

# Evaluación de Preparación en IA para Instituciones Públicas

---

El modelo de madurez Enpath — y la predicción que se cumplió en las seis dimensiones

---

Febrero de 2025

**Enpath · Investigación · Marco (modelo de madurez) · Pilar: Gobernanza de la IA Publicado:** febrero de 2025 · **Corte de la evidencia:** 24 de julio de 2026 **Idioma:** español · **Ámbito:** entidades públicas de América Latina; instrumento de 24 afirmaciones **Clasificación de la evidencia:** modelo propio, ahora validado contra medición empírica (ver §16.3)

Colombia ya tiene política nacional de inteligencia artificial: el CONPES 4144, aprobado el 14 de febrero de 2025, con 106 acciones y un horizonte a 2030. La pregunta operativa llega ahora a cada entidad: ¿qué tan preparada está usted para lo que la política le pedirá? Este documento entrega el instrumento para responderla: seis dimensiones, cinco niveles, veinticuatro afirmaciones verificables — y una regla de ambición.

## Tabla de contenido

**Parte I — Síntesis** 1. Resumen ejecutivo 2. Hallazgos principales 3. Por qué importa

**Parte II — Contexto** 4. Antecedentes: por qué un modelo de madurez y no un índice 5. El contexto de febrero de 2025, y lo que ocurrió después 6. El campo: madurez, gestión de riesgos y certificación

**Parte III — El modelo** 7. Las seis dimensiones 8. Los cinco niveles 9. El instrumento: veinticuatro afirmaciones 10. La regla de ambición 11. Cómo se aplica

**Parte IV — La prueba** 12. El modelo contra la medición de 2026 13. Rendición de cuentas: las promesas de febrero de 2025

**Parte V — Prospectiva y decisión** 14. Riesgos, revalorados 15. Recomendaciones 16. Metodología, calidad de la evidencia y limitaciones 17. Perspectivas y predicciones falsables 18. Conclusión

**Anexos** - Anexo A. Las veinticuatro afirmaciones y su evidencia aceptada - Anexo B. Protocolo de autoevaluación en dos días - Anexo C. La regla de ambición: tabla de decisión - Anexo D. Glosario - Anexo E. Índice de figuras y tablas - Referencias

## Parte I — Síntesis

### 1. Resumen ejecutivo

La aprobación del CONPES 4144 —la política nacional de inteligencia artificial de Colombia, con seis ejes estratégicos, **106 acciones** y una inversión indicativa de **479.273 millones de pesos** a 2030<sup>1</sup> — cambió la pregunta relevante para las entidades públicas. Ya no es «¿qué dice la política?» sino «¿qué puede hacer mi entidad con lo que tiene?». La distancia entre las dos preguntas es exactamente lo que este documento mide.

El **Modelo Enpath de madurez institucional en IA** evalúa seis dimensiones —gobernanza y mandato, datos e infraestructura, talento y capacidad, procesos y salvaguardas, transparencia y rendición, uso y resultados— en cinco niveles (0 a 4), mediante un instrumento de **24 afirmaciones verificables**: cada una se cumple o no se cumple, con evidencia documental, sin espacio para el «parcialmente» que convierte las autoevaluaciones en ejercicios de cortesía.

El modelo añade lo que los índices internacionales de preparación no tienen: **consecuencia operativa**. La regla de ambición vincula el nivel de madurez con el tipo de sistema que la entidad debería operar. La preparación deja de ser un puntaje para presumir y se convierte en un límite de velocidad.

### 1.1. La predicción, evaluada

El modelo cerró con una predicción falsable sobre el perfil típico. Doce meses después, el benchmark de febrero de 2026 la puso a prueba sobre quince entidades del orden nacional colombiano:

| Dimensión                      | Perfil predicho (feb. 2025) | Mediana medida (feb. 2026) | ¿Coincide? |
|--------------------------------|-----------------------------|----------------------------|------------|
| D1 · Gobernanza y mandato      | 2                           | 2                          | Sí         |
| D2 · Datos e infraestructura   | 2                           | 2                          | Sí         |
| D3 · Talento y capacidad       | 1                           | 1                          | Sí         |
| D4 · Procesos y salvaguardas   | 1                           | 1                          | Sí         |
| D5 · Transparencia y rendición | 0                           | 0                          | Sí         |
| D6 · Uso y resultados          | 1                           | 1                          | Sí         |

Tabla 1. La predicción de febrero de 2025 contra la medición de febrero de 2026. Fuente: modelo de madurez Enpath; benchmark de capacidades (quince entidades, archivo `benchmark_2026.csv`).

**Coincidencia exacta en las seis dimensiones**, y el desbalance anticipado —gobernanza y datos por encima de transparencia y uso— aparece en **quince de quince perfiles**, sin una sola excepción. Ninguna entidad alcanzó el nivel 4, como el modelo anticipaba que ocurriría.

Un modelo puede acertar por dos razones: porque describe bien la realidad o porque su predicción era trivial. Conviene decir cuál es el caso. La predicción **no** era trivial: podía fallar de varias maneras verificables —transparencia en 1 en lugar de 0, ministerios por encima de entidades técnicas, dispersión alta entre entidades, alguna entidad en nivel 4— y no falló en ninguna. Lo que sí conviene declarar como límite es que quien predijo y quien midió son el mismo equipo (\$16.6).

### 1.2. La regla de ambición, aplicada por primera vez a entidades reales

Es el hallazgo más severo de esta edición, y no estaba disponible cuando el modelo se publicó.

La regla dice: automatización de bajo impacto desde el nivel 1; decisión asistida de impacto medio desde el nivel 2; alto impacto solo desde el nivel 3; y las decisiones indelegables, en ningún nivel. Y precisa que **el nivel relevante es el de la dimensión más baja pertinente, no el promedio**.

Aplicada literalmente a las quince entidades medidas:

| Tipo de sistema                    | Nivel exigido                          | Entidades habilitadas |
|------------------------------------|----------------------------------------|-----------------------|
| Automatización de bajo impacto     | 1 en todas las dimensiones pertinentes | 3 de 15               |
| Decisión asistida de impacto medio | 2                                      | 0 de 15               |
| Sistemas de alto impacto           | 3                                      | 0 de 15               |

Tabla 2. La regla de ambición aplicada al benchmark de febrero de 2026. Fuente: cálculo de Enpath.

**El resultado es incómodo en dos direcciones a la vez.** Hacia el Estado: doce de quince entidades no estarían habilitadas ni siquiera para el tramo más bajo de la regla, y ninguna para lo que varias ya operan. Hacia el propio modelo: **una regla que descalifica al 100 % de la muestra para el uso que el 100 % de la muestra hace es una regla que, aplicada literalmente, nadie va a aplicar.** La §12.2 trata las dos lecturas y propone la corrección: la regla no debe usarse como prohibición retroactiva sino como **criterio de dotación** — no «no puede operar esto», sino «para operar esto, esta es la dimensión que debe subir y este es el plazo».

### 1.3. El techo real del Estado medido

| Dimensión          | Mínimo | Mediana | Máximo de todo el Estado medido |
|--------------------|--------|---------|---------------------------------|
| D1 · Gobernanza    | 0      | 2       | 3 (una sola entidad)            |
| D2 · Datos         | 1      | 2       | 3 (una sola entidad)            |
| D3 · Talento       | 0      | 1       | 2                               |
| D4 · Procesos      | 0      | 1       | 2                               |
| D5 · Transparencia | 0      | 0       | 1                               |
| D6 · Uso           | 0      | 1       | 2                               |

Tabla 3. El techo por dimensión. Fuente: benchmark de capacidades, febrero de 2026.

**El nivel 4 —institucionalizado, el que sobrevive al relevo de personas— no lo alcanza nadie, en ninguna dimensión.** Es el dato que más importa para 2026–2027: la definición operativa del nivel 4 es precisamente la resistencia al cambio de gobierno, y el Estado colombiano llega a la transición del 7 de agosto de 2026 sin una sola capacidad institucionalizada según su propia medición.

