglas del sistema no cambian sin aviso; los compromisos sobreviven a los cambios de administración; lo prometido en el despliegue se cumple o se explica. *Verificación:* los documentos de gobernanza del sistema tienen historial público de versiones.

La correspondencia con los determinantes OCDE es directa: los pilares 1 y 2 operacionalizan la *apertura* y la *equidad*; el 3, la *capacidad de respuesta* y la *confiabilidad*; el 4 y el 5, la *integridad*. La diferencia está en el grano: los determinantes describen percepciones; los pilares prescriben prácticas comprobables.

Dos espirales: lo que la transparencia compra y lo que la opacidad cuesta

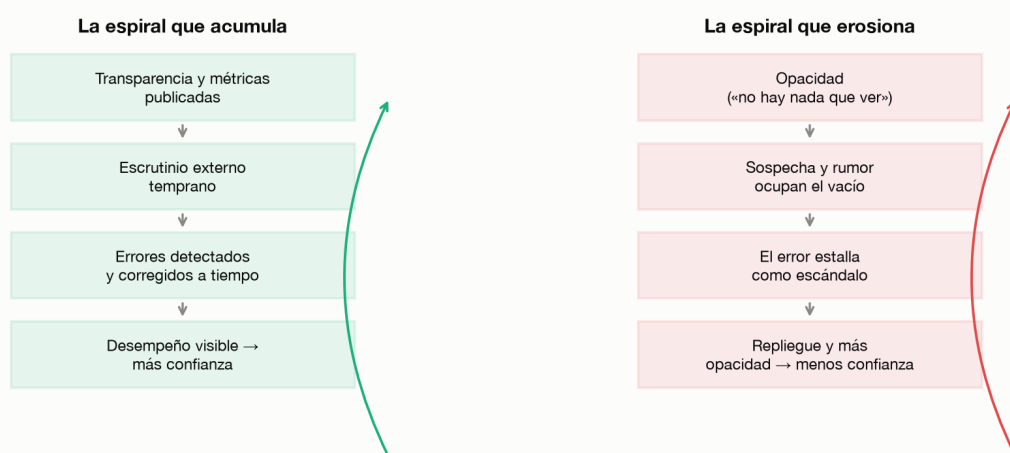


Figura 3. El mecanismo: transparencia y escrutinio corrigen temprano; la opacidad convierte errores en escándalos. Fuente: Enpath, con base en el catálogo de casos de esta serie.

Situación actual.

Colombia cerró 2024 con una paradoja instructiva: **sus primeras reglas exigibles de IA nacieron de la desconfianza, no de la promoción.** La sentencia T-323 de 2024 (transparencia, responsabilidad y no sustitución en la función judicial) y la Circular Externa 002 de 2024 de la SIC (tratamiento de datos personales en sistemas de IA) son actos de control que fijan estándar; la política de promoción (el CONPES en elaboración) llega después. Para las entidades, el mensaje institucional es claro: el escrutinio ya opera; las ventanas de transparencia que este marco propone son la diferencia entre someterse a él o colaborar con él. En la región, el ILIA 2024 (septiembre) reconfirmó a Chile, Brasil y Uruguay como líderes y documentó el crecimiento desigual de capacidades — mientras el clima social de fondo (satisfacción democrática por debajo del 30 %) mantiene el margen de tolerancia a fallas tecnológicas visibles en mínimos ([CENIA y CEPAL, 2024](#); [Latinobarómetro, 2023](#)).

Análisis de datos.

Tres lecturas de los datos disponibles sostienen el marco:

Primera: la confianza latinoamericana es baja pero sensible a la práctica. El caso colombiano en la encuesta OCDE —mejora de percepciones de gobernanza 2021–2023 acompañada de mejora en confianza— replica el patrón general del estudio: los determinantes se mueven y la confianza los sigue, con rezago. Para la agenda de IA esto significa que las prácticas verificables no son un gesto: son la palanca disponible.

Segunda: la desconfianza tecnológica regional es específica, no genérica. El contraste del ILIA entre medios (optimistas) y conversación social (pesimista) sugiere que la ciudadanía no rechaza la tecnología en abstracto: desconfía de quién la usa y para qué. Esa es exactamente la pregunta que los pilares 1 y 2 responden — y que ninguna campaña de comunicación puede responder por ellas.

Tercera: los casos confirman el mecanismo causal. En el catálogo de junio, el tiempo de detección —la variable que separa el error corregido del desastre— es función directa de las ventanas de escrutinio. Ofqual (transparente, corregido en una semana) contra Horizon (opaco, dos décadas): la espiral de la figura 3 no es una metáfora; es el resumen de los expedientes.