## 2. Hallazgos principales

**H1. La predicción se cumplió en las seis dimensiones, y el desbalance en quince de quince perfiles.** Es la validación más completa que un modelo de madurez de esta serie ha recibido. La predicción era falsable de varias maneras y no falló en ninguna. *Confianza: alta, con el sesgo de autoría declarado en §16.6.*

**H2. La regla de ambición, aplicada literalmente, descalifica a todas las entidades medidas para lo que varias ya hacen.** Tres de quince estarían habilitadas para el tramo más bajo; ninguna para los otros dos. **La regla estaba bien diseñada como diagnóstico y mal diseñada como prohibición** — y por eso se formula como criterio de dotación con plazo (§12.2). *Confianza: alta sobre el cálculo; la reformulación es una decisión de diseño.*

**H3. El nivel 4 no existe en el Estado colombiano medido.** Ninguna entidad, en ninguna dimensión. Como el nivel 4 se define por la supervivencia al relevo de personas, el dato tiene una lectura inmediata para la transición del 7 de agosto de 2026: **nada de lo construido está protegido por definición del propio modelo.** *Confianza: alta.*

**H4. La dimensión de transparencia tiene un techo de 1 en todo el Estado medido.** No es solo que la mediana sea cero: es que ninguna entidad llega al nivel 2. En las otras cinco dimensiones hay al menos una entidad en 2 o más. **D5 no es la dimensión más rezagada por poco: es la única sin ningún caso demostrativo interno.** *Confianza: alta.*

**H5. El instrumento resistió el campo sin modificaciones.** Veinticuatro afirmaciones evaluadas documentalmente sobre quince entidades, con días de trabajo por entidad y sin acceso privilegiado. La medición de capacidades de IA a escala estatal es viable hoy; la excusa logística expiró. *Confianza: alta.*

**H6. Los índices nacionales y este modelo miden cosas distintas y ambas hacen falta.** El índice regional sitúa a Colombia tercero en adopción con 67,29 puntos; este modelo encuentra a sus entidades con transparencia en cero. **No hay contradicción: un país puede adoptar rápido y no poder demostrar nada de lo que adopta.** La utilidad del modelo es exactamente ver dentro del país lo que el índice ve desde fuera. *Confianza: alta.*

**H7. La aplicación censal que el modelo recomendó no ocurrió, y esa es la brecha operativa que queda.** El modelo pedía una aplicación censal a las entidades del orden nacional en 2025 para dar a la política nacional su línea de salida. Se hizo una muestra de quince entidades en 2026 por iniciativa propia de Enpath. **La línea de salida agregada del Estado colombiano sigue sin existir.** *Confianza: alta.*

**H8. Ninguno de los dos marcos internacionales de referencia vincula ambición con capacidad.** Ni el marco de gestión de riesgos —voluntario y no certificable— ni la norma de sistemas de gestión —certificable pero de cumplimiento binario— gradúan la capacidad institucional ni limitan el uso según ella. La regla de ambición sigue sin equivalente internacional. *Confianza: media-alta (revisión de dos marcos principales).*

### 3. Por qué importa

Tres decisiones dependen de saber el nivel real: **qué proyectos aprobar** (la regla de ambición como filtro de la primera puerta del marco de enero de 2025); **dónde invertir primero** (la dimensión más baja pertinente, no la más visible); y **cómo reportar el avance de la política nacional** (perfiles medidos y comparables, no listas de actividades).

Dieciocho meses después, las tres siguen abiertas y una tiene ahora respuesta parcial. La segunda: **la dimensión más baja es la misma en las quince entidades medidas** —transparencia—, lo que significa que la respuesta a «dónde invertir primero» es idéntica para todas y no requiere diagnóstico individual. Es un resultado que ahorra trabajo: cualquier entidad del orden nacional colombiano puede saltarse la autoevaluación para decidir su primera inversión, porque la evidencia agregada ya la determinó.

**Y para la administración que se posesiona el 7 de agosto de 2026,** el dato de la §1.3 es el que importa: ninguna capacidad del Estado medido está en nivel 4, es decir, ninguna está protegida contra el

relevo. La autoevaluación de una entidad en los primeros noventa días de la nueva administración no es un ejercicio burocrático: es el inventario de lo que puede perderse.

## Parte II — Contexto

### 4. Antecedentes: por qué un modelo de madurez y no un índice

Los modelos de madurez son una tecnología de gestión probada —de la ingeniería de software a la gestión de datos— cuya utilidad depende de dos condiciones que este diseño cuida: **criterios verificables** (no percepciones) y **consecuencias operativas** (no solo puntajes).

El modelo se apoya en los instrumentos ya publicados por la serie: las veinticuatro afirmaciones no son nuevas, **son la versión evaluable de lo que la serie ya especificó** — siete preguntas, diez cláusulas, protocolo de supervisión, ruta de datos, tablero directivo, cuatro puertas. El modelo es, deliberadamente, la serie completa convertida en lista de verificación.

| Instrumento de la serie                                                       | Dimensión que alimenta              |
|-------------------------------------------------------------------------------|-------------------------------------|
| Matriz de delegación y lista de lo indelegable (feb. 2024)                    | D1 · Gobernanza (límites)           |
| Marco de gobernanza de datos de entidad (may. 2024)                           | D2 · Datos                          |
| Catálogo de fallas y taxonomía de riesgos (jun. 2024)                         | D4 · Procesos                       |
| Inventario documental de sistemas (jul. 2024)                                 | D2 · Datos, D6 · Uso                |
| Diez cláusulas mínimas de compra (ago. 2024)                                  | D4 · Procesos                       |
| Protocolo de supervisión con métricas y dotación (sep. 2024)                  | D4 · Procesos, D5 · Transparencia   |
| Cinco pilares de confianza (oct. 2024)                                        | D5 · Transparencia                  |
| Siete preguntas del despacho y tablero directivo (nov. 2024)                  | D1 · Gobernanza, D5 · Transparencia |
| Marco de gobernanza ministerial: seis componentes, cuatro puertas (ene. 2025) | Las seis dimensiones                |

Tabla 4. Los instrumentos de la serie y las dimensiones que alimentan. Fuente: Enpath.

### 5. El contexto de febrero de 2025, y lo que ocurrió después

#### 5.1. Lo que el informe vio

El CONPES 4144 acababa de aprobarse —14 de febrero de 2025— con 106 acciones y horizonte 2030, y ninguna medición pública decía desde dónde partía cada entidad. Los índices existentes medían países, no entidades. El marco de gobernanza ministerial de enero de 2025 había definido los seis componentes y las cuatro puertas, pero no cómo saber si una entidad estaba lista para adoptarlos.

## 5.2. Lo que ocurrió después

| Fecha                    | Hecho                                                                               | Relevancia para el modelo                                       |
|--------------------------|-------------------------------------------------------------------------------------|-----------------------------------------------------------------|
| Abril–septiembre de 2025 | Se publican los instrumentos de registro, compra y evaluación de riesgo de la serie | Completan las afirmaciones de D4 y D5 con plantillas            |
| Octubre de 2025          | Índice regional de IA, tercera edición: Colombia con 67,29 en adopción <sup>4</sup> | Confirma la paradoja que el modelo traduce a escala de entidad  |
| Noviembre de 2025        | Scorecard de compra: veinte procesos, método documental con protocolo binario       | Valida el método que el modelo presupone                        |
| Febrero de 2026          | Benchmark de capacidades: quince entidades evaluadas con este instrumento           | La prueba de la predicción (§12)                                |
| Marzo de 2026            | Índice nominal: dieciocho ministerios, seis obligaciones, con nombres               | El ejercicio nominal que el modelo anunció                      |
| Junio–julio de 2026      | Aplazamiento europeo del régimen de alto riesgo a diciembre de 2027 <sup>6</sup>    | No afecta al modelo; sí a la urgencia externa que lo acompañaba |

Tabla 4. Lo que ocurrió después de febrero de 2025. Fuente: Enpath, con las fuentes primarias citadas.

## 6. El campo: madurez, gestión de riesgos y certificación

Esta sección compara el modelo con los dos marcos que son sus referentes naturales.

| Instrumento                               | Unidad de análisis            | ¿Verificable por terceros?     | ¿Certificable? | ¿Vincula ambición con capacidad? |
|-------------------------------------------|-------------------------------|--------------------------------|----------------|----------------------------------|
| Índices nacionales de preparación en IA   | País                          | Sí, con datos agregados        | No             | No                               |
| Marco de gestión de riesgos de IA (2023)  | El sistema y su ciclo de vida | Autodeclarado                  | No             | No                               |
| Norma de sistemas de gestión de IA (2023) | La organización               | Por auditoría de certificación | Sí             | No                               |
| Modelo Enpath de madurez                  | La entidad pública            | Sí, documentalmente            | No             | Sí — la regla de ambición        |

Tabla 5. El campo de instrumentos y el lugar del modelo. Fuente: Enpath, con base en los marcos citados<sup>23</sup>.

**Las tres decisiones que distinguen al modelo, con su fundamento:**

**Unidad de análisis.** Los índices miden países; la unidad que decide, compra y opera es la entidad. Y la evidencia de 2026 muestra por qué importa: Colombia puntúa 67,29 en adopción regional mientras sus entidades tienen transparencia con techo de 1. Ambas cosas son ciertas y solo la segunda dice qué hacer el lunes.

**Verificabilidad.** Evidencia documental binaria, no percepción del encuestado. Es lo que permitió aplicar el instrumento **a distancia y sin cooperación de las entidades**, cosa que ninguna autoevaluación de percepción permite.

**Consecuencia.** La regla de ambición no tiene equivalente en los marcos comparados. El marco de gestión de riesgos es voluntario y no certificable —la alineación es autodeclarada— y la norma de

sistemas de gestión certifica cumplimiento organizacional sin graduar capacidad ni limitar el uso según ella<sup>23</sup>.

**Y una recomendación adicional:** la norma internacional de sistemas de gestión de IA es exigible a los proveedores. Una entidad de nivel bajo que no puede subir su propia capacidad en el corto plazo sí puede trasladar parte de la exigencia al mercado, pidiéndola en el pliego donde sea proporcional. Es la única vía rápida disponible para una entidad con perfil bajo y un sistema de alto impacto ya contratado.

## Parte III — El modelo

### 7. Las seis dimensiones

Modelo Enpath de madurez institucional en IA · Seis dimensiones, cinco niveles

|                                | 0 · Inexistente  | 1 · Inicial                   | 2 · En formación                   | 3 · Gestionado             | 4 · Institucionalizado           |
|--------------------------------|------------------|-------------------------------|------------------------------------|----------------------------|----------------------------------|
|                                | Nada verificable | Esfuerzos aislados, sin regla | Reglas adoptadas, práctica parcial | Práctica completa y medida | Sobrevive a personas y gobiernos |
| D1 · Gobernanza y mandato      |                  |                               |                                    |                            |                                  |
| D2 · Datos e infraestructura   |                  |                               |                                    |                            |                                  |
| D3 · Talento y capacidad       |                  |                               |                                    |                            |                                  |
| D4 · Procesos y salvaguardas   |                  |                               |                                    |                            |                                  |
| D5 · Transparencia y rendición |                  |                               |                                    |                            |                                  |
| D6 · Uso y resultados          |                  |                               |                                    |                            |                                  |

Cada dimensión se califica con cuatro afirmaciones verificables (0–4 cumplidas = nivel 0–4).

Fuente: Enpath. Instrumento completo de 24 afirmaciones en [data/instrumento\\_autoevaluacion.csv](#).

Figura 1. El modelo: seis dimensiones, cinco niveles. Fuente: Enpath.

| Dimensión Enpath               | Qué mide                                               | Eje del CONPES 4144 relacionado |
|--------------------------------|--------------------------------------------------------|---------------------------------|
| D1 · Gobernanza y mandato      | Responsable, comité, puertas, límites                  | Ética y gobernanza              |
| D2 · Datos e infraestructura   | Inventario, calidad, acuerdos, terceros                | Datos e infraestructura         |
| D3 · Talento y capacidad       | Perfiles, formación, memoria, relevo                   | Capacidades y talento           |
| D4 · Procesos y salvaguardas   | Preguntas, cláusulas, protocolos, incidentes           | Mitigación de riesgos           |
| D5 · Transparencia y rendición | Registro público, métricas, impugnación, auditoría     | Ética y gobernanza              |
| D6 · Uso y resultados          | Líneas de base, decisiones cambiadas, pilotos cerrados | Uso y adopción                  |

Tabla 6. Dimensiones del modelo y su correspondencia con los ejes de la política nacional. Fuente: Enpath; CONPES 4144<sup>1</sup>.



## 8. Los cinco niveles

**0 · Inexistente** (nada verificable) · **1 · Inicial** (esfuerzos aislados sin regla) · **2 · En formación** (reglas adoptadas, práctica parcial) · **3 · Gestionado** (práctica completa y medida) · **4 · Institucionalizado** (sobrevive a personas y gobiernos: documentado, auditado, con historial).

La calificación es mecánica: cada dimensión tiene cuatro afirmaciones; el nivel es el número de afirmaciones cumplidas con evidencia.

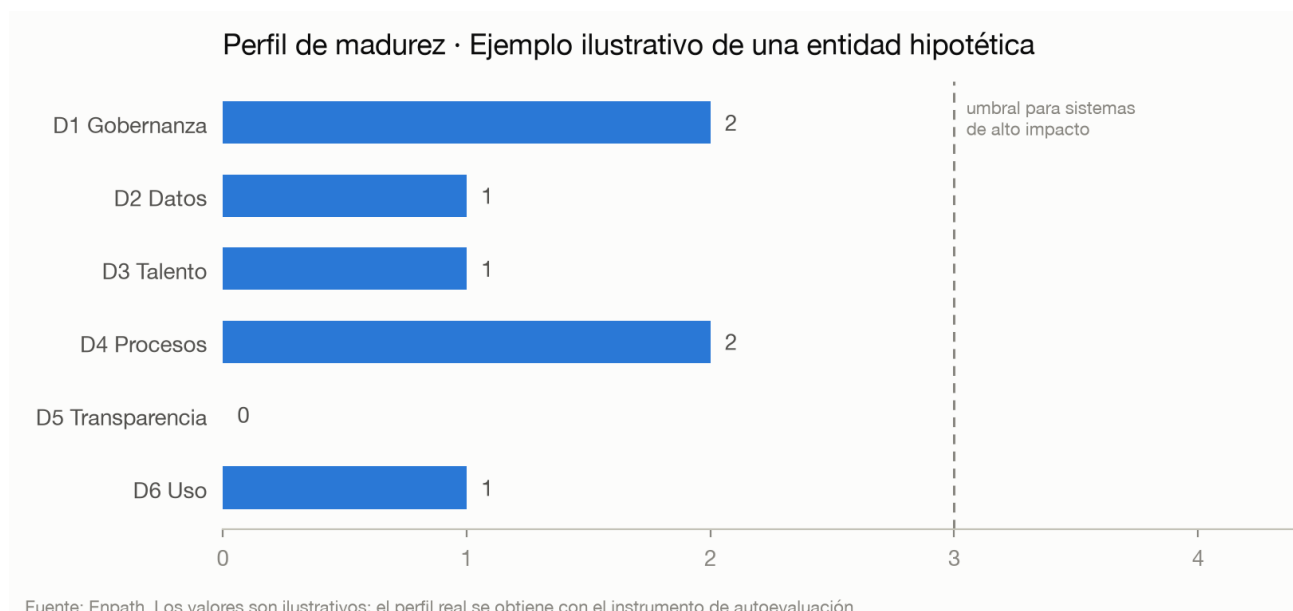


Figura 2. Perfil de una entidad hipotética. El desbalance —común— es el diagnóstico. Fuente: Enpath (ejemplo ilustrativo).

**Lo que la medición añadió sobre el nivel 4.** El modelo definió el nivel 4 como el que «sobrevive a personas y gobiernos». En febrero de 2026, **ninguna entidad lo alcanzó en ninguna dimensión** (§12.3, Figura 6). La definición dejó de ser una aspiración de escala y se convirtió en un diagnóstico: el Estado colombiano medido no tiene capacidad institucionalizada, y llega así a una transición de gobierno.

## 9. El instrumento: veinticuatro afirmaciones

Cada afirmación se cumple o no se cumple, con evidencia documental que la afirmación misma nombra. El detalle está en `data/instrumento_autoevaluacion.csv` y su desarrollo con criterios de evidencia aceptada, en el **Anexo A**.

**Lo que la aplicación de 2026 demostró sobre el instrumento:** se aplicó sin modificación, a distancia, con evidencia pública, en días por entidad, con ambigüedad relevante en el 8 % de los casos —resueltos contra el evaluado por protocolo—. **El instrumento resistió el campo**, que era la pregunta abierta más importante sobre él.

**Y lo que la aplicación dejó pendiente:** la matriz completa de veinticuatro afirmaciones por quince entidades **no se publicó**; el archivo publicado contiene los seis niveles por entidad. Las afirmaciones del benchmark sobre ítems concretos —cuáles puntúan cero en todas— no son reproducibles

externamente. Es una deuda de transparencia del propio proyecto y la §15.4 la convierte en recomendación.

## 10. La regla de ambición

La regla de ambición · El nivel de madurez limita lo que la entidad debería operar

|                               |                                                                                                     |
|-------------------------------|-----------------------------------------------------------------------------------------------------|
| <b>Nivel 1+</b>               | Automatización de bajo impacto con monitoreo<br>(trámites masivos reversibles; asistentes internos) |
| <b>Nivel 2+</b>               | Decisión asistida en impacto medio<br>(priorización con cola de excepciones y bitácora)             |
| <b>Nivel 3+</b>               | Decisión asistida de alto impacto<br>(revisión previa universal; métricas públicas)                 |
| <b>Nunca, en ningún nivel</b> | Las seis decisiones indelegables<br>(derechos fundamentales, sanciones, fuerza, fondo judicial...)  |

Fuente: Enpath. El nivel se toma de la dimensión más baja pertinente al sistema, no del promedio.

Figura 3. La regla de ambición. Fuente: Enpath.

| Tipo de sistema                             | Nivel exigido         | Fundamento                                        |
|---------------------------------------------|-----------------------|---------------------------------------------------|
| Automatización de bajo impacto y reversible | 1                     | Basta con que exista práctica, aunque sea aislada |
| Decisión asistida de impacto medio          | 2                     | Exige reglas adoptadas y práctica parcial         |
| Sistemas de alto impacto sobre personas     | 3                     | Exige práctica completa y medida                  |
| Las decisiones indelegables                 | Ningún nivel habilita | No es una cuestión de capacidad sino de límite    |

Tabla 7. La regla de ambición. Fuente: Enpath.

Dos precisiones evitan los malos usos, y la segunda exige una precisión. **Primera: el nivel relevante es el de la dimensión más baja pertinente al sistema**, no el promedio — un promedio de 2,5 con transparencia en cero no habilita nada que requiera transparencia. **Segunda: la regla limita la operación, no el aprendizaje** — una entidad de nivel 1 puede y debe pilotar, con bitácora y con fecha de fin.

La revisión que la medición obliga a hacer está en la §12.2: aplicada literalmente a entidades reales, la regla descalifica al 100 % de la muestra para el uso que buena parte de la muestra hace. Una regla con ese resultado no se aplica: se ignora. La reformulación —de prohibición a criterio de dotación con plazo— es la corrección de diseño principal de esta edición.

## 11. Cómo se aplica

**Autoevaluación:** dos días de trabajo del comité de gobernanza con los documentos sobre la mesa, sin consultores. El protocolo está en el **Anexo B**.