Pilar	Práctica mínima	Costo aproximado	Informe de esta serie que la especifica
Transparencia	Registro público de sistemas	Bajo (inventario ya recomendado)	Julio
Voz	Canal de impugnación con plazo	Bajo	Febrero
Competencia	Línea de base y métricas publicadas	Medio	Septiembre
Salvaguardas	Cláusula de auditoría + revisión externa anual	Medio	Agosto
Coherencia	Gobernanza documentada con versiones	Bajo	Mayo

Colombia.

La aplicación colombiana del marco tiene un punto de partida concreto: el inventario de julio mostró que ninguno de los sistemas documentados publica métricas ni tiene registro oficial. Con 32 % de confianza de fondo, esa opacidad es una hipoteca creciente. Tres movimientos la revierten a bajo costo: **el registro público de sistemas** (la infraestructura de datos abiertos del país lo hace viable en meses — la fortaleza documentada en mayo); **la publicación de métricas de supervisión** de los sistemas de mayor impacto (justicia y fiscalización primero, con la bitácora de septiembre como fuente); y **la respuesta institucional ejemplar al próximo incidente** — porque lo habrá, y la diferencia entre un escándalo y una demostración de competencia es la preparación (protocolo de incidentes de junio). El sector con la oportunidad más clara es la Rama Judicial: los lineamientos exhortados por T-323 pueden ser el primer documento público regional que convierta los cinco pilares en reglas operativas de una institución completa.

América Latina.

El marco viaja bien regionalmente porque sus insumos ya existen en toda la región: verificadores activos (la academia y sociedad civil que corrigieron Salta, Alerta Niñez y el reconocimiento facial porteño), tradición de leyes de transparencia y acceso, y organismos regionales capaces de sostener estándares comunes. La brecha compartida son las ventanas: ningún país de la región mantiene aún un registro nacional de sistemas públicos de IA con métricas — el instrumento que el informe de abril de 2025 de esta serie desarrollará en detalle. La cooperación de menor costo y mayor señal: que dos o tres países acuerden el formato común del registro, convirtiendo la transparencia algorítmica en estándar regional antes de que sea obligación importada.

Comparación internacional.

- **Países Bajos** ofrece la lección completa del ciclo: la opacidad de SyRI y Toeslagen destruyó un capital enorme de confianza; la respuesta — registro nacional de algoritmos, autoridad de supervisión algorítmica — es el catálogo de ventanas que este marco recomienda instalar *antes* del escándalo, no después.
- **Reino Unido** estandarizó el reporte de transparencia algorítmica para el gobierno (Algorithmic Transparency Recording Standard) — la plantilla de registro más madura disponible, adaptable al español con esfuerzo mínimo ([Gobierno del Reino Unido](#)).
- **Estados Unidos** convirtió el inventario público de casos de uso en obligación federal anual (M-24-10) y añadió en octubre de 2024 la guía de adquisición (M-24-18): la infraestructura de confianza como requisito de gestión, no como aspiración ([OMB, 2024](#)).

- **OCDE** aporta el marco de determinantes y su medición comparada — el instrumento que permitirá evaluar, en las próximas olas, si las prácticas de IA pública mueven la aguja de la confianza. La hipótesis de este documento quedará así sometida a prueba externa.

Oportunidades.

- **El dividendo del primero.** La primera entidad colombiana que publique registro y métricas fijará el estándar nacional de facto y cosechará la distinción — en un entorno de 32 % de confianza, la señal de "dejarse verificar" es escasa y por eso valiosa.
- **La confianza como ventaja operativa, no solo reputacional.** Sistemas confiados reciben mejores datos (la gente coopera), menos litigio y más margen político para iterar. Los cinco pilares son también una inversión en eficiencia.
- **El control como aliado.** Órganos de control preparados (recomendación de julio) y ventanas de transparencia convierten la relación entidad-controlador de adversarial en correctiva — el modelo Ofqual, no el modelo Horizon.
- **Medición externa disponible.** Con Colombia dentro de la encuesta OCDE, el país tiene gratis el tablero para evaluar si su política de IA construye o gasta confianza. Pocas políticas públicas tienen un indicador de resultado tan limpio esperándolas.

Riesgos.

- **Transparencia teatral.** Registros desactualizados, métricas seleccionadas, informes sin hallazgos: la simulación del pilar 1 es peor que su ausencia, porque gasta la credibilidad del instrumento. El antídoto es la verificabilidad: fechas, versiones, terceros.
- **El primer escándalo mal gestionado.** Un incidente visible manejado con negación definiría la percepción pública de la IA estatal por años. El protocolo de incidentes no es opcional; es el seguro del programa completo.
- **Confianza pedida, no ganada.** Campañas de pedagogía sobre "los beneficios de la IA" sin las prácticas de este marco — el equivalente institucional de pedir un préstamo sin garantías. La comunicación sigue a la verificabilidad; nunca la sustituye.
- **Erosión por rotación.** Cada cambio de administración que reinicia reglas, nombres y compromisos cobra su precio en el pilar 5. La gobernanza documentada con historial es la protección mínima contra la amnesia institucional.