**Evaluación documental externa:** un tercero aplica el instrumento contra evidencia pública, sin cooperación de la entidad. Es el método que el benchmark de 2026 usó y que el scorecard de noviem-

bre de 2025 validó. Su ventaja es la comparabilidad; su costo, que subestima lo que la entidad hace y no publica.

**Verificación mixta (recomendada, nueva en esta edición):** evaluación documental externa más un módulo de solicitud de información dirigido a los ceros, que distingue tres situaciones con remedios distintos: el documento no existe; existe y no se publica; sin respuesta. La calificación no cambia —el modelo mide demostrabilidad— pero el diagnóstico sí.

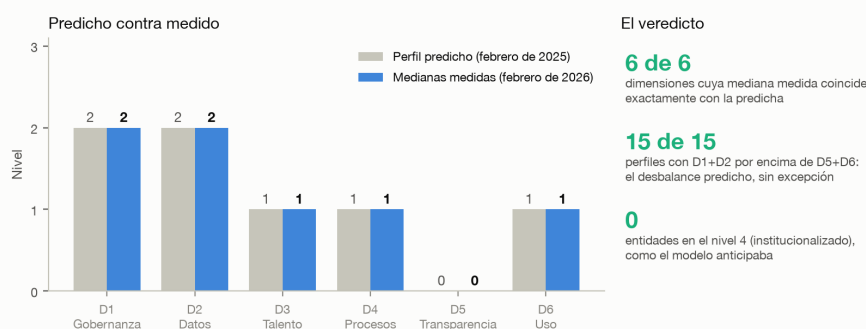
## Parte IV — La prueba

### 12. El modelo contra la medición de 2026

Esta parte es enteramente nueva. El original se publicó declarando que el modelo no había sido aplicado y que su calibración con datos era el compromiso de 2025. Esta sección usa la aplicación.

#### 12.1. La predicción, evaluada dimensión por dimensión

Figura 4 · La predicción de febrero de 2025, evaluada un año después: se cumplió en las seis dimensiones



Fuentes: perfil típico predicho por el modelo de madurez (febrero de 2025); benchmark Enpath de capacidades (febrero de 2026, quince entidades, archivo benchmark\_2026.csv). La predicción se publicó doce meses antes de la medición que la evalúa.

*Figura 4. La predicción de febrero de 2025, evaluada un año después: se cumplió en las seis dimensiones. Fuentes: modelo de madurez (febrero de 2025); benchmark de capacidades (febrero de 2026).*

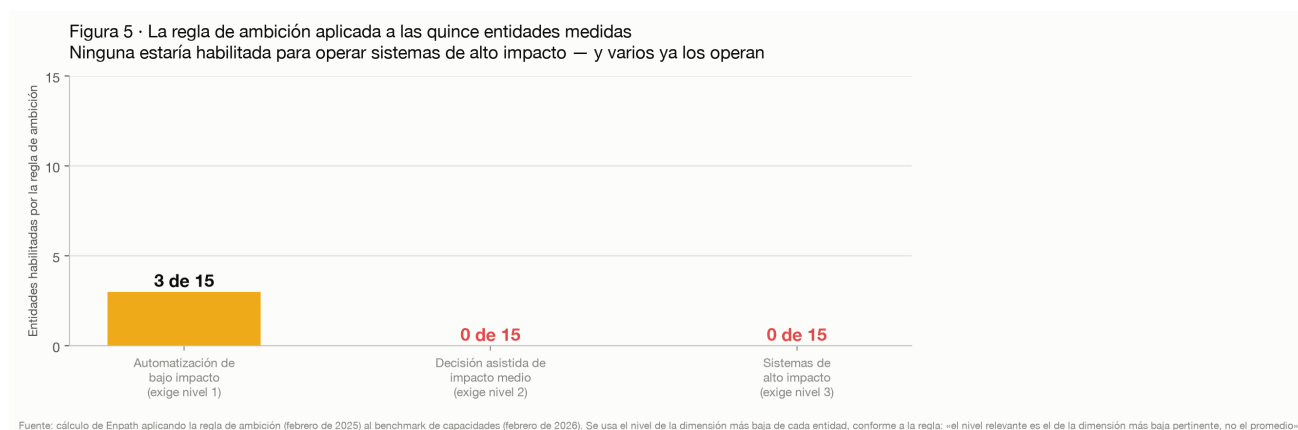
El benchmark de febrero de 2026 aplicó este instrumento a quince entidades del orden nacional colombiano —seis ministerios, tres entidades técnicas, dos superintendencias, una entidad de justicia y tres misionales— con evaluación documental pública y protocolo de réplica.

**Resultado: coincidencia exacta en las seis medianas** (Tabla 1), y **quince de quince perfiles** con el desbalance predicho.

**Por qué la predicción no era trivial.** Podía fallar de al menos cuatro maneras verificables: (i) transparencia en 1 en lugar de 0 —bastaba con que la mitad de las entidades tuviera alguna afirmación de D5—; (ii) ministerios por encima de entidades técnicas, que era la expectativa intuitiva dado que los ministerios dirigen la política; (iii) alguna entidad en nivel 4, lo que habría refutado la afirmación de que la institucionalización no existe; (iv) dispersión alta entre entidades, que habría invalidado la idea de un «perfil típico». No falló en ninguna.

**Y una precisión:** la predicción acertó incluso en la asimetría entre tipos de entidad. El modelo anticipó que el perfil típico tendría reglas y datos por delante de transparencia y resultados; la medición encontró además que **las entidades técnicas superan sistemáticamente a los ministerios** —las que operan sistemas hace años tienen más capacidad que las que dirigen la política—, una inversión que el modelo insinuaba al distinguir entre «reglas» y «resultados» pero no había formulado explícitamente.

## 12.2. La regla de ambición, aplicada — y su revisión



*Figura 5. La regla de ambición aplicada a las quince entidades medidas. Fuente: cálculo de Enpath sobre el benchmark de febrero de 2026.*

Aplicando la regla tal como está escrita —el nivel relevante es el de la dimensión más baja pertinente— a las quince entidades: **tres habilitadas para automatización de bajo impacto, cero para decisión asistida de impacto medio, cero para sistemas de alto impacto.**

**Las dos lecturas, y ambas son verdaderas.**

**Lectura hacia el Estado: la regla describe un desajuste real y grave.** El inventario documental de julio de 2024 registró sistemas que apoyan decisiones sobre personas operando en entidades que, medidas dieciocho meses después, no alcanzan el nivel que la regla exigiría. No es una objeción teórica: es la constatación de que la capacidad institucional va por detrás del despliegue, que es exactamente la condición que produjo los casos del catálogo de fallas de junio de 2024.

**Lectura hacia el modelo: una regla que descalifica al 100 % es una regla que nadie aplicará.** Una norma cuyo cumplimiento literal exigiría apagar sistemas en operación en todas las entidades medidas no se convierte en política: se convierte en excepción permanente. El propio original anticipó este riesgo —«la regla de ambición ignorada por excepción»— y lo calificó de probable. La medición lo confirma y agrega la razón: **el problema no es la voluntad de las entidades, es que la regla no ofrece salida.**

**La reformulación, que es la decisión de diseño principal.** La regla de ambición no se enuncia como prohibición y pasa a enunciarse como **criterio de dotación con plazo:**

**Regla de ambición, versión 2.** Un sistema de nivel de impacto N exige que la entidad alcance el nivel correspondiente en las dimensiones pertinentes a ese sistema. Si la entidad no lo alcanza: (a) para sistemas nuevos, la regla opera como prohibición: no se aprueba en la puerta 1 hasta alcanzar el nivel; (b) para sistemas en operación, la regla opera como plan obligatorio: el comité declara qué dimensión falta, fija plazo y designa responsable, y esa declaración se publica junto con la ficha del

sistema; (c) en ambos casos, **la brecha se declara, no se oculta** — y su publicación es lo que la hace corregible.

La diferencia práctica es grande: la versión 1 le decía a una entidad «no debería estar operando esto» y no le decía qué hacer con lo que ya opera. La versión 2 convierte la brecha en un compromiso fechado y público — que es lo que toda esta serie sostiene que funciona.

### 12.3. El techo real y la ausencia de nivel 4



Figura 6. El techo real del Estado colombiano por dimensión. Fuente: benchmark de capacidades, febrero de 2026.

Tres lecturas de la figura.

**El nivel 3 es alcanzable pero excepcional.** Se alcanza en dos dimensiones —gobernanza y datos— y en una sola entidad, una entidad técnica. Es la demostración de que el nivel gestionado no requiere condiciones extraordinarias: lo logró una organización colombiana con las reglas del país.

**Tres dimensiones tienen techo 2 y una tiene techo 1.** En talento, procesos y uso, **ninguna entidad del Estado medido llega al nivel 3.** En transparencia, ninguna llega al 2. Esto significa que para esas cuatro dimensiones no existe un caso demostrativo interno al que otras entidades puedan aspirar: **no hay a quién copiar dentro del país.**

**El nivel 4 no existe.** Ninguna entidad, ninguna dimensión. Y como el nivel 4 se define por la supervivencia al relevo de personas y de gobiernos, el dato tiene una lectura inmediata: **el Estado colombiano llega a la transición del 7 de agosto de 2026 sin una sola capacidad institucionalizada según su propia medición.** Lo que se pierda en la transición se perderá porque nada estaba protegido, no porque alguien decidiera desmontarlo.

12.4. El circuito que el modelo anunció

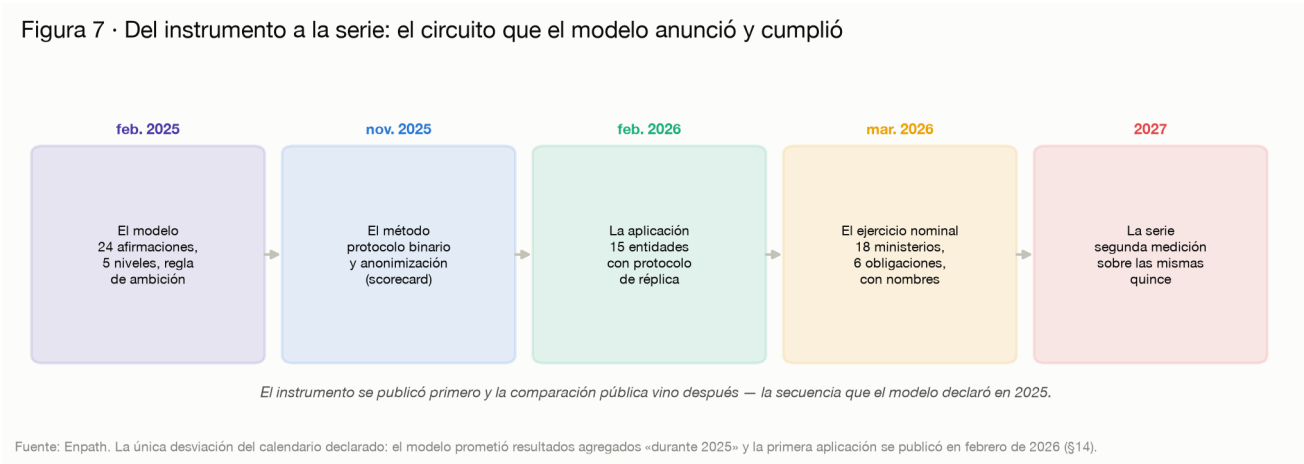


Figura 7. Del instrumento a la serie: el circuito que el modelo anunció y cumplió. Fuente: Enpath.

El original declaró una secuencia: «primero el instrumento y la práctica de medirse; después la comparación pública». Se cumplió: instrumento (febrero de 2025) → método documental validado (noviembre de 2025) → aplicación anonimizada por tipo (febrero de 2026) → ejercicio nominal con nombres (marzo de 2026) → serie comprometida (2027).

La única desviación es de calendario y está en la §13.

13. Rendición de cuentas: las promesas de febrero de 2025

| Promesa                                                                                                     | Resultado                                                                 | Veredicto                                                                                 |
|-------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| «El desbalance típico predicho aquí (D1–D2 sobre D5–D6) aparecerá en la mayoría de los perfiles reales»     | Apareció en <b>quince de quince</b>                                       | <b>Acierto pleno</b> , por encima de lo predicho («la mayoría»)                           |
| «Enpath aplicará el modelo durante 2025 con entidades voluntarias y publicará los resultados agregados»     | La aplicación se publicó en <b>febre-ro de 2026</b>                       | <b>Cumplido con retraso</b> de aproximada-mente dos meses respecto del año de-clarado     |
| «Su integración como base del Modelo de capaci-dades (febrero de 2026) y del Índice (marzo de 2026)»        | Ambos se publicaron en las fechas anunciadas                              | <b>Acierto</b>                                                                            |
| «Compromiso de revisión anual con datos de apli-cación»                                                     | Esta edición es la primera revisión con datos                             | <b>Cumplido</b> , con esta edición                                                        |
| «Aplicación censal del instrumento a las entida-des del orden nacional en 2025» (recomendación al Gobierno) | No ocurrió; se hizo una muestra de quince entidades por iniciativa propia | <b>No cumplido</b> — no era una promesa de Enpath sino una recomendación, y queda vigente |

Tabla 8. Las promesas de febrero de 2025, evaluadas. Fuente: Enpath.

**Dos observaciones de método.** La primera: la promesa de aplicación «con entidades voluntarias» se cumplió por una vía distinta de la anunciada — la aplicación de febrero de 2026 fue **documental y externa**, con periodo de réplica, no voluntaria. El resultado es mejor para la comparabilidad y peor para el detalle interno; la diferencia debió haberse anunciado y no se anunció.

La segunda: la predicción estaba formulada de manera conservadora («la mayoría de los perfiles») y el resultado fue universal (quince de quince). Una predicción que acierta por encima de su propio umbral es un acierto, pero también una señal de que el umbral era demasiado laxo. Las predicciones de la §17 se formulan con umbrales exactos.

## Parte V — Prospectiva y decisión

### 14. Riesgos, revalorados

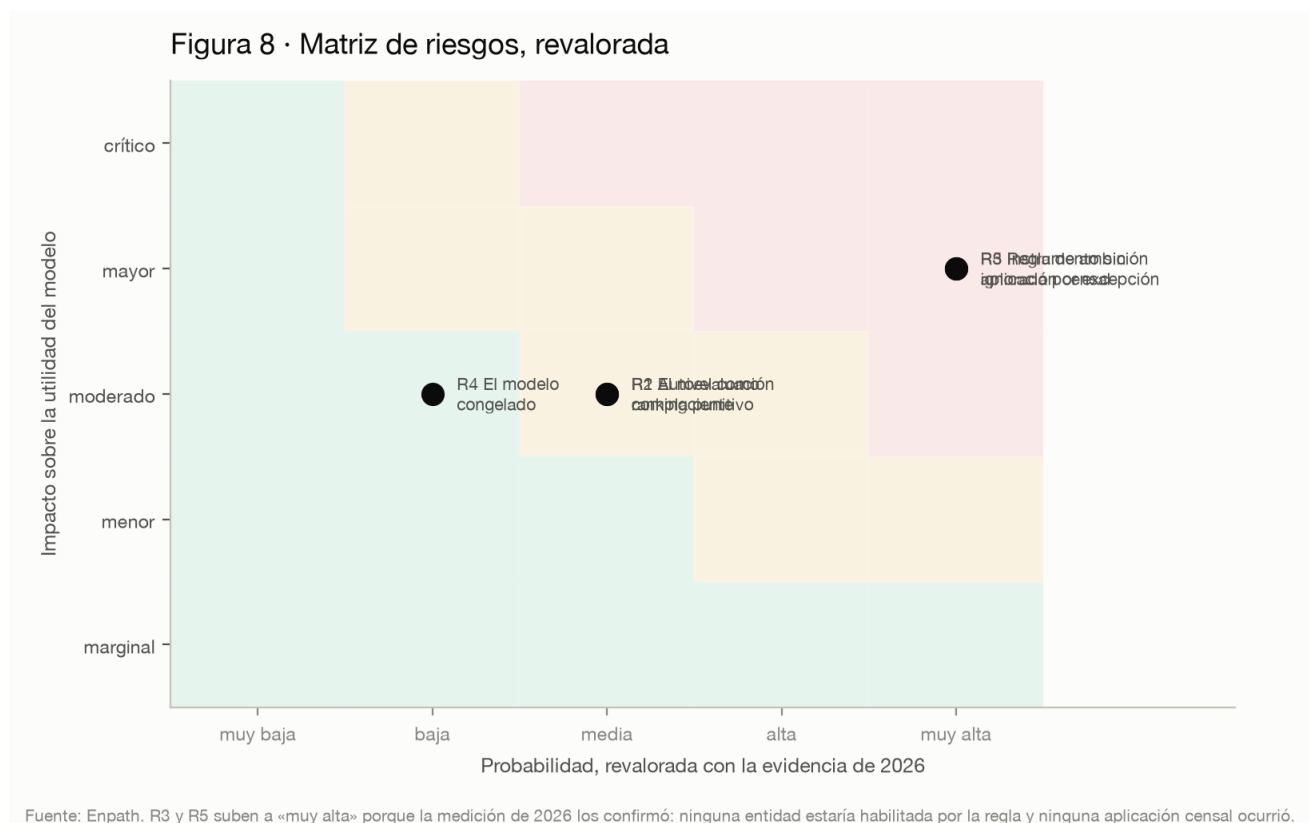


Figura 8. Matriz de riesgos, revalorada. Fuente: Enpath.

| Id | Riesgo                                      | Valoración original     | Valora-<br>ción 2026    | Fundamento                                                                                                                                                              |
|----|---------------------------------------------|-------------------------|-------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| R1 | Autoevaluación complaciente                 | Riesgo clásico          | Media / moderado        | No se materializó porque <b>no hubo autoevaluaciones</b> : la única aplicación fue documental y externa                                                                 |
| R2 | El nivel como ranking punitivo              | Enunciado               | Media / moderado        | Mitigado en la práctica por la anonimización de febrero de 2026 y el protocolo de réplica del ejercicio nominal                                                         |
| R3 | La regla de ambición ignorada por excepción | Enunciado como probable | <b>Muy alta / mayor</b> | <b>Confirmado y agravado</b> : la regla no descalifica a algunas entidades, descalifica a todas (§12.2). Corregido con la versión 2 de la regla                         |
| R4 | El modelo congelado                         | Enunciado               | Baja / moderado         | El instrumento resistió el campo sin modificación; esta edición es la primera revisión con datos                                                                        |
| R5 | Instrumento sin aplicación censal           | No contemplado          | <b>Muy alta / mayor</b> | <b>Riesgo nuevo</b> . Dieciocho meses después, la línea de salida agregada del Estado colombiano sigue sin existir, y la política nacional sigue reportando actividades |

Tabla 9. Riesgos, revalorados. Fuente: Enpath.