Perspectivas.

Con corte a octubre de 2024: la política nacional colombiana de IA definirá si la confianza es un capítulo o un criterio transversal; los lineamientos judiciales post-T-323 serán la primera prueba institucional del marco; y las próximas olas de medición (encuesta OCDE, ILIA 2025, Latinobarómetro 2024) permitirán seguir la variable de resultado. La predicción falsable de este documento: las entidades de la región que

instalen las cinco prácticas mínimas de la tabla 1 antes de 2026 enfrentarán sus primeros incidentes con costos de confianza sustancialmente menores que las que no — y esa diferencia será visible en los casos, como lo fue en el catálogo de junio.

Recomendaciones.

1. **Instale las cinco prácticas mínimas de la tabla 1 como paquete**, no como menú: los pilares se sostienen entre sí (la transparencia sin competencia demostrada expone; la competencia sin transparencia no existe públicamente).
2. **Publique el registro de sistemas antes de que sea obligatorio**. Es la práctica de mejor relación señal/costo del marco completo, y la infraestructura colombiana de datos abiertos la hace inmediata.
3. **Prepare la respuesta a incidentes hoy**. Protocolo, vocería técnica, compromiso de informe público — el pilar 3 se juega entero en la primera crisis.
4. **Pacte la verificación externa en cada contrato** (cláusula 3 de agosto) y prográmela: una revisión externa anual publicada vale más que diez autoevaluaciones.
5. **Proteja la coherencia con documentos, no con personas**. Gobernanza versionada, compromisos escritos, transiciones documentadas: el pilar 5 es el único que ningún equipo saliente puede llevarse.
6. **Mida la confianza como resultado**. Incorpore las preguntas estándar (OCDE, Latinobarómetro) en las mediciones de la entidad y reporte la serie: la política de IA que no mide confianza está declarando que no la considera resultado.

Metodología.

Marco elaborado con corte al 31 de octubre de 2024 sobre tres insumos: la medición comparada de confianza institucional (encuesta OCDE 2024, con nota de país para Colombia; Latinobarómetro 2023; componentes de percepción del ILIA), el marco de determinantes de la confianza de la OCDE, y la evidencia acumulada de esta serie (catálogo de fallas, condiciones de supervisión, cláusulas contractuales, inventario colombiano). Los cinco pilares son una síntesis normativa de Enpath: su correspondencia con los determinantes OCDE está argumentada en el texto; su eficacia causal sobre la confianza es una hipótesis explícita, con la predicción falsable enunciada en perspectivas.

Limitaciones. Las mediciones de confianza citadas no aíslan el efecto de los sistemas de IA; el vínculo entre prácticas del marco y confianza agregada se apoya en el mecanismo documentado en casos (tiempo de detección, escrutinio), no en evidencia experimental. Las cifras de encuestas comparadas tienen las limitaciones muestrales y de comparabilidad declaradas por sus autores.

Referencias.

1. CENIA y CEPAL (2024). *Índice Latinoamericano de Inteligencia Artificial (ILIA) 2024*. Septiembre de 2024. <https://indicelatam.cl/>
2. Corte Constitucional de Colombia (2024). *Sentencia T-323 de 2024*. <https://www.corteconstitucional.gov.co/>
3. Gobierno del Reino Unido. *Algorithmic Transparency Recording Standard*. <https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub>
4. Latinobarómetro (2023). *Informe Latinobarómetro 2023*. <https://www.latinobarometro.org/>
5. OCDE (2024). *OECD Survey on Drivers of Trust in Public Institutions — 2024 Results*. Julio de 2024. https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results_9a20554b-en.html
6. OCDE (2024). *Country note: Colombia — Survey on Drivers of Trust 2024*. https://www.oecd.org/en/publications/2024/06/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results-country-notes_33192204/colombia_a71f1c24.html
7. OMB (2024). *Memoranda M-24-10 y M-24-18 (gobernanza y adquisición de IA)*. <https://www.whitehouse.gov/omb/>
8. Superintendencia de Industria y Comercio (2024). *Circular Externa 002 de 2024: tratamiento de datos personales en sistemas de inteligencia artificial*. <https://www.sic.gov.co/>

Enpath · Investigación aplicada para decisiones públicas basadas en evidencia. Cada publicación incluye su metodología, sus fuentes y sus limitaciones.