## 15. Recomendaciones

### 15.1. Para cada entidad

1. **Aplice el instrumento este semestre** — dos días del comité de gobernanza con los documentos sobre la mesa, sin consultores. El protocolo está en el Anexo B.
2. **Empiece por D5, sin diagnóstico previo**. La evidencia agregada ya determinó la respuesta: la dimensión más baja es la misma en las quince entidades medidas, y su techo en todo el Estado es 1. Cualquier entidad del orden nacional puede saltarse la pregunta de dónde invertir primero.
3. **Adopte la regla de ambición en su versión 2 (§12.2)**: prohibición para sistemas nuevos, plan fechado y público para los que ya operan.
4. **Publique el perfil con su plan** — los ceros incluidos. Es la aplicación más directa del marco de confianza, y a julio de 2026 **ninguna entidad de la región lo ha hecho**: la primera que publique su perfil define el estándar.
5. **Antes de una transición de gobierno, documente lo que está en nivel 2 o 3**. Ninguna capacidad medida está en nivel 4, es decir, ninguna sobrevive por sí sola al relevo.

### 15.2. Para los coordinadores de la política nacional

**La aplicación censal sigue pendiente y sigue siendo la recomendación de mayor retorno**. Aplicar el instrumento —o uno equivalente— a las entidades del orden nacional produciría, en semanas, la línea de salida que 106 acciones necesitan y no tienen. Sin ella, la implementación se reportará en actividades durante los cinco años que le quedan.

### 15.3. Para la compra pública

**Incorpore el nivel requerido al estudio previo**. «¿Qué nivel exige este sistema y en qué nivel estamos?» es la pregunta que convierte el modelo en herramienta de compra. Y para las entidades con perfil bajo y sistemas de alto impacto ya contratados, la vía rápida disponible es trasladar



exigencia al proveedor: **la norma internacional de sistemas de gestión de IA es certificable y exigible en el pliego** donde sea proporcional (§6).

#### 15.4. Para Enpath

1. **Publicar la matriz de veinticuatro afirmaciones por quince entidades (§9).** Las afirmaciones del benchmark sobre ítems concretos no son hoy reproducibles, y una de ellas sostiene su principal recomendación de política.
2. **Formular las predicciones con umbrales exactos, no con «la mayoría» (§13).**
3. **Aplicar el módulo de composición del cero en la segunda medición, para distinguir «no existe» de «no se publica» (§11).**

### 16. Metodología, calidad de la evidencia y limitaciones

#### 16.1. Objetivo y preguntas

**Objetivo.** Entregar un instrumento verificable que permita a una entidad pública saber su nivel real de preparación para operar sistemas de IA, con consecuencia operativa sobre qué debería y qué no debería operar.

**Preguntas.** (i) ¿Qué dimensiones componen la preparación institucional en IA? (ii) ¿Cómo se mide sin depender de percepciones? (iii) ¿Qué debe poder operar una entidad según su nivel? (iv) ¿Se cumplió la predicción del perfil típico? (v) ¿Qué muestra la regla de ambición aplicada a entidades reales?

#### 16.2. Diseño

Modelo elaborado con corte al 28 de febrero de 2025 y ampliado en julio de 2026. Las 24 afirmaciones operacionalizan los instrumentos publicados por la serie, contrastados con los marcos internacionales citados y alineados con los ejes de la política nacional. Los niveles y la regla de ambición son diseño propio.

**Lo nuevo de esta edición:** la evaluación de la predicción contra el benchmark de febrero de 2026 (§12.1); la primera aplicación de la regla de ambición a entidades reales y su reformulación (§12.2); el análisis del techo por dimensión (§12.3); y la rendición de cuentas de las promesas de febrero de 2025 (§13).

### 16.3. Calidad de la evidencia

| Nivel           | Definición                                               | Aplicado a                                                                                    |
|-----------------|----------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| A               | Medición documental primaria propia con datos publicados | Las medianas, máximos y perfiles del benchmark de febrero de 2026                             |
| B               | Cálculo propio verificable sobre datos de nivel A        | La aplicación de la regla de ambición (§12.2); el techo por dimensión (§12.3)                 |
| A               | Documento oficial primario                               | CONPES 4144; marco de gestión de riesgos; norma de sistemas de gestión; índice regional de IA |
| D               | Valoración cualitativa estructurada                      | Matriz de riesgos revalorada; Tabla 6 (comparación de instrumentos)                           |
| Diseño          | Formulación propia                                       | Las seis dimensiones, los cinco niveles, las 24 afirmaciones, la regla de ambición (v1 y v2)  |
| No reproducible | Afirmación sin respaldo en archivo publicado             | Las cuatro afirmaciones en cero universal reportadas por el benchmark (§9)                    |

Tabla 10. Escala de calidad de la evidencia. Fuente: Enpath.

### 16.4. Qué contaría como error de este informe

1. Si una réplica independiente con el mismo instrumento produce perfiles sustancialmente distintos para entidades equivalentes, el instrumento no es suficientemente objetivo.
2. Si una entidad alcanza el nivel 4 en alguna dimensión antes de 2027, el diagnóstico sobre la ausencia de institucionalización era demasiado pesimista.
3. Si la regla de ambición en su versión 2 tampoco se adopta, el problema no era su formulación sino la ausencia de obligación, y la corrección de la §12.2 es insuficiente.
4. Si la publicación de la matriz de ítems revela que alguna de las cuatro afirmaciones en cero universal tiene valores positivos, la recomendación de política que se apoya en ella está mal construida.

### 16.5. Limitaciones

**El instrumento binario sacrifica matiz por verificabilidad** — decisión deliberada cuyo costo (subestimar avances parciales) queda declarado.

**Quince entidades no son el Estado**, y la selección sobrerrepresenta a las entidades con presencia en la agenda de IA; el sesgo es conservador respecto del hallazgo principal.

**La regla de ambición fija umbrales discretos sobre un continuo de riesgo**; los casos límite exigen juicio documentado, no aritmética.

**La aplicación de la regla en la §12.2 usa la dimensión mínima de cada entidad**, conforme al texto de la regla. Un lector podría argumentar que solo deben considerarse las dimensiones «pertinentes al sistema», que serían menos de seis para algunos sistemas; con ese criterio más laxo el número de entidades habilitadas subiría, aunque no cambiaría el resultado de cero para alto impacto.

**La matriz de ítems del benchmark no está publicada** (§9, §16.3).

**La correspondencia con los ejes de la política nacional se basa en el documento aprobado**, no en su implementación.

16.6. Sesgos declarados

- 1. **Sesgo de autoría en la validación.** Enpath formuló la predicción y produjo la medición que la evalúa. Es el sesgo estructural de este trabajo y no puede mitigarse desde dentro; lo único disponible es que el instrumento, el protocolo y los datos del benchmark están publicados para que un tercero repita la medición.
- 2. **Sesgo de confirmación en la formulación de la predicción.** La predicción se formuló con base en la evidencia de la propia serie (línea de base de julio de 2024), lo que la hacía más probable de acertar que una predicción a ciegas.
- 3. **Sesgo de selección de la muestra.** Entidades con presencia en la agenda de IA, no muestra probabilística.
- 4. **Sesgo de instrumento propio.** Las 24 afirmaciones derivan de los instrumentos de Enpath; una entidad que gobierne su IA con otro método podría puntuar bajo sin déficit real.

17. Perspectivas y predicciones falsables

Con umbrales exactos, según la corrección de método de la §13.

| #  | Predicción                                                                                                                            | Umbral                                           | Cómo se falsea                                                                          |
|----|---------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|-----------------------------------------------------------------------------------------|
| P1 | Ninguna entidad pública colombiana alcanzará el nivel 4 en ninguna dimensión en la medición de 2027                                   | Exacto: cero entidades                           | Si alguna lo alcanza, el diagnóstico sobre la institucionalización era pesimista        |
| P2 | El techo de D5 (transparencia) seguirá en 1 en 2027 si no se expide obligación de registro; llegará a 2 o más si se expide            | Exacto                                           | Si sube sin obligación, la hipótesis del mecanismo es falsa                             |
| P3 | Ninguna entidad de América Latina publicará su perfil de madurez completo — con los ceros — antes de julio de 2027                    | Exacto: cero entidades                           | Si alguna lo publica, se documentará como el primer caso                                |
| P4 | La brecha entre el perfil mediano de las entidades técnicas y el de los ministerios no se reducirá en 2027 sin obligación transversal | Diferencia agregada igual o mayor que la de 2026 | Si se reduce sin obligación, la asimetría no es estructural                             |
| P5 | La aplicación censal recomendada al Gobierno nacional no se habrá realizado antes de julio de 2027                                    | Exacto                                           | Si se realiza, la predicción falla y el resultado será el dato más valioso de la agenda |

Tabla 11. Predicciones falsables. Fuente: Enpath.

18. Conclusión

Un modelo de madurez se juzga por si lo que predice ocurre. Este predijo, en febrero de 2025, que el perfil típico de las entidades públicas mostraría gobernanza y datos en formación y transparencia en cero. Doce meses después, aplicado a quince entidades del Estado colombiano con evaluación documental y protocolo de réplica, **las seis medianas coincidieron exactamente con las predichas y el desbalance apareció en quince de quince perfiles**. La predicción podía fallar de al menos cuatro maneras verificables y no falló en ninguna.

Ese acierto obliga a decir lo que el mismo dato revela sobre el instrumento que lo produjo. **La regla de ambición —la pieza que distingue a este modelo de cualquier índice de preparación— aplicada**

por primera vez a entidades reales, descalifica a doce de quince para el tramo más bajo y a las quince para lo que varias ya hacen. Es simultáneamente el diagnóstico más importante del modelo y su mayor problema de diseño: una norma cuyo cumplimiento literal exigiría apagar sistemas en operación en todas las entidades medidas no se convierte en política, se convierte en excepción permanente. El propio original lo anticipó como riesgo probable; la medición lo confirmó y agregó la razón — la regla no ofrecía salida.

De ahí la corrección principal de esta edición: **la regla deja de ser una prohibición y pasa a ser un criterio de dotación con plazo**. Prohibición para los sistemas nuevos, que aún pueden no aprobarse; plan fechado, con responsable y público, para los que ya operan. La brecha se declara en lugar de ocultarse, porque declararla es lo único que la hace corregible.

Queda un dato que no es sobre el modelo sino sobre el Estado que midió, y es el que más pesa en este momento. **Ninguna de las quince entidades alcanzó el nivel 4 en ninguna de las seis dimensiones**. El nivel 4 se define, desde febrero de 2025, como aquel que sobrevive a personas y gobiernos. Colombia llegó a la transición del 7 de agosto de 2026 sin una sola capacidad institucionalizada según su propia medición — lo que significa que lo que se pierda en el relevo no se perderá por decisión de nadie, sino porque nada estaba protegido.

Y queda la recomendación de mayor retorno, que nadie ha ejecutado: **una aplicación censal del instrumento a las entidades del orden nacional**. Costaría semanas, produciría la línea de salida que 106 acciones necesitan y no tienen, y convertiría el reporte de la política nacional de una lista de actividades en una serie temporal de capacidades. Y ya no es una propuesta sin respaldo: existe la prueba de que el instrumento funciona en el campo, aplicado a distancia, en días por entidad y sin cooperación de las entidades evaluadas.

---

## Anexos

---

### Anexo A. Las veinticuatro afirmaciones y su evidencia aceptada

El instrumento completo está en `data/instrumento_autoevaluacion.csv`. Para cada dimensión se fija además **qué se acepta como evidencia** — el criterio que la aplicación en el campo obliga a explicitar.

**Regla general de calificación.** Cada afirmación se cumple (1) o no se cumple (0). El nivel de la dimensión es el número de afirmaciones cumplidas. **La duda se resuelve contra el evaluado.**

## D1 · Gobernanza y mandato

| # | Afirmación verificable                                                  |
|---|-------------------------------------------------------------------------|
| 1 | Existe un responsable de IA designado por acto administrativo           |
| 2 | Opera un comité de gobernanza que deja actas de sus decisiones          |
| 3 | Las cuatro puertas del ciclo de vida están adoptadas como regla interna |
| 4 | Las decisiones indelegables están definidas por escrito                 |

**Evidencia aceptada:** acto administrativo publicado con nombre o cargo del responsable (1); actas fechadas de al menos dos sesiones (2); acto o procedimiento interno que condicione el avance de proyectos a los requisitos de las puertas (3); documento que enumere las decisiones que no se apoyarán en sistemas automatizados (4). **No se acepta:** menciones en planes estratégicos; comités de tecnología sin mandato explícito sobre IA; declaraciones de principios sin lista de decisiones. *La afirmación 3 puntuó cero en las quince entidades del benchmark.*

## D2 · Datos e infraestructura

| # | Afirmación verificable                                                 |
|---|------------------------------------------------------------------------|
| 5 | Existe inventario de bases de datos críticas con dueño por cada una    |
| 6 | La calidad de los dominios críticos se mide periódicamente             |
| 7 | Los intercambios de datos tienen acuerdos escritos vigentes            |
| 8 | Los contratos de custodia tecnológica incluyen continuidad y auditoría |

**Evidencia aceptada:** inventario fechado con custodio identificado por base (5); medición publicada o documentada con al menos dos observaciones (6); acuerdos vigentes con finalidad y campos (7); cláusulas verificables en los contratos de custodia (8).

## D3 · Talento y capacidad

| #  | Afirmación verificable                                                  |
|----|-------------------------------------------------------------------------|
| 9  | Los perfiles técnicos críticos están identificados y cubiertos          |
| 10 | Los revisores reciben formación documentada en límites de los sistemas  |
| 11 | Los métodos y decisiones técnicas quedan documentados en la entidad     |
| 12 | La rotación no ha dejado funciones críticas sin relevo en el último año |

**Nota de esta edición:** la afirmación 10 tiene ahora respaldo normativo externo — desde el 2 de febrero de 2025, el artículo 4 del reglamento europeo exige a proveedores y responsables del despliegue asegurar un nivel suficiente de alfabetización en IA de su personal<sup>5</sup>. Ninguna entidad latinoamericana está sujeta a esa norma, pero es un estándar citable. **La afirmación 12 será la más informativa de la medición de 2027**, tras la transición de gobierno del 7 de agosto de 2026.

## D4 · Procesos y salvaguardas

| #  | Afirmación verificable                                              |
|----|---------------------------------------------------------------------|
| 13 | Todo proyecto nuevo responde las siete preguntas por escrito        |
| 14 | Los contratos vigentes de sistemas incluyen las cláusulas mínimas   |
| 15 | Cada sistema de impacto tiene protocolo de supervisión con dotación |
| 16 | Existe protocolo de respuesta a incidentes con responsable          |

**Referencia externa de magnitud:** el scorecard de noviembre de 2025 midió estas mismas exigencias sobre veinte procesos de compra reales — problema escrito en 7 de 20, cláusulas mínimas en 6 de 20, protocolo de supervisión en 3 de 14. Una entidad que se autoevalúe muy por encima de esas proporciones debería verificar la calidad de su evidencia.

## D5 · Transparencia y rendición

| #  | Afirmación verificable                                                    |
|----|---------------------------------------------------------------------------|
| 17 | El registro de sistemas es público y está actualizado                     |
| 18 | Las métricas de supervisión se publican trimestralmente                   |
| 19 | Las personas afectadas cuentan con vía de impugnación con plazo           |
| 20 | La última auditoría o revisión externa tiene fecha y hallazgos publicados |

**Es la dimensión con techo 1 en todo el Estado medido.** Las afirmaciones 17, 18 y 20 puntuaron cero en las quince entidades del benchmark; el punto de las tres entidades con D5 = 1 proviene de la afirmación 19. **Evidencia aceptada para la 18:** publicación con al menos dos observaciones en el tiempo o periodicidad declarada. Indicadores de gestión del proyecto —avance, presupuesto ejecutado— **no cuentan:** la afirmación se refiere a métricas de supervisión del sistema.

## D6 · Uso y resultados

| #  | Afirmación verificable                                                     |
|----|----------------------------------------------------------------------------|
| 21 | Cada sistema tiene línea de base y métricas de resultado pactadas          |
| 22 | Los resultados medidos han modificado al menos una decisión documentada    |
| 23 | Los pilotos terminan con evaluación y decisión formal de adopción o cierre |
| 24 | El tablero directivo se revisa trimestralmente con datos completos         |

**La afirmación 22 es la más exigente del instrumento y la más importante.** Es la única que pregunta si la evidencia cambió algo — el «eslabón del uso» que toda esta serie documenta. En la medición de febrero de 2026, la mediana de D6 fue 1 y su máximo, 2.

## Anexo B. Protocolo de autoevaluación en dos días

**Participantes:** el comité de gobernanza, con acceso a los documentos. **No** consultores; **no** el área de tecnología en solitario.

**Día 1 · mañana (3 h) — Reunir la evidencia.** Cada área trae los documentos que cree que sostienen sus afirmaciones. Se listan sin discutir. *Producto: carpeta de evidencia.*

**Día 1 · tarde (3 h) — D1 a D3.** Afirmación por afirmación, con la regla de la duda contra el evaluado. Quien afirma el cumplimiento muestra el documento; si no puede mostrarlo en la sesión, la afirmación es cero. *Producto: doce afirmaciones calificadas.*

**Día 2 · mañana (3 h) — D4 a D6.** Igual. **Es el bloque incómodo:** contiene las afirmaciones sobre supervisión, publicación y uso de resultados, donde la evidencia agregada dice que casi nadie cumple.

**Día 2 · tarde (2 h) — Perfil, regla de ambición y plan.** Se construye el perfil de seis niveles; se aplica la regla de ambición en su versión 2 (§12.2) a los sistemas en operación; se fija, para la dimensión más baja pertinente, una acción con responsable y fecha. *Producto: perfil, declaración de brechas y plan.*

**Regla de calidad de la autoevaluación.** Si el perfil resultante es sustancialmente superior al perfil mediano medido ( $2 \cdot 2 \cdot 1 \cdot 1 \cdot 0 \cdot 1$ ), la entidad debería someter su evidencia a verificación externa antes de publicarlo. **Un perfil alto sin evidencia verificable es peor que un perfil bajo con plan** — es la regla de comunicación que conviene repetir.

## Anexo C. La regla de ambición: tabla de decisión

Versión 2, reformulada en esta edición (§12.2).

| Situación                                                | Nivel exigido                    | Si la entidad lo alcanza         | Si no lo alcanza                                                                                                                            |
|----------------------------------------------------------|----------------------------------|----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Sistema <b>nuevo</b> de bajo impacto y reversible        | 1 en las dimensiones pertinentes | Aprobar en puerta 1              | <b>No aprobar</b> hasta alcanzarlo                                                                                                          |
| Sistema <b>nuevo</b> de impacto medio, decisión asistida | 2                                | Aprobar en puerta 1              | <b>No aprobar</b> hasta alcanzarlo                                                                                                          |
| Sistema <b>nuevo</b> de alto impacto sobre personas      | 3                                | Aprobar en puerta 1              | <b>No aprobar</b> hasta alcanzarlo                                                                                                          |
| Sistema <b>en operación</b> , cualquier nivel            | El correspondiente a su impacto  | Continuar con revisión ordinaria | <b>Declarar la brecha:</b> qué dimensión falta, qué acción, qué plazo, qué responsable — y <b>publicarlo junto con la ficha del sistema</b> |
| Decisión <b>indelegable</b>                              | —                                | —                                | <b>Ningún nivel habilita.</b> No es cuestión de capacidad                                                                                   |
| <b>Piloto</b> con bitácora y fecha de fin                | Ninguno                          | Permitido en cualquier nivel     | Permitido: la regla limita la operación, no el aprendizaje                                                                                  |

Tabla 12. La regla de ambición, versión 2. Fuente: Enpath.

**Por qué cambió.** La versión 1 era una prohibición sin salida para lo ya operante, y aplicada a entidades reales descalificaba al 100 % de la muestra (§12.2). La versión 2 conserva la prohibición

donde todavía puede ejercerse —lo nuevo— y convierte la brecha de lo existente en un compromiso fechado y público. Lo que se declara es corregible; lo que se prohíbe sin salida se exceptúa.

## Anexo D. Glosario

| Término                       | Definición operativa en este documento                                                                              |
|-------------------------------|---------------------------------------------------------------------------------------------------------------------|
| Madurez institucional en IA   | Capacidad demostrable de una entidad para gobernar, alimentar, dotar, procesar, transparentar y usar sistemas de IA |
| Nivel 4 · institucionalizado  | Capacidad que sobrevive a personas y gobiernos. No alcanzada por ninguna entidad medida                             |
| Regla de ambición             | Criterio que vincula el tipo de sistema que una entidad debería operar con su nivel demostrado                      |
| Dimensión más baja pertinente | La dimensión con menor nivel entre las relevantes para un sistema concreto; es la que manda, no el promedio         |
| Afirmación verificable        | Enunciado binario cuyo cumplimiento se establece con un documento nombrado                                          |
| Perfil                        | El vector de seis niveles de una entidad                                                                            |
| Desbalance                    | Patrón en que D1 y D2 superan a D5 y D6; presente en quince de quince entidades medidas                             |
| Aplicación censal             | Aplicación del instrumento a todas las entidades de un ámbito, para producir la línea de salida agregada            |

## Anexo E. Índice de figuras y tablas

**Figuras** 1. El modelo: seis dimensiones, cinco niveles (§7) 2. Perfil de una entidad hipotética (§8) 2. La regla de ambición (§10) 3. La predicción de febrero de 2025, evaluada (§12.1) 4. La regla de ambición aplicada a quince entidades (§12.2) 5. El techo real del Estado colombiano por dimensión (§12.3) 6. Del instrumento a la serie: el circuito cumplido (§12.4) 7. Matriz de riesgos, revalorada (§14)

**Tablas** 1. La predicción contra la medición (§1.1) 2. La regla de ambición aplicada (§1.2) 2. El techo por dimensión (§1.3) 4. Los instrumentos de la serie y las dimensiones (§4) 5. Lo que ocurrió después de febrero de 2025 (§5.2) 6. El campo de instrumentos y el lugar del modelo (§6) 7. Dimensiones y ejes de la política nacional (§7) 8. La regla de ambición, versión 1 (§10) 9. Las promesas de febrero de 2025, evaluadas (§13) 10. Riesgos, revalorados (§14) 11. Escala de calidad de la evidencia (§16.3) 12. Predicciones falsables (§17) 13. La regla de ambición, versión 2 (Anexo C)

## Referencias

### Fuentes primarias

- Departamento Nacional de Planeación (2025). *Documento CONPES 4144: Política Nacional de Inteligencia Artificial*, aprobado el 14 de febrero de 2025: seis ejes, 106 acciones a 2030, inversión indicativa de 479.273 millones de pesos. <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/4144.pdf> · <https://www.dnp.gov.co/publicaciones/Planeacion/Paginas/con>



pes-4144-hoja-de-ruta-colombia-inteligencia-artificial-retos-actuales-transformacion-futura.aspx

2. National Institute of Standards and Technology (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, enero de 2023. <https://airc.nist.gov/airmf-resources/airmf/> · <https://doi.org/10.6028/NIST.AI.100-1>
3. ISO/IEC (2023). *ISO/IEC 42001:2023 — Information technology, Artificial intelligence, Management system*. <https://www.iso.org/standard/42001>
4. Unión Europea (2024). *Reglamento (UE) 2024/1689*, con aplicación de los artículos 1 a 5 desde el 2 de febrero de 2025. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
5. Unión Europea (2026). *Digital Omnibus on AI*: adopción del Parlamento Europeo el 16 de junio de 2026 y del Consejo el 29 de junio de 2026.

## Mediciones y estudios comparados

1. CENIA y CEPAL (2025). *Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025*, tercera edición, presentada el 3 de octubre de 2025. <https://indichelatam.cl/>
2. Oxford Insights (2024). *Government AI Readiness Index*. <https://oxfordinsights.com/ai-readiness/> — citado como contraste de unidad de análisis (mide países, no entidades).

## Serie de investigación de Enpath

1. Enpath (2024). *Serie de investigación 2024*, doce informes; instrumentos por dimensión en la Tabla 4. Esta biblioteca.
2. Enpath (2025). *Marco de gobernanza de IA para ministerios* (enero); *Scorecard de compra pública de IA* (noviembre, con archivo resultados\_scorecard.csv). Esta biblioteca.
3. Enpath (2026). *Modelo de capacidades de IA para el gobierno: el primer benchmark* (febrero, con archivo benchmark\_2026.csv) — **la aplicación que evalúa la predicción de este modelo**; *Índice Enpath de gobernanza de IA del sector público: Colombia 2026* (marzo, con archivo igia\_2026.csv). Esta biblioteca.

## Notas

---

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

---

1. Departamento Nacional de Planeación de Colombia, *Documento CONPES 4144: Política Nacional de Inteligencia Artificial*, aprobado el 14 de febrero de 2025. Seis ejes estratégicos —ética y gobernanza; datos e infraestructura; investigación, desarrollo e

innovación; capacidades y talento; mitigación de riesgos; uso y adopción—, **106 acciones** con horizonte 2030 e inversión indicativa de **479.273 millones de pesos**.

2. National Institute of Standards and Technology (Estados Unidos), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, enero de 2023. Guía voluntaria para gestionar riesgos a lo largo del ciclo de vida de la IA, con cuatro funciones: gobernar, mapear, medir y gestionar. **Es no certificable**: no existe auditoría contra el marco y la alineación es autodeclarada. No vincula la ambición tecnológica con la capacidad demostrada.
3. ISO/IEC 42001:2023, norma internacional de sistemas de gestión de inteligencia artificial. Establece requisitos y guía para organizaciones que desarrollan, proveen o usan sistemas de IA. **Es certificable** por organismos acreditados, con adopción creciente entre grandes proveedores tecnológicos entre 2024 y 2026. Verifica cumplimiento de requisitos organizacionales; no gradúa capacidad comparable entre entidades ni limita el uso según ella.
4. CENIA y CEPAL, *Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025*, tercera edición, presentada el 3 de octubre de 2025, con más de cien subindicadores. Puntajes de adopción superiores a 60: Brasil 71,15; Perú 67,47; Colombia 67,29; Costa Rica 66,27; Chile 65,79; México 63,27.
5. El 2 de febrero de 2025 entraron en aplicación los artículos 1 a 4 del Reglamento (UE) 2024/1689 y las prohibiciones del artículo 5. El artículo 4 obliga a proveedores y responsables del despliegue a adoptar medidas para garantizar un nivel suficiente de alfabetización en inteligencia artificial de su personal y de quienes operan sistemas de IA en su nombre.
6. *Digital Omnibus on AI* de la Unión Europea: adopción por el Parlamento Europeo el 16 de junio de 2026 y por el Consejo el 29 de junio de 2026. Traslada las obligaciones de los sistemas de alto riesgo del Anexo III al 2 de diciembre de 2027 y las del Anexo I al 2 de agosto de 2028.